International Journal of
*Environmental Research
and Public Health*

IJERPH
15th Anniversary

Edited by
Paul B. Tchounwou

Printed Edition of the Special Issue Published in
International Journal of Environmental Research and Public Health

www.mdpi.com/journal/ijerph

MDPI

IJERPH

IJERPH: 15th Anniversary

Editor

Paul B. Tchounwou

MDPI • Basel • Beijing • Wuhan • Barcelona • Belgrade • Manchester • Tokyo • Cluj • Tianjin

Editor
Paul B. Tchounwou
Department of Biology, College of Science,
Engineering and Technology, Jackson State University
USA

Editorial Office
MDPI
St. Alban-Anlage 66
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *International Journal of Environmental Research and Public Health* (ISSN 1660-4601) (available at: https://www.mdpi.com/journal/ijerph/special_issues/15th-anniversary).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

LastName, A.A.; LastName, B.B.; LastName, C.C. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-03943-713-9 (Hbk)
ISBN 978-3-03943-714-6 (PDF)

Contents

About the Editor

Paul B. Tchounwou is Associate Dean of Graduate and International Programs at the College of Science, Engineering and Technology. He is a Presidential Distinguished Professor who also serves as a Director of DoD-Center of Excellence in STEM Education, Director of NIH-Center for Environmental Health, and Director of Environmental Science PhD Program at Jackson State University (JSU), Jackson, MS. He also serves as Adjunct Professor at Tulane University Health Sciences Center, New Orleans, LA, and Life Member of the scientific faculty of the International Biographical Institute in Cambridge, England.

Dr. Tchounwou is nationally and internationally known for his biomedical research. His work on arsenic trioxide pharmacology and toxicology was recently highlighted in the Fall 2011 issue of *NCRR—Reporter Magazine at the National Institutes of Health*. Dr. Tchounwou has published 178 refereed papers in top-tiered journals and books. He is the Editor-in-Chief of two international journals, including the *International Journal of Environmental Research and Public Health* (MDPI, Basel, Switzerland), and *Environmental Toxicology—An International Journal* (John Wiley & Sons, New York, USA). He was one of the Section Editors of *Encyclopedia of Environmental Health*, published by Elsevier BV in 2011. He served as Guest Editor for the *International Journal of Molecular Sciences*; *Ethnicity and Diseases*; and *Journal of Health Care for the Poor and Underserved*. He is also on the Editorial Board of several other journals, including *Journal of Cancer Science and Therapy*, *Journal of Environmental Biology*, and *Reviews on Environmental Health.* Dr. Tchounwou has served as Session Chair of scientific meetings, and has given over 450 presentations at national and international conferences. This year, he has been selected to serve as Expert Subject Reviewer for the Fulbright Specialist Program.

Preface to "IJERPH: 15th Anniversary"

The International Journal of Environmental Research and Public Health—IJERPH (ISSN 1660-4601) was found in 2004, and has experienced a tremendous amount of growth in terms of the number and quality of scientific publications. With a 2019 Impact Factor of 2.849, IJERPH now ranks among the top international journals in the emerging field of environmental research and public health. As described on our website (https://www.mdpi.com/journal/ijerph), IJERPH is a peer-reviewed journal that focuses on the publication of scientific and technical information on the impacts of natural phenomena and anthropogenic factors on the quality of our environment, the interrelationships between environmental health and the quality of life, as well as the socio-cultural, political, economic, and legal considerations related to environmental stewardship and public health. Its primary areas of research interests include: gene–environment interactions; environmental genomics and proteomics; environmental toxicology, mutagenesis and carcinogenesis; environmental epidemiology and disease control; health risk assessment and management; ecotoxicology, and ecological risk assessment and management; natural resources damage assessment; environmental chemistry and computational modeling; environmental policy and management; environmental engineering and biotechnology; emerging issues in environmental health and diseases; environmental education and public health.

Paul B. Tchounwou
Editor

International Journal of
Environmental Research and Public Health

MDPI

Article

Ethylenediamine-*N*,*N′*-Disuccinic Acid (EDDS)—Enhanced Flushing Optimization for Contaminated Agricultural Soil Remediation and Assessment of Prospective Cu and Zn Transport

Marco Race [1,*], Alberto Ferraro [2], Massimiliano Fabbricino [1], Agostino La Marca [1], Antonio Panico [3], Danilo Spasiano [4], Alice Tognacchini [5,6] and Francesco Pirozzi [1]

1 Department of Civil, Architectural and Environmental Engineering, University of Naples Federico II, via Claudio 21, 80125 Naples, Italy; massimiliano.fabbricino@unina.it (M.F.); agostino.lamarca@libero.it (A.L.M.); francesco.pirozzi@unina.it (F.P.)
2 Department of Civil and Mechanical Engineering, University of Cassino and Southern Lazio, 03043 Cassino, Italy; alberto.ferraro3@gmail.com
3 Telematic University Pegaso, Piazza Trieste e Trento 48, 80132 Naples, Italy; antonio.panico@unipegaso.it
4 Department of Environmental, Building, Civil Engineering and Chemistry, Polytechnic University of Bari, Via Orabona n.4, 70126 Bari, Italy; danilo.spasiano@poliba.it
5 Department of Forest and Soil Sciences, University of Natural Resources and Life Sciences, Vienna, Institute of Soil Research, Konrad-Lorenz-Straße 24, 3430 Tulln, Austria; alice.tognacchini@yahoo.it
6 Alchemia-nova GmbH, Baumgartenstrasse 93, 1140 Vienna, Austria
* Correspondence: marco.race@unina.it; Tel.: +39-081-768-3434

Received: 7 February 2018; Accepted: 15 March 2018; Published: 18 March 2018

Abstract: This paper presents the results of an experimental study aimed at investigating the effect of operative parameters on the efficiency of a soil flushing process, conducted on real contaminated soil containing high amounts of Cu and Zn. Soil flushing tests were carried out with Ethylenediamine-*N*,*N′*-disuccinic acid (EDDS) as a flushing agent due to its high biodegradability and environmentally friendly characteristics. Process parameters such as Empty-Bed Contact Time (EBCT) and EDDS solution molarity were varied from 21–33 h and from 0.36–3.6 mM, respectively. Effects on the mobility of cations such as Fe and Mn were also investigated. Results showed that very high performances can be obtained at [EDDS] = 3.6 mM and EBCT = 33 h. In these conditions, in fact, the amount of removed Cu was 53%, and the amount of removed Zn was 46%. Metal distribution at different depths from the top surface revealed that Cu has higher mobility than Zn. The process results were strongly dependent on the exchange of metals due to the different stability constants of the EDDS complexes. Finally, results from a comparative study showed that soil washing treatment reached the same removal efficiency of the flushing process in a shorter time but required a larger amount of the EDDS solution.

Keywords: metal mobility; soil flushing; soil reclamation; agricultural soil; EDDS deficiency

1. Introduction

Illegal disposal of industrial waste containing different pollutants, both organic and inorganic, is responsible for severe agricultural soil pollution [1]. Although such contamination can be related to the presence of so-called emerging pollutants, such as liquid-ionic pollutants or pesticides [2], in most cases it is related to the presence of potentially toxic elements (PTEs) [3,4]. Among others, the presence of Cu and Zn is quite common, because these metals have a wide application in numerous industrial processes as well as in the production of many pesticides and herbicides [5]. The remediation of

PTEs-contaminated soils can be carried out using different techniques. Among them, the soil washing process proved to be one of the few permanent contaminated soil remediation treatments [6]. Therefore, this process can achieve good remediation performances if the operative conditions are carefully selected and optimized [7–9]. However, in the case of soil washing treatment, the excavation of the contaminated soil is required, and this operation is time-consuming and expensive. This step is then followed by an in-site or on-site treatment by mixing the solid phase with a washing solution able to solubilize the pollutants in the liquid phase. The two phases are successively separated. The remediated soil can be returned to its original place [10] or used for different purposes. The contaminated washing solution, on the other hand, must be appropriately treated [8,11,12] before its final disposal. Unfortunately, this process cannot be applied if the contamination is extended to large or deep areas, because of the excessive costs related to the soil excavation. In these cases, an interesting alternative to soil washing is soil flushing [13,14]. Soil flushing consists of the direct injection in the soil of a leaching solution, avoiding the necessity of contaminated soil excavation.

Several chemical reagents have been positively tested in the past as leaching solutions for soil remediation in the flushing process (i.e., HCl, EDTA, $CaCl_2$) [15,16]. However, part of the flushing solution can be retained in the soil and/or leaches into the groundwater, representing a further source of environmental contamination. Therefore the use of a biodegradable agent, such as Ethylenediamine-*N*,*N*′-disuccinic acid (EDDS) [17,18], may be more advisable in order to decrease the potential risk of negative effects on the environmental quality. While the application of EDDS as a washing agent has been widely studied in several scientific research papers [9,19,20], its use as flushing agent has been tested in limited cases [21,22], and thus requires further investigation in more detail.

The present paper aims to provide a deeper understanding of the applicability of EDDS-enhanced flushing for the remediation of Cu- and Zn-contaminated agricultural soil. The study was conducted at the lab scale to keep all process parameters under control, as well as to analyze the removal efficiency and the extraction kinetic trend. Metals distribution at different depths was also investigated to understand the mechanisms of metal desorption and adsorption during transportation, according to their initial distribution in the different contaminated soil fractions. Moreover, the release of PTEs in the leachate after the soil flushing treatment was evaluated. Finally, results from EDDS-enhanced washing tests conducted at the lab scale on the same soil were reported in order to compare the pollutants removal efficiency and the treatment environmental suitability of the two processes (i.e., soil flushing and soil washing). All of these aspects have been rarely considered in a single experimental work, thus reflecting the novelty of this study.

2. Materials and Methods

2.1. Reagents and Analytical Standards

Hydroxylammonium chloride (reagent grade >98% *w*/*w*), ammonium acetate (>99% *w*/*w*), (S,S)-ethylenediamine-*N*,*N*′-disuccinic acid-trisodium salt solution (35% *v*/*v*), hydrogen peroxide solution (30% *v*/*v*), acetic acid (ACS reagent >97% *v*/*v*), and nitric acid (ACS reagent >67% *v*/*v*) from Sigma-Aldrich (Saint Louis, MI, USA) were used as reagents. Only ultra-pure water was used for analytical preparations and dilution. Atomic Absorption Spectrometer (AAS) standards were employed for Cd, Cr, Cu, Fe, Pb, Zn (Carlo Erba, Reagenti), Mn, and Ni (Fluka Reagents).

2.2. Soil Characterization

The investigated soil was sampled from a land used in the past for agriculture and located in the Litorale Domizio Flegreo and Agro Aversano (Campania Region, Italy) (Figure 1). In particular, this land includes an area of 1076 km^2 (57 municipalities) and has been affected by the illegal disposal of hazardous wastes [23,24]. The soil was sampled manually from the top 20 cm of the soil over an area of about 1 m^2, homogenized, and then stored in hermetic containers. Immediately after collection, the samples were dried at 40 °C in an oven (Argolab, TCN115) and then kept at room temperature.

Figure 1. The sampling point (40°96′05″ N, 14°11′84″ E).

The particle size distribution analysis was performed according to American Society for Testing and Materials (ASTM) method D 422-63 [25]. Only the fraction of soil that was smaller than 2 mm, which was assumed to be the most contaminated [26], was used in all of the tests and analytical determinations. The pH of the soil was evaluated according to the method of Violante ed Adamo [27]. Organic matter was determined through the loss on ignition (LOI) index [28]. Total PTEs concentration was measured in the liquid phase from soil acid mineralization procedure [29]. Before the analysis, the solid and liquid phases from the acid mineralization were separated through centrifugation at 4600 rpm for 20 min using an IEC CENTRA GP8R centrifuge (Needham Heights, MA, USA). Then, the resulting liquid phase was analyzed by atomic adsorption spectrometry using a Varian spectrometer, Model 55 B SpectrAA (F-AAS) (Varian Australia Pty Ltd., Victoria, Australia) equipped with a flame (acetylene/air) and a deuterium background correction. The limit of detection (LOD) values for each of the analyzed elements were 5×10^{-2} mg·L^{-1} for Cd and Zn, 0.5 mg·L^{-1} for Cr, 0.2 mg·L^{-1} for Cu, 0.25 mg·L^{-1} for Fe, 0.1 mg·L^{-1} for Mn, 0.3 mg·L^{-1} for Ni, and 1 mg·L^{-1} for Pb. A sequential extraction procedure proposed by the Community Bureau of Reference (BCR) [30,31] was performed on the contaminated soil samples to determine the metal partition among different soil fractions. This procedure was based on an initial extraction in 40 mL of acetic acid (0.11 M) (step 1—exchangeable and weak acid soluble fraction). Afterward, a volume of 40 mL of hydroxylammonium chloride solution (0.5 M) was added to the residual soil from "step 1" and acidified by the addition of a 2 M HNO_3 solution (step 2—reducible fraction). Successively, 20 mL of hydrogen peroxide (8.8 M) and 50 mL of ammonium acetate (1 M) were used for oxidizing the soil (step 3—oxidizable fraction). The last step consisted in a soil acid mineralization (step 4—residual fraction).

2.3. Soil Flushing Lab Scale Tests

Soil flushing tests were performed following the indication of Hauser et al. [21]. In particular, the columns were prepared using two polypropylene "falcon" test tubes characterized by a diameter and length of 30 and 80 mm, respectively. Tests were carried out by filling each tube with 40 g of the investigated soil. Soil was packed in the columns until no reduction of volume was observed, according to Pontoni et al. [32]. Subsequently, a total bed volume (bv) of 50 mL was obtained in each tube.

During the first series of tests, two EDDS solution molarities ([EDDS]) were tested (0.36 and 3.6 mM). The flushing solution was leached through the columns using IVEK rotary pumps (IVEK Corporation, North Springfield, Vermont) at three different percolation velocities (1.7, 2.1, and 2.7 mL·h^{-1}), which corresponded to three Empty-Bed Contact Times (EBCTs), i.e., 33, 27, and 21 h. The total treatment lasted 20 days. The solution collected at the bottom of the column was sampled at different times of the test and analyzed to measure the concentration of the following metals: Cu, Zn, Fe and Mn.

In order to evaluate metals mobility in the soil column, other tests were performed for selected experimental conditions ([EDDS] = 3.6 mM and EBCT = 33 h). After the flushing process, the soil column was divided into four sections, and the metal concentration was measured in each of them (Soil Flushing Test 1, SF1). Furthermore, sequential extraction was carried out on the soil contained in each layer to evaluate the changes of PTEs fractionation after the treatment. A comparative test using only pure water as a flushing agent was also performed. A second series of tests (Soil Flushing Test 2, SF2) was performed to evaluate metal transportation along the columns. For this purpose, before the flushing process, the columns were divided into four layers, each of them containing 10 g of the contaminated soil. The flushing solution was sampled at the bottom of each layer and analyzed for PTEs and main cations concentration. Flushing operative conditions were the same as also adopted for the previous test. Finally, a third series of tests (Soil Flushing Test 3, SF3) was conducted to study the release of PTEs in the leachate after the soil flushing treatment (i.e., release due to the rain phenomenon). In this case, only 4 bv of the EDDS solution was injected through the columns at the beginning of the test, and then only pure water was added as a flushing agent. The flushing solutions were collected at the bottom of the columns and analyzed for PTEs evaluation. Also in this case, flushing operative conditions set for SF1 and SF2 were adopted.

2.4. Soil Washing Lab Scale Tests

Comparative soil washing tests were conducted on the same soil, at the lab scale, in 50-mL plastic reactors. Two different values of the liquid to solid ratio (LSR) were chosen, namely 5:1 and 10:1 (v/w). Soil washing parameters were optimized on the basis of the results of previous studies [24,33] conducted on the same soil. In particular, an extracting solution of [EDDS] = 0.36 mM was adopted to study the process efficiency in EDDS deficiency conditions, whereas an extracting solution of [EDDS] = 3.6 mM was used to achieve the best process performance. Two reaction times (i.e., 48 and 96 h) were selected since a previous work [33] proved these times to be efficient for achieving the PTEs equilibrium conditions. All tests were conducted in triplicate to reduce experimental errors. The metal concentration in the exhausted solution was measured by atomic adsorption spectrometry following the procedure described in Section 2.2. One-way and two-way analysis of variance (ANOVA) were used to analyze the statistical differences among treatments. Comparisons were made with the post-hoc Tukey's Honestly Significant Difference HSD test. Statistical significance was assumed at $p < 0.05$. Statistical analyses were conducted in Microsoft® Excel 2013/XLSTAT©-Pro (Version 7.2, 2003, Addinsoft, Inc., Brooklyn, NY, USA) and GraphPad Prism 6.0 (GraphPad Software, San Diego, CA, USA).

3. Results and Discussion

3.1. Soil Characterization

The initial pH of the soil was close to neutral value (7.21) and the LOI index was 7.01%. The study of the particle size distribution resulted in classifying the soil as a loam, confirming its suitability for plant growth and agricultural activities. Among all of the measured PTEs, only Cu and Zn exceeded the threshold values (TVs) established by the European Regulation (Table 1) [34].

Table 1. Soil characterization.

PTEs	Cd	Cr	Cu	Fe	Mn	Ni	Pb	Zn
Soil (mg·kg^{-1})	0.13 ± 0.01	23.0 ± 1.2	296 ± 10	19842 ± 1230	913.9 ± 31.1	11.9 ± 2.1	21.9 ± 0.2	277 ± 3.0
Threshold values (mg·kg^{-1}) [34]	1	100	100	-	-	50	60	200

The PTEs reported in Table 1 are very common soil pollutants as they can be found in some pesticides and fertilizers [35]. High concentrations of Cu and Zn can represent a serious risk for human health since both metals could become toxic [36], making soil treatment necessary. Results of the sequential extraction showed that the percentage of Cu and Zn in the bioavailable fraction (sum of the first three fractions) was 81% and 70%, respectively (Figure 5). These values further confirmed the risk of Cu and Zn migration from the soil to agricultural products.

3.2. Optimization of the Soil Flushing Process

Figures 2 and 3 report the results of Cu, Fe, Mn, and Zn removal efficiency obtained during the flushing tests. Results from Figure 2 displayed that the EBCT influenced the flushing process only at high EDDS concentrations, resulting in a different removal efficiency. These differences were accentuated mainly for the lower values of bv (Figure 2). Actually, the slopes of the curves related to Cu and Zn removal were equal when the bv was higher than 2. On the contrary, no effects on the removal efficiency at different treatment times were found for tests characterized by significantly high EDDS concentrations (Figure 3). In this case, in fact, it was possible to achieve significant PTEs removal efficiency for all of the investigated treatment times. At a high EDDS concentration (3.6 mM), Cu and Zn extraction rates were characterized by a fast kinetic only in the first period of the treatment, whereas in the remaining time the kinetics were significantly slower. This observed effect was likely due to the following two phenomena: (i) an immediate extraction of the two metals from the carbonate fraction where they are in the di-valent form, i.e., Cu^{2+} and Zn^{2+}; (ii) a subsequent oxidation of the metals from the mono- and zero- to the di-valent form, as a consequence of the humic acid reduction [33,37]. Actually, metals in zero-valent form cannot be complexed as Me-EDDS [33]. On the contrary, the extraction rates of Fe and Mn increased during the overall treatment time. In detail, a lower removal was observed at the beginning of the flushing treatment followed by a gradual increase at the end of the process. Such a result was attributed to the higher affinity of EDDS for Cu and Zn compared to Fe and Mn, which mainly affected the removal efficiencies at the beginning of the flushing treatment. Subsequently, the effect of the high concentration of EDDS prevailed on the previous factors, resulting in a relevant Fe and Mn removal efficiency. Moreover, the pattern was also related to the occurrence of different metals mobilization depending on the redox potential conditions of the soil. In fact, it is generally reported that Cu and Zn have a higher mobility rate under oxidizing conditions while Fe and Mn are characterized by higher mobility at lower values of soil oxidizing potential [38]. During the tests, after the first few days the soil in the column became water-saturated. This caused a reduction of the oxidizing potential and the conversion of Fe and Mn to their more leachable reduced forms [39].

During tests conducted with [EDDS] = 0.36 mM solution at different EBCTs, the extraction process was not influenced by the different values of bv. Results in Figure 2, in fact, show an overlap of the extraction curves.

This result was due to the occurrence of a total EDDS complexation with metals, as a consequence of the low concentration of EDDS in the solution. Hence, the EDDS concentration represented a limiting factor of the treatment velocity. As the EDDS was fully complexed, the metal exchange mechanisms acquired a higher importance in the process. Indeed, in all tests with lower EDDS concentration (i.e., [EDDS] = 0.36 mM), Cu was the only metal with an excellent initial extraction rate. Results showed that for bv values up to 4 the amount of leached Cu was 77 $mg \cdot kg^{-1}$, corresponding to approximately 49 ± 5 μmol (Figure 2). This value was higher than 70% of the corresponding EDDS moles injected (72 μmol). Such a result could be ascribable to the adsorption of a small amount of EDDS onto the soil and the formation of Cu-EDDS complexes with the remaining EDDS [40]. Then, Zn-EDDS complexes started to form only for bv values higher than 4, as a result of the different values of the stability constant (Kst) of the two PTEs-EDDS. In fact, the Kst of Cu-EDDS is higher than that of Zn-EDDS [41]. For fixed values of treatment time, the EBCT affected the extraction rate of metals since lower EBCT values corresponded to a higher amount of the injected EDDS solution, and resulted in higher metal extraction rates (Figure 3). As the EDDS was totally complexed by Cu and Zn, the cations Fe and Mn could not form complexes with the flushing agent, and therefore could not be extracted. The pH and the LOI index of the soil were evaluated at the end of the tests. The soil pH displayed a value equal to 7.05 ± 0.2, assessing for the investigated soil buffering capacity in minimizing pH changes [42]. Likewise, the LOI value was observed at the end of the test confirming no soil characteristic alteration due to the EDDS involvement.

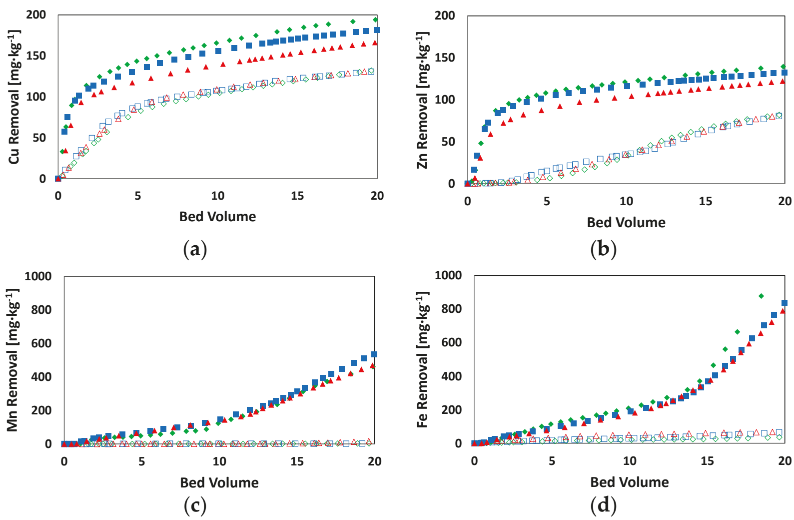


Figure 2. Breakthrough curves at different Empty-Bed Contact Times (EBCTs) with respect to bed volume. [EDDS] = 3.6 mM: ◆ EBCT = 33 h; ■ EBCT = 27 h; ▲ EBCT = 21 h; [EDDS] = 0.36 mM: ◇ EBCT = 33 h; □ EBCT = 27 h; △ EBCT = 21 h—(**a**) Cu; (**b**) Zn; (**c**) Mn; (**d**) Fe.

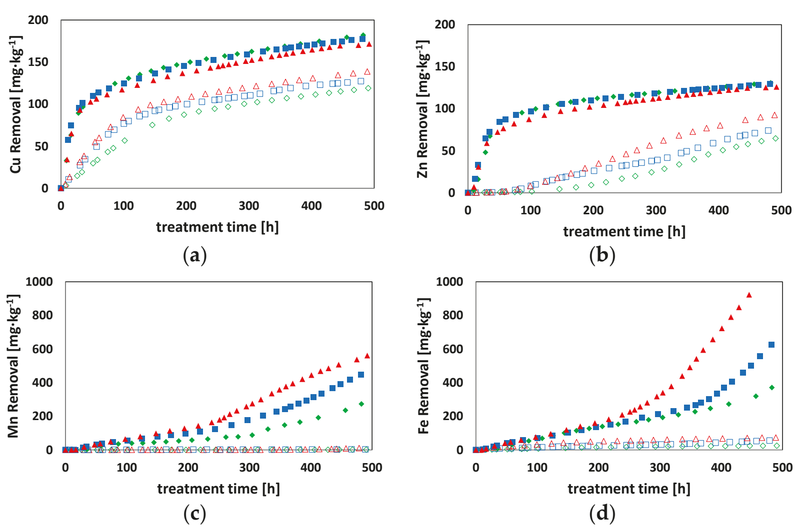


Figure 3. Breakthrough curves at different EBCTs with respect to treatment time. [EDDS] = 3.6 mM: ◆ EBCT = 33 h; ■ EBCT = 27 h; ▲ EBCT = 21 h; [EDDS] = 0.36 mM: ◇ EBCT = 33 h; □ EBCT = 27 h; △ EBCT = 21 h—(**a**) Cu; (**b**) Zn; (**c**) Mn; (**d**) Fe.

3.3. Soil Washing Process and Comparison of the Removal Efficiencies

The results obtained from the soil washing tests, with constant LSR and different EDDS concentrations, showed different removal trends for Cu and Zn (Figure 4). In all tests, a substantial increase in the removal efficiency occurred in the first 48 h, followed by a non-significant increase in the removal efficiency at 96 h. This was also confirmed by statistical analysis since no statistically significant differences between the removal efficiency at 48 and 96 h were observed ($p > 0.05$). These results were in accordance with previous investigations [33,43].

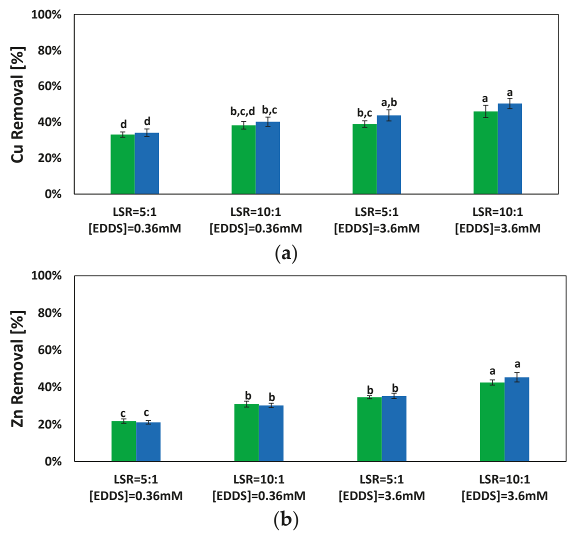


Figure 4. Cu and Zn removal by soil washing treatment at different times. 48 h; 96 h—(**a**) Cu; (**b**) Zn.

Results of tests conducted with varying the LSR showed significant differences in terms of the removal efficiency, especially for Zn, as also observed from the statistically significant differences ($p < 0.05$). It is worth noting that higher LSR values, achieved by increasing the liquid phase and keeping constant the EDDS solution molarity, induced a consequent increase of EDDS moles in the solution. Such a trend may be attributed to the following processes: (i) the presence of a certain amount of free EDDS non-complexed with metals; and (ii) the occurrence of metal exchange phenomena among the PTEs-EDDS complexes, which promotes the formation of Cu-EDDS or Zn-EDDS complexes [44].

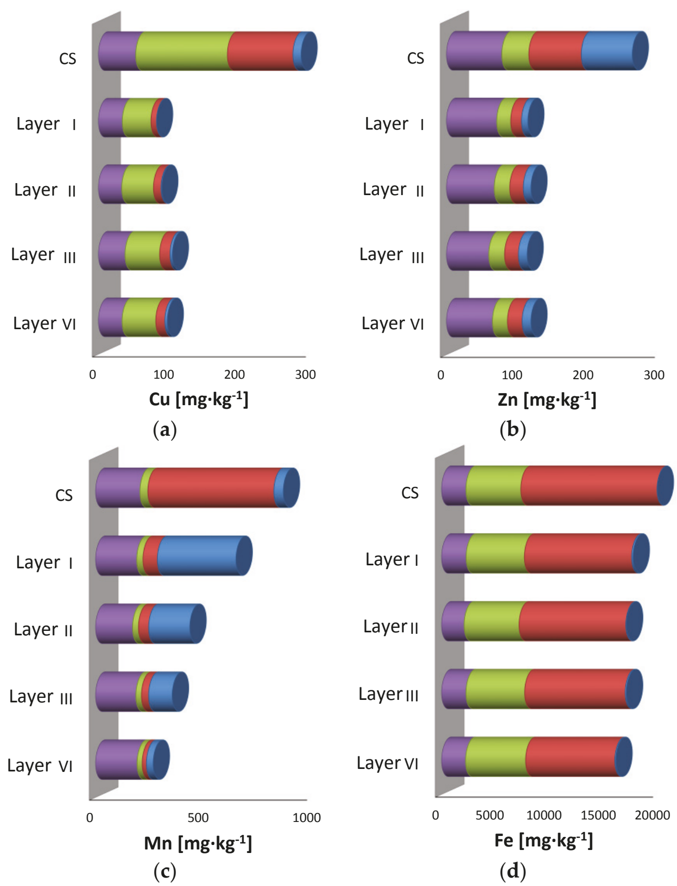
In agreement with Tsang et al. [45], tests conducted with an EDDS concentration deficiency led to Fe and Mn removal efficiencies lower than 1%. At higher EDDS concentration, Fe and Mn removals slightly increased, but the overall removed amount of these elements was negligible compared to the initial total amount in the soil [46].

Comparing the results of the soil flushing and soil washing tests, it could be concluded that a significantly higher Cu removal efficiency was obtained with the soil flushing treatment considering the same amount of EDDS solution used. On the contrary, only a slightly higher Cu removal efficiency was achieved with the soil flushing treatment compared to the soil washing when the same treatment time was considered (Figure S1).

3.4. Fractionation of PTEs/Main Cations in Different Soil Layers after Leaching

Interesting results were obtained from the sequential extraction procedure by comparing pre- and post-treatment metal distributions (Figure 5). Cu was initially bound to the organic substance and metal oxides/hydroxides complexes, while the amount of ions in the cation exchange sites was not relevant. A higher removal efficiency was observed for the Cu fraction bound to the organic substance and absorbed onto the metal oxides, as shown by the values of Cu extracted in the second and third steps. As regards Zn, instead, the highest removal efficiency was observed in the exchangeable and reducible fractions (steps 1 and 2). Small amounts of Cu and Zn were still found in the exchangeable and weak acid soluble fractions (step 1) after the treatment. This latter result was attributed to the EDDS amount adsorbed onto the soils after leaching [40,47].

Finally, Fe was extracted almost exclusively from the reducible fraction (step 2), while Mn was mobilized mainly from the reducible fraction (step 2) and partially deposited on the cation exchange sites along the soil column. The results showed (Figure 5) that the Cu and Zn concentrations at the end of treatment were lower than the threshold levels (Table 1) and their removal occurred mainly from the bioavailable fraction. This treatment displayed interesting outcomes since it allowed the reduction of the leaching of PTEs. In fact, the main removal occurred from the acid soluble fraction of the soil, which is generally characterized by fast metals mobilization [48] and represents a serious contamination risk for the environment [49].



(**a**) (**b**) (**c**) (**d**)

Figure 5. Values of metals in the four steps of the sequential extraction on the contaminated soil (CS) and on the four layers of the treated soil before and after SF1 tests. ■ Exchangeable and weak acid soluble fraction, ■ Reducible fraction, ■ Oxidizable fraction; ■ Residual Fraction (**a**) Cu; (**b**) Zn; (**c**) Mn; (**d**) Fe.

3.5. Metals Transportation through the Columns

Figure 6 reported the results of metal profiles at different column layers obtained from SF2. Data were plotted at different bv values corresponding to different amounts of injected flushing solution. Metal concentration in the soil was evaluated by mass balance through the measure of their concentration in the spent solutions sampled at the bottom of each layer.

The various metals exhibited different behaviors due to the occurrence of metal exchange phenomena or metal-EDDS complex adsorption. These phenomena were more significant for Zn than Cu.

The Cu concentration initially increased with the depth, therefore the concentration measured in the deepest layer was higher than that in the top layer. Nonetheless, after the injection of the first 0.25 bv of the flushing solution, the gradient tended to be gradually less pronounced, along with the reduction of the residual concentration in each layer. The same trend was observed for Zn only after the injection of the first 0.5 bv of the flushing solution. For this bv value, the distribution of Zn concentration with the depth was not monotonic at the beginning of the treatment, due to Zn release from the top layer and its adsorption in the following two layers. Upon increasing the bv values, the removal of Zn also occurred in the second layer but at a lower extent than in the top layer, and the released metal amount was adsorbed in the two successive layers. Finally, Zn removal also occurred in the third layer after 1 bv, and in the last layer after 2 bv injection. At the end, the Zn distribution was uniform, as was the Cu distribution.

The obtained distribution of Cu with the depth was in contrast with the findings of Hauser et al. [21]. This result was ascribable to the different operative conditions set for the flushing tests. In the work of Hauser et al. [21], leaching tests were carried out with a non-continuous flow rate and non-water saturated soil. In such conditions the reduction from Cu^{2+} to elemental Cu occurred and the complex with EDDS did not form [33].

On the other hand, the distribution of Zn with respect to the depth was similar to that obtained in previous studies [21,50]. This result was attributed to the non-occurrence of Zn^{2+} reduction to elemental Zn. In fact, it is reported [50] that Zn can be released as Zn^{2+} and $Zn(OH)^{+}$ according to the following reactions (r_1–r_2):

r_1) $Zn^{2+} + HEDDS^{3-} \leftrightarrow ZnEDDS^{2-} + H^{+}$
r_2) $Zn(OH)^{+} + HEDDS^{3-} \leftrightarrow ZnEDDS^{2-} + H_2O$

and H^{+} formation can cause ZnO dissolution (r_3–r_3):

r_3) $ZnO + 2H^{+} \leftrightarrow Zn^{2+} + H_2O$
r_4) $ZnO + H^{+} \leftrightarrow ZnOH^{+}$

Moreover, Zn oxy-hydroxides can be chelated from EDDS as follows (r_5):

r_5) $ZnO + HEDDS^{3-} \leftrightarrow ZnEDDS^{2-} + OH^{-}$.

As regards Fe and Mn distribution throughout the column depth, it was confirmed that the removal took place starting from 2 bv. A uniform removal profile was observed along the depth during the first period of the treatment. After 10 bv a relevant removal of Fe and Mn occurred in the top layer in addition to the release of the metals from the successive layers. However, the amount of Fe and Mn removed was negligible compared to their initial concentration in the soil. This confirmed that pollutants can be removed without damaging the soil if an appropriate technique is selected.

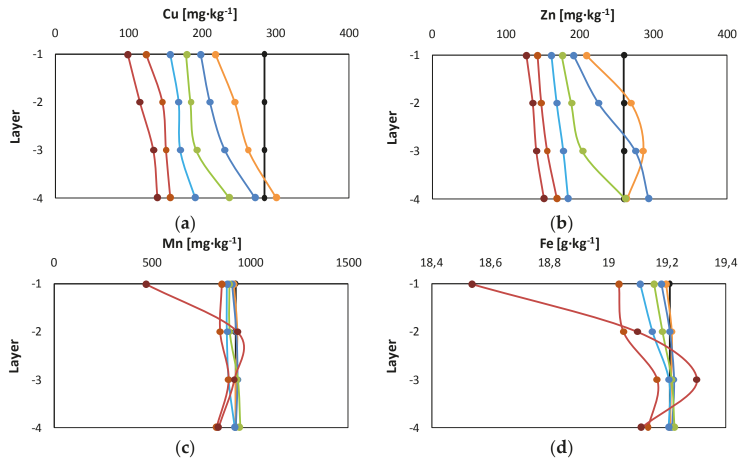


Figure 6. Cu, Zn, Mn, and Fe transport across soil column layers with respect to different values of bv: 0 bv; 0.25 bv; 0.5 bv; 1 bv; 2 bv; 5 bv; 19 bv—(**a**) Cu; (**b**) Zn; (**c**) Mn; (**d**) Fe.

3.6. Release of PTEs in the Water after the Treatment

Results from the SF3 tests are reported in Figure 7. As can be seen, there was no release of Cu, Zn, or Mn in the water once the injection of EDDS was stopped (i.e., after 4 bv of solution injection), confirming the effectiveness of the treatment. This indicates that the PTEs were likely bound in forms that exhibit a low potential release in the environment and low bioavailability for the living organisms [51]. The only exception was represented by the first 0.5 bv of removal (4–4.5 bv), characterized by the presence of a low metals concentration, due to the retention of a certain amount of EDDS in the column. This occurrence led to the recommendation of soil flushing with water as a final step of the remediation treatment. Conversely, the release of Fe persisted all along the treatment until the same removal efficiency was achieved, as observed in the control test.

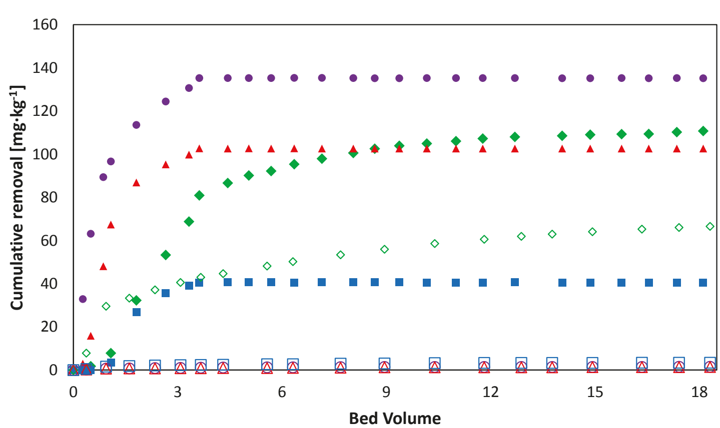


Figure 7. Cu, Zn, Fe, and Zn cumulative removal from the soil as a function of bv, during the SF3 tests. Full symbols: 0–4 bv [EDDS] = 3.6 mM, 4–18 bv [EDDS] = 0 mM—● Cu; ▲ Zn; ■ Mn; ✚ Fe. Empty symbols: 0–18 bv [EDDS] = 0 mM. ○ Cu; △ Zn; □ Mn; ◇ Fe.

4. Conclusions

The present study assessed the feasibility of performing soil flushing treatment with EDDS solution for remediating natural soils contaminated by Cu and Zn. An almost total removal of the bio-available fractions of these metals was achieved. Although the main factor that influenced the metals removal efficiency was the EDDS concentration in the extracting solution, this study proved that the effectiveness of the soil flushing was also dependent on several other factors, including contaminated soil depth as well as the different PTEs-EDDS complexes affinity. Moreover, tests under EDDS-deficiency conditions have further demonstrated that the Cu and Zn removal efficiency was lower due to their re-adsorption onto the soil and/or because of the occurrence of metal exchange processes. In order to take into account all of these aspects, it is essential to carry out bench-scale tests before performing a full-scale soil remediation treatment. Finally, the study proved that soil flushing, besides having a lower environmental impact and lower operative costs compared to soil washing (e.g., no soil excavation is required), is capable of achieving the same high removal efficiency, producing a smaller volume of contaminated spent solution and consuming a minor amount of chemical agents.

Supplementary Materials: The following are available online at http://www.mdpi.com/1660-4601/15/3/543/s1, Figure S1: Cu and Zn removal by soil flushing treatment at different times. ■ 48 h; ■ 96 h and at different values of bv—(a) Cu 4 bv; (b) Zn 4bv; (c) Cu 8 bv; (d) Zn 8 bv.

Acknowledgments: This study was carried out within the EUProject LIFE11 ENV/IT/000275 (ECOREMED).

Author Contributions: Marco Race and Alberto Ferraro performed all the tests, conducted all the analyses, and contributed to writing the manuscript. Massimiliano Fabbricino designed the study and participated in performing the tests as well as in writing the manuscript up to its final version. Antonio Panico supervised the whole work, edited the manuscript, and revised it critically. Alice Tognacchini participated in writing the manuscript and conducting analyses. Danilo Spasiano participated in performing the chemical analyses and edited the manuscript. Francesco Pirozzi participated in designing the study as well as in writing the manuscript and revising it critically.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dahlin, S.; Witter, E.; Mårtensson, A.; Turner, A.; Bååth, E. Where's the limit? Changes in the microbiological properties of agricultural soils at low levels of metal contamination. *Soil Biol. Biochem.* **1997**, *29*, 1405–1415. [CrossRef]
2. Spasiano, D.; Siciliano, A.; Race, M.; Marotta, R.; Guida, M.; Andreozzi, R.; Pirozzi, F. Biodegradation, ecotoxicity and UV_{254}/H_2O_2 treatment of imidazole, 1-methyl-imidazole and *N*,*N*′-alkyl-imidazolium chlorides in water. *Water Res.* **2016**, *106*. [CrossRef] [PubMed]
3. Gurung, B.; Race, M.; Fabbricino, M.; Komínková, D.; Libralato, G.; Siciliano, A.; Guida, M. Assessment of metal pollution in the Lambro Creek (Italy). *Ecotoxicol. Environ. Saf.* **2018**, *148*. [CrossRef] [PubMed]
4. Liang, Y.; Yi, X.; Dang, Z.; Wang, Q.; Luo, H.; Tang, J. Heavy Metal Contamination and Health Risk Assessment in the Vicinity of a Tailing Pond in Guangdong, China. *Int. J. Environ. Res. Public Health* **2017**, *14*, 1557. [CrossRef] [PubMed]
5. Meers, E.; Ruttens, A.; Hopgood, M.J.; Samson, D.; Tack, F.M.G. Comparison of EDTA and EDDS as potential soil amendments for enhanced phytoextraction of heavy metals. *Chemosphere* **2005**, *58*, 1011–1022. [CrossRef] [PubMed]
6. Dermont, G.; Bergeron, M.; Mercier, G.; Richer-Laflèche, M. Soil washing for metal removal: A review of physical/chemical technologies and field applications. *J. Hazard. Mater.* **2008**, *152*, 1–31. [CrossRef] [PubMed]
7. Ferraro, A.; Fabbricino, M.; van Hullebusch, E.D.; Esposito, G.; Pirozzi, F. Effect of soil/contamination characteristics and process operational conditions on aminopolycarboxylates enhanced soil washing for heavy metals removal: A review. *Rev. Environ. Sci. Biol./Technol.* **2015**, 1–35. [CrossRef]
8. Lestan, D.; Luo, C.; Li, X. The use of chelating agents in the remediation of metal-contaminated soils: A review. *Environ. Pollut.* **2008**, *153*, 3–13. [CrossRef] [PubMed]
9. Satyro, S.; Race, M.; Marotta, R.; Dezotti, M.; Guida, M.; Clarizia, L. Photocatalytic processes assisted by artificial solar light for soil washing effluent treatment. *Environ. Sci. Pollut. Res.* **2017**, *24*. [CrossRef] [PubMed]

10. Komínková, D.; Fabbricino, M.; Gurung, B.; Race, M.; Tritto, C.; Ponzo, A. Sequential application of soil washing and phytoremediation in the land of fires. *J. Environ. Manag.* **2018**, *206*. [CrossRef]
11. Onotri, L.; Race, M.; Clarizia, L.; Guida, M.; Alfè, M.; Andreozzi, R.; Marotta, R. Solar photocatalytic processes for treatment of soil washing wastewater. *Chem. Eng. J.* **2017**, *318*. [CrossRef]
12. D'Acunto, B.; Esposito, G.; Frunzo, L.; Mattei, M.R.; Pirozzi, F. Mathematical modeling of heavy metal biosorption in multispecies biofilms. *J. Environ. Eng.* **2015**, *142*, C4015020. [CrossRef]
13. Yang, L.; Jiang, L.; Wang, G.; Chen, Y.; Shen, Z.; Luo, C. Assessment of amendments for the immobilization of Cu in soils containing EDDS leachates. *Environ. Sci. Pollut. Res.* **2015**, *22*, 16525–16534. [CrossRef] [PubMed]
14. Fedje, K.K.; Modin, O.; Strömvall, A.-M. Copper recovery from polluted soils using acidic washing and bioelectrochemical systems. *Metals* **2015**, *5*, 1328–1348. [CrossRef]
15. Reed, B.; Carriere, P.; Moore, R. Flushing of a Pb(II) Contaminated Soil Using HCl, EDTA, and $CaCl_2$. *J. Environ. Eng.* **1996**, *122*, 48–50. [CrossRef]
16. Yun, S.M.; Kang, C.S.; Kim, J.; Kim, H.S. Evaluation of soil flushing of complex contaminated soil: An experimental and modeling simulation study. *J. Hazard. Mater.* **2015**, *287*, 429–437. [CrossRef] [PubMed]
17. Vandevivere, P.; Hammes, F.; Verstraete, W.; Feijtel, T.; Schowanek, D. Metal Decontamination of Soil, Sediment, and Sewage Sludge by Means of Transition Metal Chelant [S,S]-EDDS. *J. Environ. Eng.* **2001**, *127*, 802–811. [CrossRef]
18. Race, M. Applicability of alkaline precipitation for the recovery of EDDS spent solution. *J. Environ. Manag.* **2017**, *203*. [CrossRef] [PubMed]
19. Ferraro, A.; Fabbricino, M.; van Hullebusch, E.D.; Esposito, G. Calibration and validation of a two-step kinetic mathematical model for predicting Cu extraction efficiency in an EDDS-enhanced soil washing. *Water Air Soil Pollut.* **2016**, *227*, 71. [CrossRef]
20. Yan, D.Y.S.; Yip, T.C.M.; Yui, M.M.T.; Tsang, D.C.W.; Lo, I.M.C. Influence of EDDS-to-metal molar ratio, solution pH, and soil-to-solution ratio on metal extraction under EDDS deficiency. *J. Hazard. Mater.* **2010**, *178*, 890–894. [CrossRef] [PubMed]
21. Hauser, L.; Tandy, S.; Schulin, R.; Nowack, B. Column extraction of heavy metals from soils using the biodegradable chelating agent EDDS. *Environ. Sci. Technol.* **2005**, *39*, 6819–6824. [CrossRef] [PubMed]
22. Lo, I.M.C.; Tsang, D.C.W.; Yip, T.C.M.; Wang, F.; Zhang, W. Significance of metal exchange in EDDS-flushing column experiments. *Chemosphere* **2011**, *83*, 7–13. [CrossRef] [PubMed]
23. Greyl, L.; Vegni, S.; Natalicchio, M.; Cure, S.; Ferretti, J. The waste crisis in Campania. In *Ecological Economics from the Ground Up*; Healy, H., Martinez-Alier, J., Temper, L., Walter, M., Gerber, J.F., Eds.; Routledge: London, UK, 2013; pp. 273–308.
24. Satyro, S.; Race, M.; Marotta, R.; Dezotti, M.; Spasiano, D.; Mancini, G.; Fabbricino, M. Simulated solar photocatalytic processes for the simultaneous removal of EDDS, Cu(II), Fe(III) and Zn(II) in synthetic and real contaminated soil washing solutions. *J. Environ. Chem. Eng.* **2014**, *2*. [CrossRef]
25. American Society for Testing and Materials. ASTM D422-63, Standard test method for particle-size analysis of soils. In *Annual Book of ASTM Standards*; American Society for Testing and Materials: West Conshohocken, PA, USA, 2003; Volume 4–8, pp. 10–17.
26. Gallego, J.L.R.; Ordóñez, A.; Loredo, J. Investigation of trace element sources from an industrialized area (Avilés, northern Spain) using multivariate statistical methods. *Environ. Int.* **2002**, *27*, 589–596. [CrossRef]
27. Violante, P.; Adamo, P. pH determination. In *Official Methods of Soil Chemical Analysis*; Francoangeli: Rome, Italy, 2000; pp. 10–13.
28. Schulte, E.E. Recommended soil organic matter tests. *Recomm. Soil Test. Proced. North East. USA Northeast. Reg. Publ.* **1995**, 52–60.
29. *USEPA Method 3051, Microwave Assisted Digestion of Sediments, Sludges, Soils and Oils, Official Methods/US EPA Methods*; Environmental Protection Agency: Washington, DC, USA, 2001.
30. Pueyo, M.; Mateu, J.; Rigol, A.; Vidal, M.; López-Sánchez, J.F.; Rauret, G. Use of the modified BCR three-step sequential extraction procedure for the study of trace element dynamics in contaminated soils. *Environ. Pollut.* **2008**, *152*, 330–341. [CrossRef] [PubMed]
31. Race, M.; Nábělková, J.; Fabbricino, M.; Pirozzi, F.; Raia, P. Analysis of Heavy Metal Sources for Urban Creeks in the Czech Republic. *Water Air Soil Pollut.* **2015**, *226*, 322. [CrossRef]

32. Pontoni, L.; van Hullebusch, E.D.; Fabbricino, M.; Esposito, G.; Pirozzi, F. Assessment of trace heavy metals dynamics during the interaction of aqueous solutions with the artificial OECD soil: Evaluation of the effect of soil organic matter content and colloidal mobilization. *Chemosphere* **2016**, *163*, 382–391. [CrossRef] [PubMed]
33. Race, M.; Marotta, R.; Fabbricino, M.; Pirozzi, F.; Andreozzi, R.; Cortese, L.; Giudicianni, P. Copper and zinc removal from contaminated soils through soil washing process using ethylenediaminedisuccinic acid as a chelating agent: A modeling investigation. *J. Environ. Chem. Eng.* **2016**, *4*, 2878–2891. [CrossRef]
34. Tóth, G.; Hermann, T.; Da Silva, M.R.; Montanarella, L. Heavy metals in agricultural soils of the European Union with implications for food safety. *Environ. Int.* **2016**, *88*, 299–309. [CrossRef] [PubMed]
35. Gall, J.E.; Boyd, R.S.; Rajakaruna, N. Transfer of heavy metals through terrestrial food webs: A review. *Environ. Monit. Assess.* **2015**, *187*, 201. [CrossRef] [PubMed]
36. Zhang, D.; Wang, X.; Zhou, Z. Impacts of Small-Scale Industrialized Swine Farming on Local Soil, Water and Crop Qualities in a Hilly Red Soil Region of Subtropical China. *Int. J. Environ. Res. Public Health* **2017**, *14*, 1524. [CrossRef] [PubMed]
37. Maurer, F.; Christl, I.; Fulda, B.; Voegelin, A.; Kretzschmar, R. Copper redox transformation and complexation by reduced and oxidized soil humic acid. 2. Potentiometric titrations and dialysis cell experiments. *Environ. Sci. Technol.* **2013**, *47*, 10912–10921. [CrossRef] [PubMed]
38. Charlatchka, R.; Cambier, P. Influence of reducing conditions on solubility of trace metals in contaminated soils. *Water. Air. Soil Pollut.* **2000**, *118*, 143–168. [CrossRef]
39. Colombo, C.; Palumbo, G.; He, J.-Z.; Pinton, R.; Cesco, S. Review on iron availability in soil: Interaction of Fe minerals, plants, and microbes. *J. Soils Sediments* **2014**, *14*, 538–548. [CrossRef]
40. Yip, T.C.M.; Tsang, D.C.W.; Ng, K.T.W.; Lo, I.M.C. Kinetic interactions of EDDS with soils. 1. Metal resorption and competition under EDDS deficiency. *Environ. Sci. Technol.* **2009**, *43*, 831–836. [CrossRef] [PubMed]
41. Orama, M.; Hyvönen, H.; Saarinen, H.; Akselab, R. Complexation of [S,S] and mixed stereoisomers of *N*,*N*′-ethylenediaminedisuccinic acid (EDDS) with Fe(III), Cu(II), Zn(II) and Mn(II) ions in aqueous solution. *J. Chem. Soc. Dalt. Trans.* **2002**, 4644–4648. [CrossRef]
42. McLean, E.O. Soil pH and lime requirement. In *Methods Soil Analysis Part 2, Chemical Microbiological Properties*; ACSESS: Madison, WI, USA, 1982; pp. 199–224.
43. Yang, R.; Luo, C.; Zhang, G.; Li, X.; Shen, Z. Extraction of heavy metals from e-waste contaminated soils using EDDS. *J. Environ. Sci.* **2012**, *24*, 1985–1994. [CrossRef]
44. Ferraro, A.; Fabbricino, M.; van Hullebusch, E.D.; Esposito, G. Investigation of different ethylenediamine-N, N′-disuccinic acid-enhanced washing configurations for remediation of a Cu-contaminated soil: Process kinetics and efficiency comparison between single-stage and multi-stage configurations. *Environ. Sci. Pollut. Res.* **2017**, *24*, 21960–21972. [CrossRef] [PubMed]
45. Tsang, D.C.W.; Yip, T.C.M.; Lo, I.M.C. Kinetic interactions of EDDS with soils. 2. Metal− EDDS complexes in uncontaminated and metal-contaminated soils. *Environ. Sci. Technol.* **2009**, *43*, 837–842. [CrossRef] [PubMed]
46. Fabbricino, M.; Ferraro, A.; Luongo, V.; Pontoni, L.; Race, M. Soil Washing Optimization, Recycling of the Solution, and Ecotoxicity Assessment for the Remediation of Pb-Contaminated Sites Using EDDS. *Sustainability* **2018**, *10*, 636. [CrossRef]
47. Lo, I.M.C.; Tsang, D.C.W.; Yip, T.C.M.; Wang, F.; Zhang, W. Influence of injection conditions on EDDS-flushing of metal-contaminated soil. *J. Hazard. Mater.* **2011**, *192*, 667–675. [CrossRef] [PubMed]
48. Kumar, M.; Ramanathan, A.L. Vertical Geochemical Variations and Speciation Studies of As, Fe, Mn, Zn, and Cu in the Sediments of the Central Gangetic Basin: Sequential Extraction and Statistical Approach. *Int. J. Environ. Res. Public Health* **2018**, *15*, 183. [CrossRef] [PubMed]
49. Vesali Naseh, M.R.; Noori, R.; Berndtsson, R.; Adamowski, J.; Sadatipour, E. Groundwater Pollution Sources Apportionment in the Ghaen Plain, Iran. *Int. J. Environ. Res. Public Health* **2018**, *15*, 172. [CrossRef] [PubMed]
50. Ding, Z.; Hu, X.; Wu, X.; Yin, D. Metal contents and fractionation in contaminated soil after column leaching using [S, S]-EDDS. *Chem. Speciat. Bioavailab.* **2010**, *22*, 247–255. [CrossRef]
51. Arfania, H.; Asadzadeh, F. Mobility of heavy metals after spiking in relation to sediment and metal properties: Leaching column study. *J. Soils Sediments* **2015**, *15*, 2311–2322. [CrossRef]

International Journal of
Environmental Research and Public Health

MDPI

Article

Medical and Legal Aspects of Child Sexual Abuse: A Population-Based Study in a Hungarian County

Andrea Enyedy [1], Panagiotis Tsikouras [2,*] and Roland Csorba [3,4]

1 Department of Obstetrics and Gynecology, City Hospital of Nyíregyháza, Mák u. 10–14, 4400 Nyíregyháza, Hungary; enyedy.andrea@szszbmk.hu
2 Department of Obstetrics and Gynecology, Faculty of Medicine, Democritus University of Thrace, 68100 Alexandroupolis, Greece
3 Department of Obstetrics and Gynecology, City Hospital of Aschaffenburg, Am Hasenkopf 1, 63739 Aschaffenburg, Germany; drcsorbaroland@gmail.com
4 Department of Obstetrics and Gynecology, Faculty of Medicine, University of Debrecen, 4033 Debrecen, Hungary
* Correspondence: ptsikour@med.duth.gr; Tel.: +30-697-472-8272

Received: 1 March 2018; Accepted: 5 April 2018; Published: 9 April 2018

Abstract: Background: Very few studies focus on childhood sexual abuse in middle European countries. Aim: The purpose of our study is to describe the medical and legal characteristics of children who experience sexual abuse and explore common features that may result in strategies for prevention. Methods: Between 2000 and 2015, 400 girls and 26 boys under the age of 18, suspected of being sexually abused, visited one of the four hospitals in a Hungarian county. Results: Mean age at onset was 10.81 years for boys, 13.46 years for girls. In 278 cases (65.3%), the perpetrator was known to the victim, and a stranger was suspected in 148 cases (34.7%). In 79 cases (30.7% of boys and 17.7% of girls), a family member was the accused perpetrator. In more than one-third (boys) and in one-fifth (girls) of cases, sexual abuse had occurred on multiple occasions. In the case of boys, child and adolescent sexual abuse (CSA) included oral genital, genital touching and genital to genital contact in 14 cases (53.8%) and anal intercourse in 12 (46.2%) cases. In case of girls, sexual abuse included coitus in 219 (54.8%), oral genital, genital touching, genital to genital contact in 164 (41.0%), anal abuse in 14 (3.5%) cases, physical injury was incurred in 15 cases. Legal proceedings followed the CSA in 205 (48.1%) cases. Conclusion: The results highlight the urgent need to address the issue of sexual abuse in Hungary and minimize its impact. Prevention requires a systematic and lifelong approach to educating children about personal space safety and privacy to reduce vulnerability and is the responsibility of parents and professionals.

Keywords: child sexual abuse; adolescents; girls; boys; gender differences; legal proceedings; child-friendly justice; prevention

1. Introduction

Child and adolescent sexual abuse (CSA) is a widespread problem not only in the western countries but also in the eastern and middle European countries, as well. Sexual abuse in childhood and adolescence is defined as the involvement of the young individual in a sexual activity that he/she does not fully comprehend or is not developmentally prepared for, or violates the law or social taboos of the society [1,2]. The sexual activities may include all forms of oral-genital, genital, or anal contact by or to the child, or nontouching abuses, such as exhibitionism, voyeurism, or using the child in the production of pornography. Sexual abuse includes a spectrum of activities ranging from rape to physically less intrusive sexual abuse. 'Child sexual abuse is more frequent than childhood cancer, juvenile diabetes, and congenital heart disease combined...' [3].

Various studies have researched the prevalence of child sexual abuse and the international estimates show prevalence rates that are frightening high. It is estimated that 10–20% of girls and 5–10% of boys are victims of child sexual assault [4], according to data collected between 1994 and 2007 in 39 predominance studies from 28 countries. These figures are in accordance with those of previous surveys [5]. In a meta-analysis of 323 studies worldwide, featuring a total of 9.9 million abused children, the worldwide incidence was revealed to be 12.7% (18.0% for girls, 7.6% for boys) [6]. In the USA, where the reporting of child sexual assault is obligatory, 60,000 to 80,000 confirmed cases are reported per year, showing a general decrease [7]. The incidence of child sexual assault is 20% in the UK and Switzerland [8,9], while it is to a certain degree higher in Germany, with 20.1%, and in Spain, with 22% [10,11]. In north European countries, the incidence appeared to be significantly lower: 17% in Norway and 13% in Sweden [12,13].

The literature reports a lifelong association between sexual victimization in childhood and adolescence and chronic mental and physical illness in adulthood [14]. CSA has been related to several negative health problems, including both physical and mental conditions [15–17]. Victims with child maltreatment histories are more likely to manifest greater risk for violent crimes [18], substance use disorders [19], more school discipline problems and suspensions [20], poor long-term intellectual and academic achievement [21], and greater likelihood of becoming a teenage parent [22,23]. There are also significant economic and societal consequences [24]. Women with a history of CSA account for a significantly higher proportion of primary care and outpatient costs than women without any such history [25].

There are limited studies that focus on childhood abuse in middle and eastern European countries. The available data from Hungary are inadequate and in many cases unreported; trustworthy data on the frequency of subtypes of sexual abuse do not exist.

There are no Hungarian National or Country statistics obtainable. The incidence and nature of child sexual abuse in Hungary have not been researched, apart from a recent study that investigated sexual abuse in only female children [26] and some unscientific evidence from split sources.

In Hungary, when a disclosure occurs, the emergency, gynecology and pediatric departments are commonly the portal for entry, and gynecologists (for female victims), pediatricians, and traumatologists (for male victims) are generally the 'first reporters' when sexual abuse is suspected. This is followed by reporting and the involvement of social services and law enforcement. Although reporting of child abuse is not mandatory, in the case of a victim under 18 years of age, legal proceedings can be initiated by the victim's parent, carer or guardian, but not by the doctor.

The aim of this population-based study was to collect data from a Hungarian county with regard to the characteristic features of both male and female child abuse cases in an effort to further understand the extent of the problem, to progress towards reporting, and to approach the services and suggest educational and prevention strategies.

2. Methods

Our population-based study took place in 1 out of 19 Hungarian counties (Szabolcs-Szatmár-Bereg County, inhabitans: 585,000). In this county, there are only four cities, with four hospitals where these cases can be seen. The case notes of sexually abused girls and boys under 18 years, seen at the four hospitals of this county (University Teaching Hospital of Nyíregyháza, University Teaching Hospital of Fehérgyarmat, City Hospital of Mátészalka and the City Hospital of Kisvárda) between 1 January 2000 and 31 December 2015, were reviewed. These are the departments in Szabolcs-Szatmár-Bereg County to which adolescent cases with a suspicion of sexual abuse are referred. The female cases are exclusively referred to the gynecological department of the above listed hospitals; the boys are seen at the pediatric, urological and traumatological departments of these hospitals. Data were collected, recorded and extracted from the files of children who had allegedly experienced CSA and received care and follow-up services. Guidelines for standard management, including the definition of CSA and the purpose and procedure for the examination, were elaborated in detail at these departments.

In our population-based study, sexual abuse includes non-contact sexual abuse, sexual touch, oral sex, sexual penetration, anal sex and sexual exploitation.

The data collected included the age and education of the victim, family relationship between victim and perpetrator, frequency of abuse, type and place of abuse, season of the year and time of day when the abuse occurred, family relationship between victim and person accompanying her/him on presentation at the hospital, the time interval between alleged sexual contact and clinical examination and the findings on the medical examination.

An adolescent gynecologist for females, or a pediatrician, urologist or a traumatologist for males, performed the examination. All cases were managed according to the standard guidelines [27]. We obtained full medical and social histories and conducted physical examinations following a standard protocol. The clinical investigation included a 'head-to-toe' physical examination, assessment of sexual development, identification of any injury with specific reference to colposcopic appearances of the hymenal membrane and surrounding surfaces (girls), identification of any injury on the penis, scrotum and anus (boys), signs of abuse. Diagnosis of a possible pregnancy was ruled out by a urine test or ultrasound examination, collection of forensic evidence (sperm, saliva, other trace evidence) and evaluation for sexually transmitted infections were also part of our examination. Appropriate medical or surgical treatment was provided based on clinical findings. All relevant findings were documented and incorporated into a report and distributed in response to official requests. We analyzed the difference between CSA and rape, emphasizing the diagnostic challenges. The results of the legal proceedings were also recorded. The criminal, judicial and medical records were continuously monitored and compared.

Descriptive analyses were conducted using SPSS (Statistical Package for Social Science, SPSS Inc., Chicago, IL, USA). Estimates of mean value, standard deviation, frequency, *t*-test, Mann-Whitney test, confidence interval and correlation were assessed. Prevalence rates were tested using Pearson's X^2 tests. The significance level was kept at 0.05 level.

3. Results

Between 2000 and 2015, we enrolled 400 girls and 26 boys under 18 years of age who were seen at one of the four hospitals following sexual abuse in the study. Mean age at onset was 10.81 years (SD 3.453) for boys, 13.46 years (SD 3.364) for girls and 13.30 years (SD 3.425) for total victims. The sample characteristics are summarized in Table 1. Forty-four percent of children were aged between 11 and 14 years, the majority of boys (46.2%) were younger than 10 years old, the majority of girls (44.5%) were between 11 and 14 years old. Of the victims, 243 (57.0%) were pupils, 26 (6.1%) preschool children and 157 (36.9%) were working or unemployed. In 278 cases (65.3%), the perpetrator was known to the victim, and a stranger was suspected in 148 cases (34.7%). In 79 of all cases (30.7% by boys and 17.7% by girls) a family member was the accused perpetrator. Perpetrators were most often familiar to the children (38.5% for boys and 47.3% for girls). A relatively low proportion of the intrafamilial abuses were committed by the victim's cousin or stepbrother (3–3 cases) for boys and the stepfather (24 cases) or father (12 cases) for girls. The distribution of the accompanying persons is presented in Table 1. In most cases, male victims were escorted to the hospital by their mother (9, 34.6%), and female victims by a police officer (200, 50.0%).

The characteristics of the CSA itself are presented in Table 2. Analysis of the time interval between the crime and the medical assessment showed a significant difference between the two genders. In case of male victims, only 2 victims (7.7%) had been able to get immediate emergency care, while 5 (19.2%) boys were seen at the hospital within 72 h and the majority—19 (73.1%)—were examined more than 72 h after the abuse. In case of female victims, the majority of the girls—152 (38.0%)—were seen more than 72 h after the abuse, but 98 (24.5%) cases had been able to receive immediate care, and 150 (37.5%) cases within 72 h. Furthermore, we found that in more than one-third (boys) and in one-fifth (girls) of cases, sexual abuse had occurred on multiple occasions. In the case of boys, CSA included sexual perversion (oral genital, genital touching and genital to genital contact) in 14 cases (53.8%) and anal intercourse in 12 (46.2%) cases; in our sample, the boys had not reported physical abuse. In the case of girls, sexual abuse included coitus in 219 (54.8%), sexual perversion in 164 (41.0%), and anal abuse in 14 (3.5%) cases; physical injury was incurred in 15 cases. Physical examination of the children focused on two points: sign of injury and identification/collection of forensic evidence whether genital, anal or extragenital. Injury occurred in 3.5% of cases, but none of them required surgical treatment. The presence of sperm was confirmed in 118 cases (27.7%). In 12 (3.0%) cases, the pregnancy test was positive. On looking at the diurnal timing of individual cases, we found dissonant results between the genders: the boys were more likely abused in the afternoon, while the girls were mostly abused in the evening. Seasonal occurrence showed comparable results, with CSA occurred mostly during the summer, 10 cases (38.5%) for boys, 117 cases (29.3%) for girls, when children were on holiday from school. In the majority of the cases (31%), the location of the abuse was the victim's home (34.6% for boys and 30.85 for girls). Other locations included: at a familiar home (19.5%), public space (14.3%), in a forest/field (11%) and at a children's home (4%). During the study period, legal proceedings followed the CSA in 205 (48.1%) cases. The number of perpetrators who were sentenced was 41 (9.6%), 22 of them were found guilty on charges of rape, 4 on charges of sexual perversion and 15 on charges of illegal sexual intercourse.

Table 1. Characteristics of the victims (n = 426).

Characteristics	Category	Male N (%)	Female N (%)	Total N (%)
Age (years)	<10	12 (46.2)	56 (14.0)	68 (16.0)
	11–14	10 (38.5)	178 (44.5)	188 (44.1)
	>14	4 (15.4)	166 (41.5)	170 (39.9)
Education	Preschool	4 (15.4)	22 (5.5)	26 (6.1)
	Pupil	19 (73.1)	224 (56.0)	243 (57.0)
	Other	3 (11.5)	154 (38.5)	157 (36.9)
Perpetrator	Father	1 (3.8)	12 (3.0)	13 (3.1)
	Stepfather	0	24 (6.0)	24 (5.6)
	Stepbrother	3 (11.5)	2 (0.5)	5 (1.2)
	Cousin	3 (11.5)	2 (0.5)	5 (1.2)
	Grandfather	1 (3.8)	0	1 (0.2)
	Other relative	0	31 (7.8)	31 (7.3)
	Familiar	10 (38.5)	189 (47.3)	199 (46.7)
	Stranger	8 (30.8)	140 (35.0)	148 (34.7)
Escorter	Alone	0	43 (10.8)	43 (10.1)
	Mother	9 (34.6)	96 (24.0)	105 (24.6)
	Parents	6 (23.1)	21 (5.3)	27 (6.3)
	Other relative	2 (7.7)	9 (2.3)	11 (2.6)
	Familiar	4 (15.4)	18 (4.5)	22 (5.2)
	Paramedical officer	0	13 (3.3)	13 (3.1)
	Police	5 (19.2)	200 (50.0)	205 (48.1)

Table 2. Characteristics of CSA cases.

Characteristics	Category	Male N (%)	Female N (%)	Total N (%)	p Value
Time of the examination	Immediate	2 (7.7)	98 (24.5)	100 (23.5)	* $p = 0.002$
	Within 72 h	5 (19.2)	150 (37.5)	155 (36.4)	
	More than 72 h	19 (73.1) *	152 (38.0)	171 (40.1)	
Type of abuse	Vaginal penetration	0	219 (54.8)	219 (51.4)	* $p < 0.001$
	Sexual perversion	14 (53.8)	164 (41.0)	178 (41.8)	
	Anal penetration	12 (46.2) *	14 (3.5)	14 (3.3)	
	Physical abuse	0	15 (3.75)	15 (3.5)	

Table 2. *Cont.*

Characteristics	Category	Male N (%)	Female N (%)	Total N (%)	*p* Value
Pregnancy Test	Negative	-	5 (1.3)	-	-
	Positive	-	12 (3.0)	-	
	Not done	-	383 (95.8)	-	
Sperm diagnostic	Positive	1 (3.8)	117 (29.3)	118 (27.7)	* p = 0.005
	Negative	25 (96.2) *	283 (70.8)	308 (72.3)	
Occurrence	single	18 (69.2)	320 (80.0)	338 (79.3)	* p = 0.189
	multiple	8 (30.8)	80 (20.0)	88 (20.7)	
Location of abuse	Familiar home	2 (7.7)	81 (20.3)	83 (19.5)	
	Car	0	13 (3.3)	13 (3.1)	
	Children's refuge	4 (15.5)	13 (3.3)	17 (4.0)	
	School	3 (11.5)	21 (5.3)	24 (5.6)	
	Bar, disco	0	34 (8.5)	34 (8.0)	
	Public space	4 (15.4)	57 (14.3)	61 (14.3)	
	Forest, field	3 (11.5)	44 (11.0)	47 (11.0)	
	Victim's home	9 (34.6)	123 (30.8)	132 (31.0)	
	other	1 (3.8)	14 (3.5)	15 (3.5)	
Seasonal occurrence	Spring	6 (23.1)	112 (28.0)	118 (27.7)	
	Summer	10 (38.5)	117 (29.3)	127 (29.8)	
	Autumn	5 (19.2)	84 (21.0)	89 (20.9)	
	Winter	5 (19.2)	87 (21.8)	92 (21.6)	
Diurnal occurrence	Morning	5 (19.2)	54 (13.5)	59 (13.8)	
	Afternoon	11 (42.3)	132 (33.0)	143 (33.6)	
	Evening	7 (26.9)	150 (37.5)	157 (36.9)	
	Night	3 (11.5)	64 (16.0)	67 (15.7)	

4. Discussion

Health care, child protection and legal systems continue to be challenged by the recognition of the societal scourge of child sexual abuse and strive to understand its complexities and how best to meet the myriad of interdisciplinary challenges. For more than 30 years, in many countries, medical doctors have been charged with the diagnosis, treatment and management of victims of sexual abuse. Their effort has been generally independent of the larger systems needed to provide protection, prevention and prosecution.

Hungary has just begun to deal with this societal issue. To the best of our knowledge, this is the first Hungarian population-based study to have researched characteristics of CSA for both girls and boys. We have been working to address the common myths regarding CSA and raise public awareness that CSA in Hungary can no longer to be 'another hidden pediatric problem', as stated by Kempe in 1977 [28].

In Hungary, during recent decades, child sexual abuse has become a priority for medical and criminal law professionals due to its frequent prevalence, serious adverse health effects and potential for lifetime consequence for the victim. In Hungary, when a disclosure occurs, the emergency, gynecology and pediatric departments are commonly the portal for entry, and gynecologists (for female victims), and pediatricians and traumatologists (for male victims), are generally the 'first reporters' when sexual abuse is suspected. Reporting and engagement of social services and law enforcement brings resources and critically important collaboration. Despite the robust association between gynecological symptoms and a history of CSA, no population-based central/eastern European data are available. Knowledge of the exact number of sexually abused children and the establishment of a systematic approach to address the medical and legal needs of alleged child victims are yet to be developed.

We were able to identify differences between CSA and rape. In the majority of cases (96.5%), force had not been used: children were mostly abused by perpetrators (65.3%) they knew and trusted. In one-third (boys) and one-fifth (girls) of the victims, the accused perpetrator of CSA was a family member and it was a domestic sexual attack (31%). In these cases, because of the long time interval between the sexual contact and the revelation (40.1%), the chance of identifying forensic evidence was lost. The time that passes between the abusive event and the physical examination is of great importance. Most sexually abused children will not have signs of genital or anal injury, especially when examined non-acutely.

Children who may have been abused should be examined by a physician to diagnose and treat any effect of the alleged sexual contact, which includes the identification of injuries, treatment for sexually transmitted infection (STI) and the collection of forensic evidence if present. The reality is that less than 5% of children will have diagnostic findings of sexual contact, and 3–5% will have an STI. Biological evidence (sperm) of recent abuse can be successfully secured (abuse within the past 24 h if before puberty, within the past 72 h in pubertal girls), and for medical reasons if there is any bleeding [29]. The most available evidence is what the child victim provides in their account of the inappropriate sexual contact.

As many as 30.8% of boys and 20% of girls reported histories of multiple occurrence of sexual contact. A delayed disclosure is typical for abuse, and since most injuries that children incur as a result of sexual contact will be superficial and heal without any lasting residual, most medical examinations show neither acute nor healed findings [30]. It has been proved that a single incomplete hymenal rupture can heal in 9 days, and a complete rupture in 24–30 days, after the trauma [31]. The most important reason for the paucity of abnormal findings is the nature of the abuse itself, as most perpetrators have little intent of actually harming the child, and if they do, the injuries are generally superficial. Although many nonmedical professionals may believe that it should be possible for a doctor to determine if a child has been 'penetrated', this is not always the case once children enter puberty and the hymenal membrane increases its elasticity and distensability. Determining 'penetration' in the prepubertal child is less of a challenge because of the lack of elasticity of the hymen, and when true vaginal penetration occurs, there is an accompanying history of pain, bleeding and diagnostic residual. That said, most prepubertal children experience penetration into the vaginal vestibule and not the vagina. Illustrative of the challenge of determining whether a pubertal child is a virgin is confirmed by a study in which only 2 (6%) of 36 pregnant teenagers manifested clear evidence of a prior penetration injury, and only 4 (11%) had suspicious, though not definitive, findings: "'Normal' does not mean 'nothing happened'" [32]. Normal findings are the rule, not the exception, in victims of child sexual abuse, with or without penetration, whether chronic or acute [33–35].

Special attention is required in cases in which a member of the child's family carried out sexual abuse. During the study-period, we saw only 8 male (30.7%) and 71 female (17.7%) cases in which the first instances of abuse were reported. The highest frequency was observed under the age of 10 years for boys and between the ages of 11 and 14 years for girls. The actual number was probably much higher than 79, because of underreporting. In almost every case, the child reported threats to maintain secrecy including threats of physical abuse if they disclosed. The mothers feared losing their homes, and possibly their children. This is the foundation on which multiple and chronic sexual abuse can continue. Most of the CSA cases occurred during the summer months when children were on holiday from school. Conversely, those children who were the victims of rape perpetrated by strangers (30.8% by boys and 35% by girls) presented immediately, and with more obvious findings (sperm found in 118 cases), as they had family support and immediate examination could be performed (in 100 cases, 23.5%). Unfortunately, few children reveal sexual contact immediately following CSA, which limits the opportunity to recognize injuries and collect forensic evidence. Sexual abuse is usually a chronic, complex, and often particularly traumatizing incident for the victim, frequently committed by family members or other trusted persons in the setting of relationship dependence and strong authority relationships [36].

In Hungary, since the decision made by the constitutional court on 4 September 2002, the perpetrator and the victim in CSA cases can be of the same sex. With that decision, the Law took account of the results of reported studies in male victims. In the case of a victim under 18 years of age, legal proceedings can be initiated by the victim's parent, carer or guardian. The new act (from 1 July 2003) was intended to ensure considerate behavior towards the victim. The evidence given by the victim can be recorded on videotape and this tape can be used later during legal proceedings. In 2012 a new Penal Code was introduced dealing with sexual crimes. The essence of the legal change was that where the previous Penal Code had focused on the defense of gender morality and public interest, the new Penal Code focused on the individual, gender integrity and gender self-determination. None of the laws prescribe obligation of charges, but the ordinances of the private proposal and the desuetude of criminality have been changed. In 2012 our government announced the Year of Child-friendly Justice; within the framework of the National Court of Justice, a work-team was established in order to evolve the concept of child-friendly justice. The elements of child-friendly justice have been continuously becoming the part of legal practice. In Hungary, the greatest challenge is not only to create the Penal Code, but also to assure its enforcement. The experience of child sexual abuse has the potential to result in long-term emotional and behavioral consequences for the victim. When children are suspected of experiencing sexual victimization, they deserve professionals who are knowledgeable, skilled, sensitive professionals capable of formulating objective opinions that do not further betray their trust, allowing justice to be served.

In our study, only 205 (48%) of 426 cases had legal proceedings initiated. Delivery of the judgements against the 41 (9.6%) perpetrators who were ultimately sentenced took several years in each of the cases. The low proportion of charges and the long time interval needed for a judgement in the Hungarian legal system demonstrate the challenges associated with successful prosecution. It is of great importance to organize more family planning centers for young families, children and teenagers to have access to appropriate health care services and suitable education to protect them from sexual abuse.

5. Conclusions

The prevention of sexual abuse is the responsibility not only of parents, but law enforcement agencies, health professionals and educators. In Hungary the training of professionals in the recognition and reporting of CSA is an urgent matter, the absence of a legal obligation to report needs reconsideration. The results stress the importance of the need for a public awareness campaign regarding the vulnerability of children of ages to sexual victimization with a focus on the development of primary prevention programs. Professionals in health care, child protection, mental health and law enforcement would all benefit from efforts to raise their awareness of child sexual abuse and what they can do in their professional role to address abuse when suspected.

Geolocation Info: The study took place in four cities (University Teaching Hospital of Nyíregyháza, University Teaching Hospital of Fehérgyarmat, City Hospital of Mátészalka and the City Hospital of Kisvárda) in Szabolcs-Szatmár-Bereg County in the northeastern part of Hungary. The county has approximately 585,000 inhabitants of the 9,900,000 inhabitants of Hungary. Hungary is located in Middle Eastern Europe, between Austria and Romania.

Author Contributions: Andrea Enyedy wrote the main part of the manuscript. Panagiotis Tsikouras contributed to evaluation of the study. Roland Csorba supervised the study and manuscript.

Conflicts of Interest: The authors declare no conflict of interest regarding this published material.

References

1. Häuser, W.; Schmutzer, G.; Brähler, E.; Glaesmer, H. Maltreatment in childhood and adolescence: Results from a survey of a representative sample of the German population. *Dtsch. Arztebl. Int.* **2011**, *108*, 287–294. [PubMed]
2. Herrmann, B.; Banaschak, S.; Csorba, R.; Navratil, F.; Dettmeyer, R. Physical examination in child sexual abuse, approaches and current evidence. *Dtsch. Arztebl. Int.* **2014**, *111*, 692–703.
3. Kaplan, R.; Adams, J.A.; Staling, S.P.; Giardino, A.P. *Medical Response to Child Abuse: A Resource for Professionals Working with Children and Families*; STM Learning: St. Louis, MO, USA, 2011.

4. Pereda, N.; Guilera, G.; Forns, M.; Gómez-Benito, J. The international epidemiology of child sexual abuse: A continuation of Finkelhor (1994). *Child Abuse Negl.* **2009**, *33*, 331–342. [CrossRef] [PubMed]
5. Finkelhor, D. The international epidemiology of child sexual abuse. *Child Abuse Negl.* **1994**, *18*, 409–417. [CrossRef]
6. Stoltenborgh, M.; van Ijzendoorn, M.H.; Euser, E.M.; Bakermans-Kranenburg, M.J. A global perspective on child sexual abuse: Meta-analysis of prevalence around the world. *Child Maltreat.* **2011**, *16*, 79–101. [CrossRef] [PubMed]
7. U.S. Department of Health & Human Services: Child Maltreatment 2012. Available online: www.acf.hhs.gov/programs/cb/resource/child-maltreatment-2012 (accessed on 17 December 2013).
8. Keane, F.E.; Young, S.M.; Boyle, H.M. The prevalence of previous sexual assault among routine female attenders at a department of genitourinary medicine. *Int. J. STD AIDS* **1996**, *7*, 480–484. [CrossRef] [PubMed]
9. Tschumper, A.; Narring, F.; Meier, C.; Michaud, P.A. Sexual victimization in adolescent girls (age 15–20 years) enrolled in post-mandatory schools or professional training programmes in Switzerland. *Acta Paediatr.* **1998**, *87*, 212–217. [CrossRef] [PubMed]
10. Peschers, U.M.; Du Mont, J.; Jundt, K.; Pfürtner, M.; Dugan, E.; Kindermann, G. Prevalence of sexual abuse among women seeking gynecologic care in Germany. *Obstet. Gynecol.* **2003**, *101*, 103–108. [PubMed]
11. López, F.; Carpintero, E.; Hernández, A.; Murtin, M.J.; Fuertes, A. Prevalence and sequelae of childhood sexual abuse in Spain. *Child Abuse Negl.* **1995**, *19*, 1039–1050. [CrossRef]
12. Pedersen, W.; Skrondal, A. Alcohol and sexual victimization: A longitudinal study of Norwegian girls. *Addiction* **1996**, *91*, 565–581. [CrossRef] [PubMed]
13. Edgardh, K.; Ormstad, K. Prevalence and characteristics of sexual abuse in a national sample of Swedish seventeen-year-old boys and girls. *Acta Paediatr.* **2000**, *89*, 310–319. [CrossRef] [PubMed]
14. De Bellis, M.D.; Spratt, E.G.; Hooper, S.R. Neurodevelopmental biology associated with childhood sexual abuse. *J. Child Sex. Abus.* **2011**, *20*, 548–587. [CrossRef] [PubMed]
15. Afifi, T.O.; MacMillan, H.L.; Boyle, M.; Taillieu, T.; Cheung, K.; Sareen, J. Child abuse and mental disorders in Canada. *CMAJ* **2014**, *186*, E324–E332. [CrossRef] [PubMed]
16. MacMillan, H.L.; Fleming, J.E.; Streiner, D.L.; Lin, E.; Boyle, M.H.; Jamieson, E.; Duku, E.K.; Walsh, C.A.; Wong, M.Y.; Beardslee, W.R. Childhood abuse and lifetime psychopathology in a community sample. *Am. J. Psychiatry* **2001**, *158*, 1878–1883. [CrossRef] [PubMed]
17. Scott, K.M.; McLaughlin, K.A.; Smith, D.A.; Ellis, P.M. Childhood maltreatment and DSM-IV adult mental disorders: Comparison of prospective and retrospective findings. *Br. J. Psychiatry* **2012**, *200*, 469–475. [CrossRef] [PubMed]
18. Deykin, E.Y.; Buka, S.L. Prevalence and risk factors for posttraumatic stress disorder among chemically dependent adolescents. *Am. J. Psychiatry* **1997**, *154*, 752–757. [PubMed]
19. De Bellis, M.D. Developmental traumatology: A contributory mechanism for alcohol and substance use disorders. *Psychoneuroendocrinology* **2002**, *27*, 155–170. [CrossRef]
20. Eckenrode, J.; Laird, M.; Doris, J. School performance and disciplinary problems among abused and neglected children. *Dev. Psychol.* **1993**, *29*, 53–62. [CrossRef]
21. Perez, C.M.; Widom, C.S. Childhood victimization and long-term intellectual and academic outcomes. *Child Abuse Negl.* **1994**, *18*, 617–633. [CrossRef]
22. Anda, R.F.; Chapman, D.P.; Felitti, V.J.; Edwards, V.; Williamson, D.F.; Croft, J.B.; Giles, W.H. Adverse childhood experiences and risk of paternity in teen pregnancy. *Obstet. Gynecol.* **2002**, *100*, 37–45. [PubMed]
23. Hillis, S.D.; Anda, R.F.; Dube, S.R.; Felitti, V.J.; Marchbanks, P.A.; Marks, J.S. The association between adverse childhood experiences and adolescent pregnancy, long-term psychosocial consequences, and fetal death. *Pediatrics* **2004**, *113*, 320–327. [CrossRef] [PubMed]
24. Dube, S.R.; Anda, R.F.; Felitti, V.J.; Chapman, D.P.; Williamson, D.F.; Giles, W.H. Childhood abuse, household dysfunction, and the risk of attempted suicide throughout the life span: Findings from the Adverse Childhood Experiences Study. *JAMA* **2001**, *286*, 3089–3096. [CrossRef] [PubMed]
25. Walker, E.A.; Unutzer, J.; Rutter, C.; Gelfand, A.R.; Saunders, K.; VonKorff, M.; Koss, M.; Katon, W. Costs of health care use by women HMO members with a history of childhood abuse and neglect. *Arch. Gen. Psychiatry* **1999**, *56*, 609–613. [CrossRef] [PubMed]

26. Csorba, R.; Tsikouras, P.; Lampé, R.; Póka, R. The sexual abuse of female children in Hungary: 20 years' experience. *Arch. Gynecol. Obstet.* **2012**, *286*, 161–166. [CrossRef] [PubMed]
27. Enyedy, A.; Csorba, R. Female child sexual abuse. *Orv. Hetil.* **2017**, *158*, 910–917. [CrossRef] [PubMed]
28. Kempe, C.H. Sexual abuse, another hidden pediatric problem: The 1977 C. Anderson Aldrich lecture. *Pediatrics* **1978**, *62*, 382–389. [PubMed]
29. Palusci, V.J.; Cox, E.O.; Shatz, E.M.; Schultze, J.M. Urgent medical assessment after child sexual abuse. *Child Abuse Negl.* **2006**, *30*, 367–380. [CrossRef] [PubMed]
30. DeLago, C.; Deblinger, E.; Schroeder, C.; Finkel, M.A. Girls who disclose sexual abuse: Urogenital symptoms and signs after genital contact. *Pediatrics* **2008**, *122*, e281–e286. [CrossRef] [PubMed]
31. McCann, J.; Voris, J.; Simon, M. Genital injuries resulting from sexual abuse: A longitudinal study. *Pediatrics* **1992**, *89*, 307–317. [PubMed]
32. Kellogg, N.D.; Menard, S.W.; Santos, A. Genital anatomy in pregnant adolescents: "normal" does not mean "nothing happened". *Pediatrics* **2004**, *113*, e67–e69. [CrossRef] [PubMed]
33. Slaughter, L.; Henry, T. Rape: When the exam is normal. *J. Pediatr. Adolesc. Gynecol.* **2009**, *22*, 7–10. [CrossRef] [PubMed]
34. Adams, J.A.; Harper, K.; Knudson, S.; Revilla, J. Examination findings in legally confirmed child sexual abuse: It's normal to be normal. *Pediatrics* **1994**, *94*, 310–317. [PubMed]
35. Heger, A.; Ticson, L.; Velasquez, O.; Bernier, R. Children referred for possible sexual abuse: Medical findings in 2384 children. *Child Abuse Negl.* **2002**, *26*, 645–659. [CrossRef]
36. Zimmermann, P.; Neumann, A.; Celik, F. *Sexuelle Gewalt Gegen Kinder in Familien*; Deutsches Jugendinstitut: München, Germany, 2011.

International Journal of
Environmental Research and Public Health

MDPI

Article

Comparative Evaluation of Arabin Pessary and Cervical Cerclage for the Prevention of Preterm Labor in Asymptomatic Women with High Risk Factors

Panagiotis Tsikouras [1,*], George Anastasopoulos [2], Vasileios Maroulis [1], Anastasia Bothou [3], Anna Chalkidou [1], Dorelia Deuteraiou [1], Xanthoula Anthoulaki [1], Georgios Tsatsaris [2], Arzou Halil Bourazan [1], George Iatrakis [4], Stefanos Zervoudis [3], Georgios Galazios [1], Lola-Katerina Inagamova [1], Roland Csorba [5] and Alexander-Tobias Teichmann [5]

1 Department of Obstetrics and Gynecology, Democritus University of Thrace, 68100 Alexandroupolis, Greece; gbmaroul@yahoo.com (V.M.); annachalkidou@yahoo.gr (A.C.); Dr.dorelia@hotmail.com (D.D.); xanthi_vatsidou@hotmail.com (X.A); arzouhalil@hotmail.com (A.H.B.); ggalaz@med.duth.gr (G.G.); lola-katerina@hotmail.gr (L.-K.I.)
2 Medical Informatics Laboratory, Medical School, Democritus University of Thrace, 68100 Alexandroupolis, Greece; anast@med.duth.gr (G.A.); tsatsarisg3@gmail.com (G.T.)
3 Department of Obstetrics and Gynecology, Rea Hospital, 17564 Athens, Greece; natashabothou@windowslive.com (A.B.); szervoud@otenet.gr (S.Z.)
4 Department of Obstetrics and Gynecology, Technological Educational Institute, 17564 Athens, Greece; giatrakis@teiath.gr
5 Department of Obstetrics and Gynecology, Clinicum Aschaffenburg, Teaching Hospital of University, 97070 Würzburg, Germany; drcsorbaroland@gmail.com (R.C.); buero.teichmann@gmx.de (A.-T.T.)
* Correspondence: ptsikour@med.duth.gr; Tel.: +30-6974-728-272

Received: 17 March 2018; Accepted: 12 April 2018; Published: 18 April 2018

Abstract: Objective: Preterm labor is one of the most significant obstetric problems associated with high rate of actual and long-term perinatal complications. Despite the creation of scoring systems, uterine activity monitoring, cervical ultrasound and several biochemical markers, the prediction and prevention of preterm labor is still a matter of concern. The aim of this study was to examine cervical findings for the prediction and the comparative use of Arabin pessary or cerclage for the prevention of preterm birth in asymptomatic women with high risk factors for preterm labor. Material and methods: The study group was composed of singleton pregnancies (spontaneously conceived) with high risk factors for preterm labor. Cervical length, dilatation of the internal cervical os and funneling, were estimated with transvaginal ultrasound during the first and the second trimesters of pregnancy. Results: Cervical funneling, during the second trimester of pregnancy, was the most significant factor for the prediction of preterm labor. The use of Arabin cervical pessary was found to be more effective than cerclage in the prolongation of pregnancy. Conclusion: In women at risk for preterm labor, the detection of cervical funneling in the second trimester of pregnancy may help to predict preterm labor and to apply the appropriate treatment for its prevention. Although the use of cervical pessary was found to be more effective than cerclage, more studies are needed to classify the effectiveness of different methods for such prevention.

Keywords: cervical Arabin pessary; cerclage; second trimester of pregnancy

1. Introduction

Premature birth is defined as birth between 23 weeks + 5 days and 37 weeks of gestation, occurs in 10–12% of deliveries and is the leading cause of perinatal morbidity and mortality [1–3]. When it occurs, it leads to 75% perinatal mortality and morbidity and, in 50%, of cases there are long-term permanent neurological complications [4]. According to Chang's report, approximately 1.1 million neonates die from preterm birth complications [5]. In 2010, it was estimated that 14.9 million neonates were born in 11.1% of all births globally, 5% in Europe and 18% in Africa [5]. More than 60% of preterm births are observed in countries of South Asia and Sub-Saharan Africa [6]. Finnström suggested that the prolongation of pregnancy especially, to 23–26 weeks of gestation increases the survival rate of premature neonate by 3% [7].

Preterm birth is a clinical syndrome with various causes which can be genetic, maternal (smoking, multiple pregnancies, diabetes mellitus, hypertension cervical insufficiency, and age of <18 or >40), fetal (malpresentation), hormonal, social and environmental and it may be difficult to recognize the exact mechanism that provokes the labor. Until now, it is believed that the majority of preterm births have an idiopathic cause [8,9]. As medicine evolved in the detection of infections, both asymptomatic and symptomatic, the ability to study microorganisms (bacteria, viruses, and parasites) and their action (toxins, immune response, prostaglandins, proteases, etc.) helped to understand various things about the importance of infections in the provocation of preterm birth [8,9]. Although a lot of research has been performed on this important condition, the frequency has increased annually in the last years [6]. Ultrasound examination of the cervix can recognize women with increased risk of preterm birth based on sonographic measurement of the cervical length and funneling of the cervical internal os in the mid trimester and, consequently, prevention and intervention lead to decreased incidence of preterm birth. However, the diagnosis of cervical insufficiency (short cervical length and dilation of internal cervical os) is difficult and there are no existing objective diagnostic criteria [10–12]. In cases of cervical insufficiency, the value of cervical cerclage and cervical pessaries is still a matter of controversy [13,14].

The aim of this study was to examine cervical findings for the prediction and the comparative use of Arabin pessary or cerclage for the prevention of preterm birth in asymptomatic women with high risk factors for preterm labor.

2. Materials and Methods

In the interval between 2007 and 2012 in a retrospective cohort research study, we studied the cervical length and the dilatation of the internal os of pregnant women in the 1st trimester (10–14 weeks of gestation) and in the 2nd trimester (14–28 weeks of gestation). Ninety-five percent of all women were followed in Teaching Hospital Aschaffenburg of Germany and the rest in the Democritus University, Department of Obstetrics and Gynecology. The study was approved by the scientific committee as a clinical audit in the two departments.

All studied pregnancies were singleton asymptomatic pregnancies with high risk factors in the past such as recurrent miscarriage, previous preterm birth, recurrent vaginal bleeding, cervical surgical procedures, and infections.

Exclusion criteria included: stillbirths, fetal congenital anomalies, uterine anomalies, preeclampsia and metabolic diseases.

The transvaginal measurements were performed by experienced clinicians in the two departments according to standard recommended techniques.

Cervical length less of 2.7 cm, dilatation of the internal os more than 10 mm or indentation V or U shaped of the internal os were determined early as important signs for a cervical insufficiency.

In our pregnant women, we collected the following data: maternal age, parity, past obstetric and gynecological history, cervical length, funneling, cerclage vs. pessary insertion or conservative supervision of pregnancy gestation age at delivery and labor modus.

All study participants had normal findings in the sonographical measurements in the first trimester and were asymptomatic. Women with previous history of abnormal cervical factors were found to have abnormal measurements in reevaluation during the second trimester, especially during 14–28 weeks of gestation. The treatment of the cervical insufficiency was performed either with cerclage or Arabin pessary. In some cases, the performance of both methods was obligatory while in the rest neither was necessary except conservative medical therapy.

This study evaluates the possible role of cervical length, dilatation of the internal cervical os, funneling and the use of cerclage or pessary in avoiding this problem.

Our target was to study the correlation between the mentioned three abnormal parameters and preterm labor under 33 weeks. We defined "early preterm labor" as delivery under 33 weeks, since our patients have a medical history of high-risk factors and all of them had a preterm delivery. We used linear and logistic regression to make statistical analysis. We also tried to find which of the two invasive procedures (pessary and the placement of cervical sutures) is more effective in prolonging the week of delivery and which is less harmful. A p value <0.05 was considered as statistically significant.

3. Results

In total, 166 women were examined sonographically and we found that 95 (57.2%) had cervical length less than 2.7 cm. Dilatation of the internal cervical os >10 mm was found in 138 (83.1%) women and funneling was found in 51 (30.7%). Thirty-nine women (23.5%) who had transvaginal sonographic assessment of the cervix had preterm labor under 33 weeks. Only 20 of 166 women of the study (12%) had a cervical cerclage procedure and 124 of them (74%) had a pessary placement. About the mode of delivery, 61 participants (36.7%) had normal vaginal delivery, 9 (5.4%) had vacuum extraction and 96 (57.8%) had a low segment cesarean section (Table 1).

Table 1. Cervical characteristics of participants and labor modus.

		# of Women	%
Cervical length <2.7 cm		95	57.2
Dilatation of internal cervical os $\geq$10 mm		138	83.1
Cervical funneling		51	30.7
Cervical cerclage		20	12
Pessary		124	74
Preterm labor <33 weeks		39	23.5
Mode of delivery	Normal vaginal delivery	61	36.7
	Vacuum extraction	9	5.4
	Cesarean section	96	57.8

Table 2 describes the relationship between early labor under 33 weeks and the results of the transvaginal sonography. Logistic regression was used for the need of the specific study. As shown in Table 2, the prediction of labor under 33 weeks of gestation cannot be based on the results of transvaginal ultrasonography about cervical length and dilatation of the internal cervical os, since there is not enough statistical significance to support these specific cases. Of the 39 women who had an early preterm delivery, 25.6% had abnormal cervical length, whereas 21.12% had a normal one. However, this inference cannot be considered as a factor of early preterm delivery ($p = 0.534 > 0.05$). Women who had dilatation of cervical os >10 mm (25.3%) were more likely to have preterm delivery than those who did not (14.2%). However, for the same reason, a p value of 0.210 means that this factor cannot be considered. These results can also be perfectly explained in Figures 1 and 2. Concluding, if a woman has cervical diameter >10 mm or cervical length <2.7 cm or both, we cannot be definitely sure that she will have a delivery under 33 weeks. Funneling seems to be a predicting factor of great importance. It was found that there is a relationship between funneling and delivery under 33 weeks (YES = 39.2%, NO = 19.8%, $p < 0.05$). The treatment modus (cervical cerclage and pessary) were also checked in Table 3. Fifty percent of the ten women who had a placement of a cervical suture delivered under 33 weeks, whereas only 19.6% those who did not have a cerclage procedure had preterm labor under 33 weeks. The p value is <0.05 and this means that the result is statistically significant. In addition, it is concluded that the incidence of early preterm labor was significantly higher in women with cervical cerclage compared to women with no medical intervention. A group of women had a pessary placement. The outcome from the specific group was very encouraging. It was found that pessary does not affect the week of delivery in an adverse way. According to our findings, statistical significance of age ($p = 0.419$), parity ($p = 0.295$) and labor modus was not confirmed.

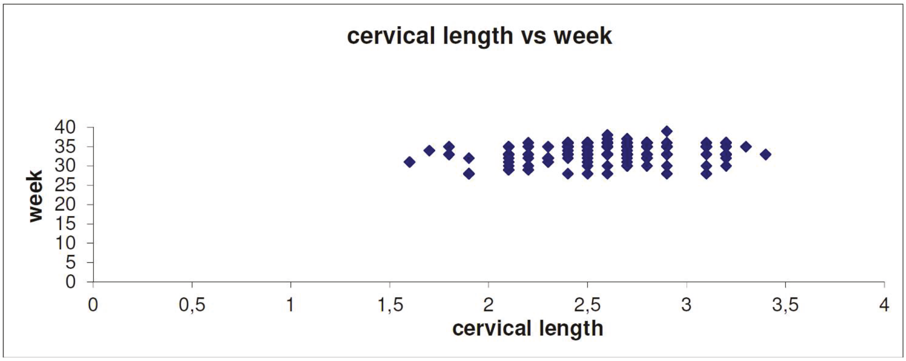


Figure 1. The association of cervical length and pregnancy week ($p = 0.534$).

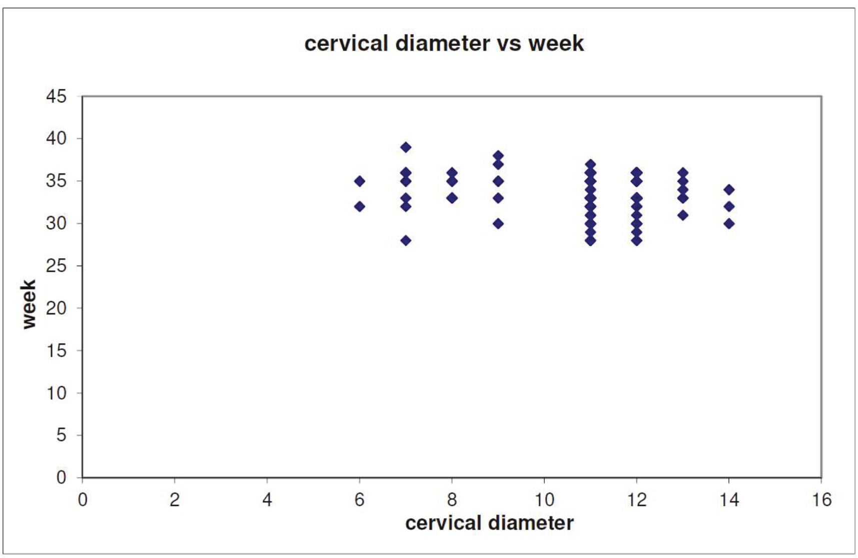


Figure 2. The association of cervical diameter and pregnancy week ($p = 0.210$).

Table 2. Cervical characteristics of the early preterm labor (<33 weeks) women. Bold type means statistical significance.

Cervical Characteristics		# of Women	%	p
Cervical length	<2.7 cm ≥2.7 cm	24 15	25.6 21.12	0.534
Dilatation of internal cervical os	>10 mm ≤10 mm	35 4	25.3 14.2	0.210
Cervical funneling	No Yes	19 20	19.8 39.2	**<0.001**

Table 3. Treatment of cervical insufficiency (cervical cerclage, pessary) of the early preterm labor (<33 weeks) women. Bold type means statistical significance.

Caption		# of Women	%	*p*
Cervical cerclage	No	29	19.6	**0.003**
	Yes	10	50	
Pessary	No	9	21.4	0.717
	Yes	30	24.2	

As presented in Table 4, if a funneling exists in a woman, there is a three times higher risk of early preterm (Odds Ratio (OR) = 3.260, with Confidence Interval (CI): 1.544–6.881 and $p < 0.05$) and if she has a cervical cerclage placement, there is a four times higher risk of labor under 33 weeks. The placement of the pessary has a 1.7 higher risk of labor (OR = 1.750, CI: 0.503–2.721 and $p > 0.05$) under 33 weeks and it is not statistically significant, which means that it does not affect negatively as the cerclage procedure.

Table 4. Correlation between preterm labor and cervical characteristics expressed as odds ratio (OR) with 95% confidence intervals (CI).

Cervical characteristics	OR	CI	*p*
Cervical length <2.7 cm	0.792	0.38–1.651	0.537
Dilatation of internal cervical os >10 cm	2.039	0.665–6.285	0.215
Funneling	3.260	1.544–6.881	0.002
Cervical cerclage	4.034	1.535–10.603	0.005
Pessary	1.170	0.503–2.721	0.715

Finally, in Table 5, the groups of women with one, two or no invasive characteristics are examined. Forty women with both cerclage and pessary had a great possibility of early preterm delivery ($p < 0.05$). For women who had no surgical intervention and had only pessary, there was less possibility ($p = 0.139$) for preterm delivery, which means that pessary is a procedure that does not affect negatively the week of labor. As far as the group of women who underwent a cerclage is concerned, there is no statistical significance ($p = 0.401$), which means that it cannot cause early preterm labor. Unfortunately, there is not enough evidence to rule it out (because there were only two cases).

Table 5. Labor outcome after treatment of cervical insufficiency (cervical cerclage, pessary). Bold type means statistical significance.

Pessary	Cerclage	# of Women	# of Early Preterm Labor Women	*p*
No	No	40	8	**<0.05**
Yes	No	106	21	0.139
No	Yes	2	1	0.401
Yes	Yes	18	9	<0.05

On the other hand, judging from the second and the last line of Table 4, we can see that pessary placement (second line) does not cause preterm labor under 33 weeks, but what make women of the last group have a delivery under 33 weeks is that they have also a placement of a cervical suture during the second trimester and it is presumed harmful for these women concerning their preterm labor.

By using logistic regression, we also managed to create three equations, in which we can predict if the week of labor is under 33, taking into account only the physical features (Equation (1)), only the invasive characteristics (Equation (2)), or both physical and invasive characteristics. If every feature is substituted by 0 or 1 according to the sample, then multiplied with the specific coefficient and added to the constant number of the first column, we find a number which will be very close to 0 or 1, representing the week of labor (early preterm labor <33 is 0, preterm labor is 1). The formulas are the following:

$$\text{Week} = \text{cervical length} \times (0.136) + \text{cervical diameter} \times (-0.242) + \text{funneling} \times (-1.112) + 1.745 \quad (1)$$

$$\text{Week} = \text{pessary} \times (0.011) + \text{cerclage} \times (-1.397) + 1.387 \quad (2)$$

$$\begin{aligned}\text{Week} = {} & \text{cervical length} \times (0.123) + \text{cervical diameter} \times (-0.398) + \text{funneling} \times (-0.983) + \\ & \text{pessary} \times (0.503) + \text{cerclage} \times (-0.740) + 1.568\end{aligned} \quad (3)$$

with cervical length (<2.7 = 0, >2.7 = 1), cervical diameter (>10 = 0, <10 = 1), funneling (yes = 1, no = 0), cerclage (yes = 1, no = 0), and pessary (yes = 1, no = 0).

4. Discussion

The preterm labor contribution to adverse outcome is largely related to pregnancy age at delivery. Despite continuous research, not a single effective method for satisfactory prognosis of preterm labor and prevention of preterm birth exists. Main risk factors have been suggested to increase the risk of prematurity, however, in most cases, it is not exactly possible to recognize clearly identifiable risk factors. In 25% of cases, the clinical symptomatology does not occur simultaneously with uterine activity [15]. Cervical evaluation by transvaginal ultrasonography in early pregnancy (first and second trimester) is a useful predictor of the risk for spontaneous preterm labor in asymptomatic pregnant women [16].

The cervix is a spindle-shaped structure, around 2 cm long and 1–2 cm wide with main fibrous structure elements, including collagen 80% type I, 20% type III and only a small amount of smooth cells about 10% [17]. Normally, in the late pregnancy prior to early phases of delivery at full term pregnancy, the cervix undergoes cervical ripening depending on biochemical changes, such as reduction of collagen synthesis and increased collagenase activity, which leads to delivery of the fetus [18,19]. In cases of preterm labor the cervical changes like cervical softening associated with painless dilatation, and shortening is explained by increased synthesis of interleukin (IL)-6 amd IL-8 and prostaglandin synthesis, and monocyte chemotactic protein I in absence of infection [20,21]. Asymptomatic pregnant women more commonly in the second than in the first trimester have a prediction to preterm labor in cases with abnormal sonographic cervical findings including the length of cervix (distance between the triangular area of echo density at the external cervical os and the V shaped notch at the internal one) [22,23].

According to previously published literature, the shortened cervical length is mainly a powerful biological marker of preterm labor andg a strong inverse association between cervical length and risk of preterm labor exists [24–26]. This risk is especially very high in cases with length less than 15 mm and is equivalent in multiple pregnancies with occurrence at 25 mm [27,28].

Based on our findings, we confirm that, in asymptomatic pregnant women with risk factors mainly in the past, the performance of transvaginal ultrasound cervical assessment in the second trimester is of great importance, even if the ultrasound examination in the first trimester was without abnormal findings. In our cohort, cervical funneling is a main prognostic factor of the prediction of preterm labor: it has three times higher risk compared to the rest of the participants with only abnormal cervical length. This finding is in accordance to previously published papers [29,30].

Concerning the other examinations parameter with cervical length, we found no statistically significant correlation between this one and preterm labor. This finding is surprisingly against previously published literature and our report 10 years ago in which we included abnormal ultrasound cervical assessments in the first and second trimester [31–34].

The fact that the frequency of preterm birth is not decreasing and is associated with significant costs, the aim of preventing treatment with cervical cerclage, progesterone, and vaginal pessaries is to prolong the duration of pregnancy and to decrease the perinatal morbidity and mortality [35–39].

Prevention of Preterm Birth

In recent years, the use of vaginal pessary has been returned to the forefront. The vaginal pessary that is used is Arabin pessary [40]. A multifactorial research in Spain proved that, in women with cervical length <25 mm in 18–22 weeks of pregnancy, the use of pessary decreased premature birth in 34 weeks of gestation by 88% and also decreased neonate complications [41]. On the other hand, a new study proved that the preventive usage of vaginal pessary in twin pregnancies does not improve the perinatal result [42]. We have to wait also for the result of other studies to prove or not its usage, and whether the use of vaginal pessaries in women with decreased length of the cervix is helpful. When they do not want or they cannot undergo cervical cerclage, particularly for women with decreased cervical length after 25 weeks of gestation, cervical cerclage is not helpful. We found no negative association between early preterm labor and pessary and so we can confirm the effectiveness of this conservative procedure. No complications were noticed [43,44]. In cases where a cerclage was performed, the contribution of the surgical procedure was not positive to prolong the pregnancy duration, although no serious side effects occurred. However, the necessity of cerclage was not confirmed in our participants, similar to several previously published papers [45–48].

5. Conclusions

Transvaginal sonographic detection of cervical funneling in the second trimester of pregnancy is the most important marker for the prediction of preterm labor. Arabin pessary (probably in combination with progesterone) seems to be effective in its prevention. Further multicenter studies are necessary to confirm these findings and determine as guidelines in the future.

Author Contributions: For research articles with several authors, a short paragraph specifying their individual contributions must be provided. The following statements should be used. Roland Csorba, George Iatrakis, Stefanos Zervoudis, Georgios Galazios and Alexander-Tobias Teichmann conceived and designed the experiments; performed the experiments; Vasileios Maroulis, Anastasia Bothou, Anna Chalkidou, Dorelia Deuteraiou, Xanthoula Anthoulaki, Arzou Halil Bourazan, Lola-Katerina Inagamova and George Anastasopoulos, Georgios Tsatsaris. analyzed the data; George Anastasopoulos, Georgios Tsatsaris contributed reagents/materials/analysis tools; Panagiotis Tsikouras wrote the paper. Authorship must be limited to those who have contributed substantially to the work reported.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wise, J. NICE guideline aims to cut premature birth rates. *BMJ* **2015**, *351*, h6253. [CrossRef] [PubMed]
2. Lim, K.; Butt, K.; Crane, J.M. SOGC Clinical Practice Guideline. Ultrasonographic cervical length assessment in predicting preterm birth in singleton pregnancies. *J. Obstet. Gynaecol. Can.* **2011**, *33*, 486–499. [CrossRef]
3. Slattery, M.M.; Morrison, J.J. Preterm delivery. *Lancet* **2002**, *360*, 1489–1497. [CrossRef]
4. Goldenberg, R.L.; Culhane, J.F.; Iams, J.D.; Romero, R. Epidemiology and causes of preterm birth. *Lancet* **2008**, *371*, 75–84. [CrossRef]
5. Chang, H.H.; Larson, J.; Blencowe, H.; Spong, C.Y.; Howson, C.P.; Cairns-Smith, S.; Lackritz, E.M.; Lee, S.K.; Mason, E.; Serazin, A.C.; et al. Born Too Soon Preterm prevention analysis group. Preventing preterm births: Analysis of trends and potential reductions with interventions in 39 countries with very high human development index. *Lancet* **2013**, *381*, 223–234. [CrossRef]
6. Blencowe, H.; Cousens, S.; Oestergaard, M.Z.; Chou, D.; Moller, A.B.; Narwal, R.; Adler, A.; Vera Garcia, C.; Rohde, S.; Say, L.; et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: A systematic analysis and implications. *Lancet* **2012**, *379*, 2162–272. [CrossRef]
7. Finnström, O.; Olausson, P.O.; Sedin, G.; Serenius, F.; Svenningsen, N.; Thiringer, K.; Tunell, R.; Wennergren, M.; Wesström, G. The Swedish national prospective study on extremely low birthweight (ELBW) infants. Incidence, mortality, morbidity and survival in relation to level of care. *Acta Paediatr.* **1997**, *86*, 503–511. [CrossRef] [PubMed]

8. Villar, J.; Papageorghiou, A.T.; Knight, H.E.; Gravett, M.G.; Iams, J.; Waller, S.A.; Kramer, M.; Culhane, J.F.; Barros, F.C.; Conde-Agudelo, A.; et al. The preterm birth syndrome: A prototype phenotypic classification. *Am. J. Obstet. Gynecol.* **2012**, *206*, 119–123. [CrossRef] [PubMed]
9. Goldenberg, R.L.; Gravett, M.G.; Iams, J.; Papageorghiou, A.T.; Waller, S.A.; Kramer, M.; Culhane, J.; Barros, F.; Conde-Agudelo, A.; Bhutta, Z.A.; et al. The preterm birth syndrome: Issues to consider in creating a classification system. *Am. J. Obstet. Gynecol.* **2012**, *206*, 113–118. [CrossRef] [PubMed]
10. Berghella, V.; Tolosa, J.E.; Kuhlman, K.; Weiner, S.; Bolognese, R.J.; Wapner, R.J. Cervical ultrasonography compared with manual examination as a predictor of preterm delivery. *Am. J. Obstet. Gynecol.* **1997**, *177*, 723–730. [CrossRef]
11. Berghella, V.; Kuhlman, K.; Weiner, S.; Texeira, L.; Wapner, R.J. Cervical funneling: Sonographic criteria predictive of preterm delivery. *Ultrasound Obstet. Gynecol.* **1997**, *10*, 161–166. [CrossRef] [PubMed]
12. Owen, J.; Yost, N.; Berghella, V.; MacPherson, C.; Swain, M.; Dildy, G.A., 3rd; Miodovnik, M.; Langer, O.; Sibai, B. Maternal-Fetal Medicine Units Network. Can shortened midtrimester cervical length predict very early spontaneous preterm birth? *Am. J. Obstet. Gynecol.* **2004**, *191*, 298–303. [CrossRef] [PubMed]
13. Drakeley, A.J.; Roberts, D.; Alfirevic, Z. Cervical cerclage for prevention of preterm delivery: Meta-analysis of randomized trials. *Obstet. Gynecol.* **2003**, *102*, 621–627. [CrossRef] [PubMed]
14. Arabin, B.; Halbesma, J.R.; Vork, F.; Hübener, M.; van Eyck, J. Is treatment with vaginal pessaries an option in patients with a sonographically detected short cervix? *J. Perinat. Med.* **2003**, *31*, 122–133. [CrossRef] [PubMed]
15. Iams, J.D. National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network What have we learned about uterine contractions and preterm birth? The HUAM Prediction Study. *Semin Perinatol.* **2003**, *27*, 204–211. [CrossRef]
16. Miller, E.S.; Tita, A.T.; Grobman, W.A. Second-Trimester Cervical Length Screening Among Asymptomatic Women: An Evaluation of Risk-Based Strategies. *Obstet. Gynecol.* **2015**, *126*, 61–66. [CrossRef] [PubMed]
17. Shennan, A.; Jones, B. The cervix and prematurity: Aetiology, prediction and prevention. *Semin. Fetal Neonatal Med.* **2004**, *9*, 471–479. [CrossRef] [PubMed]
18. Westergren-Thorsson, G.; Norman, M.; Björnsson, S.; Endrésen, U.; Stjernholm, Y.; Ekman, G.; Malmström, A. Differential expressions of mRNA for proteoglycans, collagens and transforming growth factor-beta in the human cervix during pregnancy and involution. *Biochim. Biophys. Acta* **1998**, *1406*, 203–213. [CrossRef]
19. Iwahashi, M.; Muragaki, Y.; Ooshima, A.; Umesaki, N. Decreased type I collagen expression in human uterine cervix during pregnancy. *J. Clin. Endocrinol. Metab.* **2003**, *88*, 2231–2235. [CrossRef] [PubMed]
20. Törnblom, S.A.; Klimaviciute, A.; Byström, B.; Chromek, M.; Brauner, A.; Ekman-Ordeberg, G. Non-infected preterm parturition is related to increased concentrations of IL-6, IL-8 and MCP-1 in human cervix. *Reprod. Biol. Endocrinol.* **2005**, *3*, 39. [CrossRef] [PubMed]
21. Törnblom, S.A.; Maul, H.; Klimaviciute, A.; Garfield, R.E.; Byström, B.; Malmström, A.; Ekman-Ordeberg, G. mRNA expression and localization of bNOS, eNOS and iNOS in human cervix at preterm and term labour. *Reprod. Biol. Endocrinol.* **2005**, *3*, 33. [CrossRef] [PubMed]
22. Honest, H.; Bachmann, L.M.; Coomarasamy, A.; Gupta, J.K.; Kleijnen, J.; Khan, K.S. Accuracy of cervical transvaginal sonography in predicting preterm birth: A systematic review. *Ultrasound Obstet. Gynecol.* **2003**, *22*, 305–322. [CrossRef] [PubMed]
23. Li, Q.; Reeves, M.; Owen, J.; Keith, L.G. Precocious cervical ripening as a screening target to predict spontaneous preterm delivery among asymptomatic singleton pregnancies: A systematic review. *Am. J. Obstet. Gynecol.* **2015**, *212*, 145–156. [CrossRef] [PubMed]
24. Spong, C.Y. Prediction and prevention of recurrent spontaneous preterm birth. *Obstet. Gynecol.* **2007**, *110*, 405–415. [CrossRef] [PubMed]
25. Owen, J.; Yost, N.; Berghella, V.; Thom, E.; Swain, M.; Dildy, G.A., 3rd; Miodovnik, M.; Langer, O.; Sibai, B.; McNellis, D. National Institute of Child Health and Human Development, Maternal-Fetal Medicine Units Network. Mid-trimester endovaginal sonography in women at high risk for spontaneous preterm birth. *JAMA* **2001**, *286*, 1340–1348. [CrossRef] [PubMed]
26. Owen, J.; Szychowski, J.M.; Hankins, G.; Iams, J.D.; Sheffield, J.S.; Perez-Delboy, A.; Berghella, V.; Wing, D.A.; Guzman, E.R.; Vaginal Ultrasound Trial Consortium. Does midtrimester cervical length $\geq$25 mm predict preterm birth in high-risk women? *Am. J. Obstet. Gynecol.* **2010**, *203*, 393e1–393e5. [CrossRef] [PubMed]

27. Goldenberg, R.L.; Iams, J.D.; Miodovnik, M.; Van Dorsten, J.P.; Thurnau, G.; Bottoms, S.; Mercer, B.M.; Meis, P.J.; Moawad, A.H.; Das, A.; et al. The preterm prediction study: Risk factors in twin gestations. National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. *Am. J. Obstet. Gynecol.* **1996**, *175*, 1047–1053. [CrossRef]
28. To, M.S.; Skentou, C.; Cicero, S.; Liao, A.W.; Nicolaides, K.H. Cervical length at 23 weeks in triplets: Prediction of spontaneous preterm delivery. *Ultrasound Obstet. Gynecol.* **2000**, *16*, 515–518. [CrossRef] [PubMed]
29. Mancuso, M.S.; Szychowski, J.M.; Owen, J.; Hankins, G.; Iams, J.D.; Sheffield, J.S.; Perez-Delboy, A.; Berghella, V.; Wing, D.A.; Guzman, E.R.; et al. Cervical funneling: Effect on gestational length and ultrasound-indicated cerclage in high-risk women. *Am. J. Obstet. Gynecol.* **2010**, *203*, 259e1–259e5. [CrossRef] [PubMed]
30. Tsikouras, P.; Galazios, G.; Zalvanos, A.; Bouzaki, A.; Athanasiadis, A. Transvaginal sonographic assessment of the cervix and preterm labor. *Clin. Exp. Obstet. Gynecol.* **2007**, *34*, 159–162. [PubMed]
31. Mella, M.T.; Mackeen, A.D.; Gache, D.; Baxter, J.K.; Berghella, V. The utility of screening for historical risk factors for preterm birth in women with known second trimester cervical length. *J. Matern. Fetal Neonatal Med.* **2013**, *26*, 710–715. [CrossRef] [PubMed]
32. Subramaniam, A.; Harper, L.M.; Szychowski, J.M.; Owen, J. Predictive Value of Initial Cervical Length for Subsequent Cervical Length Shortening in Women with a Prior Preterm Birth. *Am. J. Perinatol.* **2016**, *33*, 350–355. [PubMed]
33. Orzechowski, K.M.; Boelig, R.; Nicholas, S.S.; Baxter, J.; Berghella, V. Is universal cervical length screening indicated in women with prior term birth? *Am. J. Obstet. Gynecol.* **2015**, *212*, 234e1–234e5. [CrossRef] [PubMed]
34. Geitona, M.; Hatzikou, M.; Hatzistamatiou, Z.; Anastasiadou, A.; Theodoratou, T.D. The economic burden of treating neonates in Intensive Care Units (ICUs) in Greece. *Cost Eff. Resour. Alloc.* **2007**, *5*, 9. [CrossRef] [PubMed]
35. Boelig, R.C.; Orzechowski, K.M.; Berghella, V. Cervical length, risk factors, and delivery outcomes among women with spontaneous preterm birth. *J. Matern. Fetal Neonatal Med.* **2016**, *29*, 2840–2844. [CrossRef] [PubMed]
36. Kuusela, P.; Jacobsson, B.; Söderlund, M.; Bejlum, C.; Almström, E.; Ladfors, L.; Hagberg, H.; Wennerholm, U.B. Transvaginal sonographic evaluation of cervical length in the second trimester of asymptomatic singleton pregnancies, and the risk of preterm delivery. *Acta Obstet. Gynecol. Scand.* **2015**, *94*, 598–607. [CrossRef] [PubMed]
37. Szychowski, J.M.; Owen, J.; Hankins, G.; Iams, J.D.; Sheffield, J.S.; Perez-Delboy, A.; Berghella, V.; Wing, D.A.; Guzman, E.R. Vaginal Ultrasound Trial Consortium Can the optimal cervical length for placing ultrasound-indicated cerclage be identified? *Ultrasound Obstet. Gynecol.* **2016**, *48*, 43–47. [CrossRef] [PubMed]
38. Chandiramani, M.; Shennan, A. Preterm labour: Update on prediction and prevention strategies. *Curr. Opin. Obstet. Gynecol.* **2006**, *18*, 618–624. [CrossRef] [PubMed]
39. Fonseca, E.B.; Celik, E.; Parra, M.; Singh, M.; Nicolaides, K.H. Fetal Medicine Foundation Second Trimester Screening Group. Progesterone and the risk of preterm birth among women with a short cervix. *N. Engl. J. Med.* **2007**, *357*, 462–469. [CrossRef] [PubMed]
40. Kyvernitakis, I.; Khatib, R.; Stricker, N.; Arabin, B. Is Early Treatment with a Cervical Pessary an Option in Patients with a History of Surgical Conisation and a Short Cervix? *Geburtshilfe Frauenheilkd.* **2014**, *74*, 1003–1008. [CrossRef] [PubMed]
41. Goya, M.; Pratcorona, L.; Merced, C.; Rodó, C.; Valle, L.; Romero, A.; Juan, M.; Rodríguez, A.; Muñoz, B.; Santacruz, B.; et al. Pesario Cervical para Evitar Prematuridad (PECEP) Trial GroupCervical pessary in pregnant women with a short cervix (PECEP): An open-label randomised controlled trial. *Lancet* **2012**, *379*, 1800–1806. [CrossRef]
42. Liem, S.; Schuit, E.; Hegeman, M.; Bais, J.; de Boer, K.; Bloemenkamp, K.; Brons, J.; Duvekot, H.; Bijvank, B.N.; Franssen, M.; et al. Cervical pessaries for prevention of preterm birth in women with a multiple pregnancy (ProTWIN): A multicentre, open-label randomised controlled trial. *Lancet* **2013**, *382*, 1341–1349. [CrossRef]
43. Barinov, S.V.; Shamina, I.V.; Lazareva, O.V.; Tirskaya, Y.I.; Ralko, V.V.; Shkabarnya, L.L.; Dikke, G.B.; Kochev, D.M.; Klementyeva, L.L. Comparative assessment of arabin pessary, cervical cerclage and medical management for preterm birth prevention in high-risk pregnancies. *J. Matern. Fetal Neonatal Med.* **2017**, *30*, 1841–1846. [CrossRef] [PubMed]

44. Mendoza, M.; Goya, M.; Gascón, A.; Pratcorona, L.; Merced, C.; Rodó, C.; Valle, L.; Romero, A.; Juan, M.; Rodríguez, A.; et al. Modification of cervical length after cervical pessary insertion: Correlation weeks of gestation. *J. Matern. Fetal Neonatal Med.* **2017**, *30*, 1596–1601. [CrossRef] [PubMed]
45. Mateus, J. Clinical management of the short cervix. *Obstet. Gynecol. Clin. N. Am.* **2011**, *38*, 367–385. [CrossRef] [PubMed]
46. Nam, K.H.; Kwon, J.Y.; Kim, Y.H.; Park, Y.W. Pregnancy outcome after cervical conization: Risk factors for preterm delivery and the efficacy of prophylactic cerclage. *J. Gynecol. Oncol.* **2010**, *21*, 225–229. [CrossRef] [PubMed]
47. Berghella, V.; Ciardulli, A.; Rust, O.A.; To, M.; Otsuki, K.; Althuisius, S.; Nicolaides, K.; Roman, A.; Saccone, G. Cerclage for Short Cervix on Ultrasound in Singleton Gestations without Prior Spontaneous Preterm Birth: A Systematic Review and Meta-analysis of Trials using individual patient-level data. *Ultrasound Obstet. Gynecol.* **2017**, *50*, 569–577. [CrossRef] [PubMed]
48. Norman, J.E. Preterm labour. Cervical function and prematurity. *Best Pract. Res. Clin. Obstet. Gynaecol.* **2007**, *21*, 791–806. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

Exposure to Night-Time Traffic Noise, Melatonin-Regulating Gene Variants and Change in Glycemia in Adults

Ikenna C. Eze [1,2,*], Medea Imboden [1,2], Maria Foraster [1,2], Emmanuel Schaffner [1,2], Ashish Kumar [1,2,3], Danielle Vienneau [1,2], Harris Héritier [1,2], Franziska Rudzik [4], Laurie Thiesse [4], Reto Pieren [5], Arnold von Eckardstein [6], Christian Schindler [1,2], Mark Brink [7], Jean-Marc Wunderli [5], Christian Cajochen [4], Martin Röösli [1,2] and Nicole Probst-Hensch [1,2]

1. Swiss Tropical and Public Health Institute, Socinstrasse 57, P.O. Box, CH-4002 Basel, Switzerland; medea.imboden@swisstph.ch (M.I.); maria.foraster@swisstph.ch (M.F.); emmanuel.schaffner@swisstph.ch (E.S.); ashish.kumar@unibas.ch (A.K.); danielle.vienneau@swisstph.ch (D.V.); harris.heritier@swisstph.ch (H.H.); christian.schindler@swisstph.ch (C.S.); martin.roosli@swisstph.ch (M.R.); nicole.probst@swisstph.ch (N.P.-H.)
2. University of Basel, Petersplatz 1, CH-4003 Basel, Switzerland
3. Karolinska Institutet, SE-171 77 Stockholm, Sweden
4. Center for Chronobiology, Psychiatric Hospital of the University of Basel, Wilhelm Klein-Strasse 27, 4056 Basel, Switzerland; franziska.rudzik@upkbs.ch (F.R.); laurie.thiesse@upkbs.ch (L.T.); christian.cajochen@upkbs.ch (C.C.)
5. Empa, Laboratory for Acoustics/Noise Control, Swiss Federal Laboratories for Material Science and Technology, Überlandstrasse 129, CH-8600 Dübendorf, Switzerland; reto.pieren@empa.ch (R.P.); jean-marc.Wunderli@empa.ch (J.-M.W.)
6. Institute of Clinical Chemistry, University Hospital Zurich, Rämistrasse 100, 8091 Zurich, Switzerland; arnold.voneckardstein@usz.ch
7. Federal Office for the Environment, 3003 Bern, Switzerland; mark.brink@bafu.admin.ch

* Correspondence: ikenna.eze@swisstph.ch; Tel.: +41-61-284-8395

Received: 8 November 2017; Accepted: 29 November 2017; Published: 1 December 2017

Abstract: Traffic noise has been linked to diabetes, with limited understanding of its mechanisms. We hypothesize that night-time road traffic noise (RTN) may impair glucose homeostasis through circadian rhythm disturbances. We prospectively investigated the relationship between residential night-time RTN and subsequent eight-year change in glycosylated hemoglobin (ΔHbA1c) in 3350 participants of the Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults (SAPALDIA), adjusting for diabetes risk factors and air pollution levels. Annual average RTN (Lnight) was assigned to participants in 2001 using validated Swiss noise models. HbA1c was measured in 2002 and 2011 using liquid chromatography. We applied mixed linear models to explore RTN–ΔHbA1c association and its modification by a genetic risk score of six common circadian-related *MTNR1B* variants (MGRS). A 10 dB difference in RTN was associated with a 0.02% (0.003–0.04%) increase in mean ΔHbA_{1c} in 2142 non-movers. RTN–ΔHbA1c association was modified by MGRS among diabetic participants ($P_{interaction} = 0.001$). A similar trend in non-diabetic participants was non-significant. Among the single variants, we observed strongest interactions with rs10830963, an acknowledged diabetes risk variant also implicated in melatonin profile dysregulation. Night-time RTN may impair glycemic control, especially in diabetic individuals, through circadian rhythm disturbances. Experimental sleep studies are needed to test whether noise control may help individuals to attain optimal glycemic levels.

Keywords: transportation noise; *MTNR1B* gene; rs10830963; diabetes; glycosylated hemoglobin; circadian sleep-wake cycle; gene-environment interactions; adults; cohort study

1. Introduction

Glycosylated hemoglobin (HbA1c) reflects three-month average glycemia [1], predicts diabetes in non-diabetic individuals, and complications in diabetic individuals [2,3]. Despite control efforts, the prevalence of poorly-controlled diabetes remains high [4]. Recent interest in the role of environmental stressors in cardio-metabolic diseases towards improved prevention [5] makes it pertinent to explore the environmental determinants of glycemic control. Hyperglycemia is an important risk factor for cardiovascular diseases [6–8] and diabetes [2,3]. Understanding the role of environmental stressors also entails investigations of their mediating or modifying mechanisms, in persons with and without diabetes.

Road traffic noise (RTN) represents the most common transportation noise, and together with air pollution, constitutes more than 75% of environmental disease burden in Europe [9]. Growing evidence suggests a link between RTN and diabetes morbidity [10–14] and mortality [15] (independent of traffic-related air pollution), with unclear mechanisms [5,16].

Melatonin, a pineal hormone known to act on numerous organs and co-regulate various neural and endocrine processes, exhibits a circadian rhythm which is closely linked to sleep propensity [17] and insulin secretion [18]. Administration of exogenous melatonin elicited subjective sleepiness [19] and glucose intolerance when given close to mealtime [20]. Melatonin mediates its regulatory function through binding to receptors including the high-affinity melatonin receptor 1B encoded by *MTNR1B*, a gene widely expressed in the human retina and pancreatic islet cells. Recent genome-wide association studies (GWAS) showed seven common variants near *MTNR1B* to be associated with glucose homeostasis or type 2 diabetes, mainly through impairments in insulin secretion [21]. The glucose-raising allele of the lead variant, rs10830963, was shown to significantly prolong the duration of elevated melatonin levels by 41 min, and delayed the circadian phase of dimlight melatonin offset by 1.37 h, as well as increased diabetes risk among individuals who woke up early [22]. Individuals who are evening chronotypes were also shown to have higher risk of diabetes [23]. In both instances, it is thought that waking up early or sleeping late (during high melatonin and low insulin phase) predisposes to having meals, leads to high insulin levels and subsequent hyperglycemia, as a result of altered timing of food intake. Thus, the glucose-raising allele of rs10830963 could modify susceptibility to noise exposure in cases of early awakenings during melatonin secretion (Figure 1).

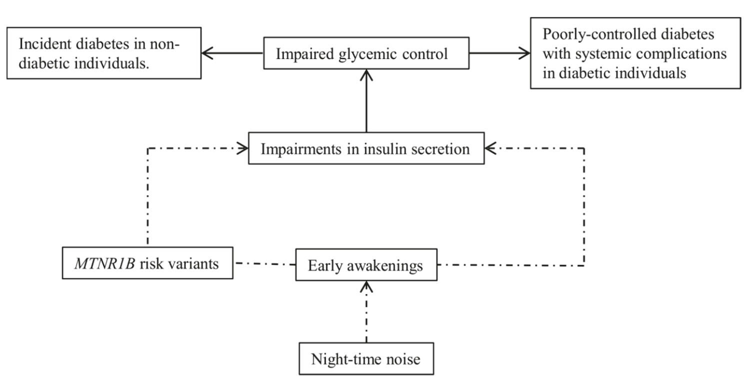


Figure 1. Hypothesized mechanisms of noise effects on glucose homeostasis explored in the present study.

Night-time noise exposure could impair glucose homeostasis through disturbances in circadian sleep–wake cycles. Circadian pathway was captured using melatonin-regulating genetic variants as a proxy, possibly reflecting genetic risk for melatonin profile dysregulation.

Although there are at least seven common variants on *MTNR1B* influencing glucose metabolism through a common pathway [21], sleep and circadian studies have only focused on rs10830963 variant [24]. In essence, following the approach of using scores involving multiple variants for improved prediction of genetic risks [25,26], an *MTNR1B* genetic risk score (MGRS) based on variants that are not in strong linkage disequilibrium (LD), may better capture individuals at high-risk of melatonin profile dysregulation than single variants.

Understanding the implications of noise exposure on glucose homeostasis through interactions with circadian-related parameters will improve our understanding and offer preventive channels towards optimal glucose control in diabetic and non-diabetic individuals.

We therefore tested the hypothesis that MGRS (and component single variants representing the circadian pathway) may modify the potential association between night-time RTN and eight-year change in HbA1c (ΔHbA1c) in diabetic and non-diabetic participants of the Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults (SAPALDIA).

2. Materials and Methods

2.1. Study Population

Participants included 3350 adults aged 29–81 years who completed health interviews and examinations at the first (SAP2) and second (SAP3) SAPALDIA follow-up surveys, and had complete information on relevant covariates. Details of inclusion are shown in Figure S1. The SAPALDIA study began in 1991 (SAP1), with 9651 randomly-selected adults from eight geographically-representative Swiss areas [27]. Two follow-up surveys in 2001/2002 and 2010/2011 retained 8047 and 6088 participants, respectively. At each survey, participants responded to detailed questions concerning their health and lifestyle. At SAP2 and SAP3, blood was sampled into a biobank for biomarker assays, including HbA1c and genotyping. Participants provided written informed consent and ethical clearances were obtained from the Ethics Committees of the Swiss Academy of Medical Sciences and the participating cantons (National Ethics Committee for Clinical Research (UREK); Project Approval Number 123/00; date of approval: September 2001).

2.2. Measurement of HbA1c and Identification of Diabetes Cases

HbA1c was measured in EDTA-buffered whole blood samples collected at SAP2 and SAP3 using the ARK-RAY ADAMS A1c HA-8180V Analyser (Menarini, Florence, Italy) based on high-performance liquid chromatography, which has minimal interference from alternate hemoglobin variants [28]. HbA1c values were measured in mmol/mol, and converted into percentage [29]. Using a combination of questionnaire data and HbA1c values, we identified participants as diabetes cases if they (I) self-reported physician-diagnosed diabetes, (II) used diabetes medication, or (III) had a HbA1c value $\geq$6.5% at SAP2 or SAP3. We also defined confirmed/advanced diabetes, restricted to only participants taking medication, thus, yielding three comparative groups: no diabetes (at SAP2 and SAP3), diabetes (at SAP2 or SAP3), and diabetes on medication (at SAP2 or SAP3). We defined ΔHbA1c as the absolute difference between HbA1c at SAP3 (2011) and HbA1c at SAP2 (2002).

2.3. Exposures

Average annual exposure to RTN was assigned to participants based on the most exposed façade of their residential floors in 2001 and 2011 (corresponding to the survey time points) using the SonROAD emission and STL-86 propagation models combining high-resolution spatial and temporal road traffic information, as previously described [30]. This was done in the framework of the SiRENE (Short and Long-Term Effects of Transportation Noise Exposure) project, where railway and aircraft noise were also assigned to the same façade using sonRAIL propagation and SEMIBEL emission models for railway noise, and FLULA2 simulation model for aircraft noise—all validated Swiss noise models [30]. Day (Lday; 07–23 h), night (Lnight; 23–07 h), and day–evening–night (Lden) noise (dB) were computed for road, railway, and aircraft sources. Participants without substantial night-time railway/aircraft exposures were assigned truncated values of 20 dB, and were assigned a truncation indicator. Lnight was highly correlated with constituent time points (23–01 h, 01–05 h and 05–07 h; Spearman R > 0.9). Day and night-time correlations (Spearman R) of road traffic, aircraft, and railway noise were 0.99, 0.33, and 0.93, respectively. Since our study was aimed at modification of noise–HbA1c association by circadian-related parameters, and RTN is the most common noise with consistently reported associations with diabetes mortality/morbidity [12], we focused our analyses on night-time RTN, considering night-time railway and aircraft noise as potential confounders, with secondary exploration of their main associations. Noise levels were quite stable over follow-up (Spearman R > 0.9 for both time points), thus, the noise levels at 2001 would capture long-term exposures in statistical models predicting health outcomes after 2001. To limit exposure misclassification to an extent, we also restricted analyses to SAPALDIA participants who did not change their residence during the follow-up period.

2.4. Potential Confounders

From the questionnaire data at SAP2, we extracted additional potential confounders of the association between RTN at SAP2 and change in HbA1c between SAP2 and SAP3. These included age (continuous), sex (male/female), formal education (≤9 (primary level)/10–13 (secondary, apprenticeship level)/>13 years (tertiary level)), neighborhood socio-economic index (SEI) (continuous; incorporating education and income of household head, median rent and crowding of households within the neighborhood) [31], and green areas within 2 km residential buffer, available from the European Environment Agency hectare resolution dataset (continuous; CORINE CLC-2006 Version 13). We also considered smoking status (never/former/current), exposure to passive smoke (yes/no), consumption of alcohol (including liquor, wines, beers and spirits; ≤1/>1 glass/day), and body mass index (BMI), defined as weight and height-squared ratio (continuous). Towards sensitivity analyses, we extracted participants' responses to the question "do you often feel that you have not slept enough after you wake up in the morning?" from which we derived insufficient sleep (yes/no), as well as their responses to the Short Form 36-item mental health questionnaire (continuous) [32]. In SAP2, average annual residential exposure to nitrogen dioxide (NO_2), a marker of road traffic-related air pollution and potential confounder of RTN [33], was assigned to participants' residences based on a hybrid model (adjusted $R^2 = 0.83$) derived by fitting NO_2 sampler measurements using land-use and traffic variables, and predictions from a Gaussian dispersion model as previously described [34].

2.5. MTNR1B Gene Score

Genomic DNA was extracted from EDTA-buffered whole blood using Puragene™ DNA Isolation Kit (Gentra Systems, Plymouth, UK). Genotyping of SAPALDIA participants was done on DNA samples collected at SAP2. In the framework of the EU-funded GABRIEL consortium to identify genetic determinants of asthma [35], 1612 participants were genotyped using Illumina Human 610Kquad BeadChip (G1; Illumina, San Diego, CA, USA) covering ~570,000 variants. An additional 3015 participants were genotyped using Illumina Human OmniExpress-Exome BeadChip (G2; Illumina, San Diego, CA, USA), covering ~1 million variants. Quality control criteria were applied to both genotyping arrays: samples with <97% genotyping success rate, or of non-European origin, with cryptic relatedness or sex inconsistencies, were excluded. Variants with minor allele frequency (MAF) of <5% or deviation from Hardy–Weinberg equilibrium (HWE) at a threshold of 10^{-6} were also excluded. G1 and G2 datasets were phased using ShapeIT version 2.r790 [36] and imputed using MiniMac2 [37]. The imputed datasets were then merged, after excluding variants with low imputation quality ($R^2 < 0.3$), yielding ~14 million markers for 4324 participants across G1 and G2, from where we identified seven common *MTNR1B* variants (rs1387153, rs10830962, rs4753426, rs8192552, rs10830963, rs3781638 and rs2166706) involved in glucose dysregulation [21]. Imputations were of high quality. All but rs10830963 ($R^2 \geq 0.87$) had an imputation $R^2 \geq 0.92$. All variants had similar allele frequencies in comparison with other studies [38–49] and the 1000 Genomes Central European Population [50]. All variants were in HWE ($p > 0.2$), with MAF ≥ 7% (Table S1). Two variants, rs10830962 and rs2166706, were in high LD ($R^2 = 0.89$; Figure S2), thus, we excluded rs2166706 from the analyses. The LD R^2 of the included MTNR1B variants ranged between 0.2 and 0.7. Since the glucose-raising allele of rs10830963 was the allele implicated in melatonin profile dysregulation, expressed as a significant delay in melatonin offset [22], we coded the other variants in their reported direction of association in glucose alterations, such that each variant contained 0–2 quantities of the risk allele. We created MGRS by summing up risk alleles across the six variants, yielding a minimum, median, and maximum of 2, 6, and 12 risk alleles, respectively. We also created a categorical variable—low risk (MGRS ≤ 6) and high risk (>6).

2.6. Statistical Analyses

We summarized the characteristics of participants by inclusion and exclusion status. Differences in proportions and means were tested using chi-squared and *t*-tests respectively. We built predictive statistical models using noise exposure and covariates measured at SAP2, for ΔHbA1c. Both crude and adjusted associations between MGRS (as well as the single variants, in separate models) and ΔHbA1c; and between RTN and ΔHbA1c were assessed using mixed linear models with a random intercept at the level of study area. Adjusted models included age, sex, education, neighborhood SEI, smoking status, alcohol consumption, and BMI. Adjusted RTN–ΔHbA1c models additionally included NO_2, railway and aircraft noise, noise truncation indicators, and green space within a 2 km residential buffer. We included interaction terms between RTN and MGRS (as well as the single variants, in separate models), in the RTN–ΔHbA1c models, to explore the presence of interactions between these variables. We also stratified the RTN–ΔHbA1c model by categories of MGRS. Furthermore, we restricted analyses to participants who did not change their residence during follow-up. Since our study included only 35% of SAP1 participants, we limited potential selection bias by applying the inverse of the probability of participating in present analyses as weights in our models. These probability weights were derived from a logistic regression model using predictor variables from SAP1. We performed sensitivity analyses: we explored cross-sectional associations between MGRS (as well as the single variants, in separate models) and HbA1c using repeated mixed linear models with random intercepts at the level of participants. Using the RTN–ΔHbA1c model, we tested sensitivity to removal of BMI from the main and interaction models, performed complete case analyses without adjusting for potential selection bias, and excluded asthmatic participants in the RTN–MGRS interaction model to explore potential genotyping selection bias. We explored interactions with self-reported sleep insufficiency reported at SAP2. All analyses were stratified into three categories—no diabetes, diabetes, and diabetes on medication—to limit confounding/effect modification by diabetes status or medication. All analyses were done using STATA version 14 (STATA Corporation, College Station, TX, USA). Statistical significance was defined at two-sided alpha-values of 0.05 and 0.1 for main associations and interactions, respectively.

3. Results

3.1. Characteristics of Participants

Over an approximate eight-year follow-up, the mean increase in HbA1c was 0.04%, whereas diabetes prevalence was 3% at SAP2 and 7% at SAP3, with 8% combined prevalence in the 3350 included participants. On average, diabetic participants gained more HbA1c than non-diabetic participants (Table 1), in line with evidence for poor glucose control in Switzerland [51].

Table 1. Characteristics of study participants included in the study.

Categorical Variables	(%)
Females at SAP2	48
Formal education at SAP2: ≤9 years	4
10–13 years	64
>13 years	32
Ever-smoker at SAP2	55
Exposure to passive smoke at SAP2	23
Alcohol consumption >once/day at SAP2	40
Sleep insufficiency at SAP2	28
Diabetes status at SAP2	3
Diabetes status at SAP3	7
Diabetes medication at SAP2	3
Diabetes medication at SAP3	5
Change of residence (movers) between SAP2 and SAP3	36
Continuous Variables (Units)	**Mean (SD)**
Age at SAP2 (years)	51 (11)
Body mass index at SAP2 (kg/m^2)	25.6 (4)
Neighborhood socio-economic index at SAP2 (%)	64 (10)
SF-36 mental health score at SAP2 (%) [a]	76 (15)
Lnight, road at SAP2 (dB)	45.8 (8)
Lnight, railway at SAP2 (dB)	28.3 (10)
Lnight, aircraft at SAP2 (dB)	22.7 (6)
Nitrogen dioxide at SAP2 (μg/m^3)	22.4 (10)
Green space within 2 km radius at SAP2 (km^2)	0.3 (0.4)
MTNR1B variants at SAP2: rs1387153, T allele	0.6 (0.6)
rs10830962, G allele	0.8 (0.7)
rs4753426, C allele	1.0 (0.7)
rs8192552, G allele	1.9 (0.4)
rs10830963, G allele	0.6 (0.6)
rs37816938, G allele	1.1 (0.7)
MTNR1B genetic risk score at SAP2	6 (3)
Glycosylated hemoglobin, HbA1c at SAP2 (%)	5.2 (0.4)
HbA1c at SAP3 (%)	5.3 (0.5)
Change in HbA1c between SAP2 and SAP3 (%): All	0.04 (0.4)
Non-diabetic participants	0.03 (0.3)
Diabetic participants	0.17 (1.1)
Diabetic participants on medication	0.12 (1.2)

[a] N (included) = 2933.

SAP2 and SAP3 represent the first and second follow-up surveys of the Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults in 2002 and 2010/2011, respectively. Lnight represents night-time (23–07 h) noise levels assigned to the SAPALDIA participants based on the most exposed façade of their residential floors. Number of included participants is 3350.

Mean (SD) annual exposures to night-time road traffic, railway, and aircraft noise at SAP2 were 46 (8) dB, 28 (10) dB, and 23 (6) dB, respectively. Figure S3 shows the distribution of noise exposures in the included participants. Mean (SD) MGRS was 6 (3) risk alleles. Compared to the participants excluded due to non-participation during follow-up or missing data, included participants were younger, more likely males, better educated, and less likely to be overweight or diabetic. They also had lower exposure to RTN and NO_2, but higher exposure to aircraft noise. There were no significant differences in ΔHbA1c, MGRS, and change of residence between both groups (Table S2).

3.2. Associations between RTN and ΔHbA1c

In non-diabetics, we observed generally non-significant associations between transportation noise and ΔHbA1c. Among diabetics, associations with railway and aircraft noise were positive (reaching significance only for aircraft noise), whereas associations with road traffic noise were negative. Limiting analyses to non-movers revealed consistent positive associations between RTN and ΔHbA1c among both diabetic and non-diabetic participants. Among non-diabetic non-movers, mean ΔHbA1c showed a statistically significant increase by 0.01% (95% CI 0, 0.03) per 10 dB exposure to RTN. Among diabetic non-movers, associations were reversed compared to all participants, became positive, and were stronger than we observed in non-diabetic participants. Associations also reached ten-fold in those reporting the use of diabetic medication compared to non-diabetics, but were not significant (Table 2).

All estimates represent increase (+) or decrease (−) in mean change in HbA1c per 10 dB of night-time road traffic noise. Adjusted models included age, sex, education, neighborhood socio-economic index, smoking status, passive smoking, alcohol consumption, green space within a 2 km residential buffer, residential levels of nitrogen dioxide, night-time railway, aircraft noise and their truncation indicators. All models include random intercepts at the level of the study areas, and were adjusted for potential selection bias by applying the probability of participation in present analyses as weights derived from a logistic regression with predictors from the baseline study in 1991.

In the adjusted models, we also found significantly positive associations of ΔHbA1c with aircraft noise in non-diabetics, where mean HbA1c increased by 0.02% (95% CI 0, 0.03) per 10 dB difference in Lnight. There were no significant associations with railway noise. Similar to RTN, associations of ΔHbA1c with aircraft noise in non-movers were consistently positive across comparison groups (Table 2). All models were stable to confounder adjustments, including BMI. Irrespective of subpopulation studied or confounder adjustments, we did not observe any association of railway noise with ΔHbA1c.

Table 2. Exposure to night-time transportation noise in 2001 (10 dB difference) and subsequent change in HbA1c (%).

All Participants	Source	Model	All N = 3350	No Diabetes N = 3098	Diabetes N = 251	Diabetes on Medication N = 168
			β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)
	Road	Crude	0.01 (−0.01, 0.03)	0.01 (0.002, 0.02) *	−0.02 (−0.22, 0.18)	−0.03 (−0.25, 0.18)
		Adjusted	0.01 (−0.02, 0.03)	0.01 (−0.0004, 0.02) †	−0.07 (−0.26, 0.11)	−0.13 (−0.36, 0.10)
		Adjusted+BMI	0.01 (−0.02, 0.03)	0.01 (−0.001, 0.02) †	−0.08 (−0.27, 0.11)	−0.11 (−0.35, 0.15)
	Railway	Crude	0.004 (−0.01, 0.02)	0.003 (−0.01, 0.01)	0.03 (−0.06, 0.12)	0.03 (−0.11, 0.16)
		Adjusted	0.001 (−0.01, 0.02)	0.001 (−0.01, 0.01)	0.01 (−0.05, 0.08)	0.05 (−0.05, 0.13)
		Adjusted+BMI	0.002 (−0.01, 0.02)	0.001 (−0.01, 0.01)	0.02 (−0.05, 0.08)	0.04 (−0.04, 0.12)
	Aircraft	Crude	−0.0003 (−0.01, 0.01)	−0.01 (−0.03, 0.01)	0.24 (−0.08, 0.55)	0.26 (0.05, 0.47) *
		Adjusted	0.005 (−0.02, 0.03)	−0.004 (−0.03, 0.02)	0.34 (0.12, 0.55) †	0.34 (0.13, 0.54) †
		Adjusted+BMI	0.01 (−0.01, 0.02)	−0.003 (−0.02, 0.02)	0.35 (0.08, 0.62) †	0.31 (0.03, 0.60) †
Non-movers	**Source**	**Model**	**All N = 2142**	**No Diabetes N = 1960**	**Diabetes N = 179**	**Diabetes on Medication N = 117**
	Road	Crude	0.02 (0, 0.04) *	0.01 (0.0004, 0.03) *	0.06 (−0.21, 0.33)	0.07 (−0.26, 0.39)
		Adjusted	0.02 (0.003, 0.04) *	0.01 (0, 0.03) *	0.04 (−0.16, 0.24)	0.10 (−0.13, 0.33)
		Adjusted+BMI	0.02 (0.003, 0.04) *	0.01 (0, 0.03) *	0.03 (−0.14, 0.19)	0.10 (−0.09, 0.29)
	Railway	Crude	0.001 (−0.03, 0.03)	−0.01 (−0.02, 0.01)	0.06 (−0.15, 0.26)	0.05 (−0.31, 0.41)
		Adjusted	−0.001 (−0.03, 0.03)	−0.01 (−0.02, 0.01)	0.02 (−0.29, 0.31)	−0.03 (−0.53, 0.47)
		Adjusted+BMI	−0.001 (−0.03, 0.03)	−0.01 (−0.02, 0.01)	0.02 (−0.24, 0.29)	−0.03 (−0.46, 0.40)
	Aircraft	Crude	0.03 (0.01, 0.05)	0.01 (−0.01, 0.03)	0.24 (−0.04, 0.53) †	0.22 (−0.03, 0.47) †
		Adjusted	0.04 (0.01, 0.07) †	0.02 (−0.002, 0.03) †	0.30 (−0.15, 0.74)	0.14 (−0.44, 0.71)
		Adjusted+BMI	0.04 (0.01, 0.07) †	0.02 (0, 0.03) *	0.32 (−0.22, 0.86)	0.13 (−0.49, 0.74)

* $p < 0.05$; † $p < 0.1$.

3.3. Associations between MGRS and ΔHbA1c

Results of main associations between MGRS (and component variants) and HbA1c are presented in Table S4. In non-diabetic participants, MGRS showed a positive association with ΔHbA1c and cross-sectional HbA1c, but only reached significance in the cross-sectional analysis. All six single variants were positively associated with HbA1c, with associations of 0.01–0.03% per risk allele. Among diabetic participants, there was a non-significant tendency for associations of MGRS (and component variants) with ΔHbA1c and cross-sectional HbA1c to be negative.

3.4. Modification of RTN-ΔHbA1c Association by MGRS

Among non-movers, we observed significant interactions between MGRS and RTN that were restricted to diabetic participants. The interactions were stronger in persons who reported medication use where mean ΔHbA1c changed by 0.90% (0.31, 1.49%) in diabetic participants on medication with high MGRS, and by −0.32% (−0.50, −0.14%) per 10 dB, in those with low MGRS ($P_{\text{interaction}}$ = 0.001) (Figure 2). All single variants showed positive interaction with RTN in diabetic participants. The lead functional variant, rs10830963, showed the strongest significant interactions, where mean ΔHbA1c increased by 0.80% (0.14, 1.47%) and by 1.21% (0.59, 1.83%) per 10 dB and per risk allele, in diabetics and medicated diabetics, respectively (Table 3).

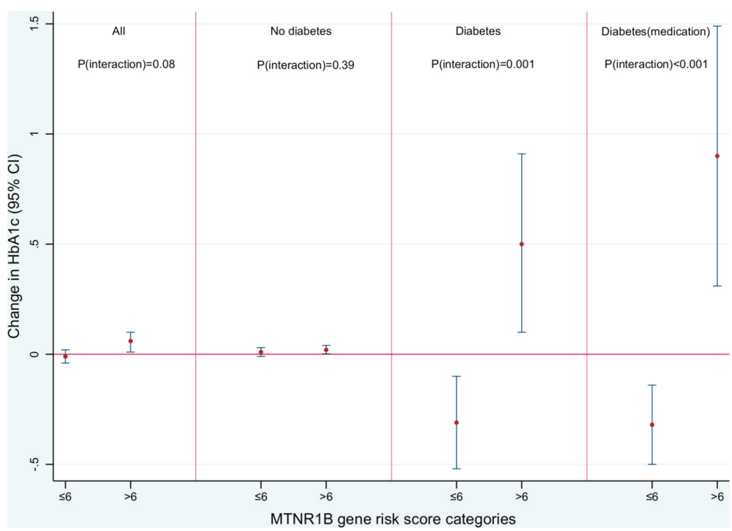


Figure 2. Modification of the association between night-time road traffic noise and change in HbA1c by melatonin dysregulation risk score, in non-movers.

Interaction terms included night-time road traffic noise and *MTNR1B* variants/score. *MTNR1B* score represents the sum of the risk alleles across six included variants. Positive sign of beta coefficient means increase in HbA1c per 10 dB change in night-time road traffic noise and per risk allele. All models were adjusted for age, sex, education, neighborhood socio-economic index, smoking status, passive smoking, alcohol consumption, green space within a 2 km residential buffer, residential levels of nitrogen dioxide, night-time railway, aircraft noise, and their truncation indicators. All models included random intercepts at the level of the study areas, and were corrected for potential selection bias by applying the probability of participation in present analyses as weights derived from a logistic regression with predictors from the baseline study in 1991.

Table 3. Interaction between *MTNR1B* variants (single/score) and night-time road traffic noise (10 dB difference) on change in HbA1c, in non-movers.

MTNR1B Variant	Risk/Other Allele	Risk Allele Frequency	All N = 2142	No Diabetes N = 1960	Diabetes N = 179	Diabetes on Medication N = 117
			β Interaction Term (95% CI)	β Interaction Term (95% CI)	β Interaction Term (95% CI)	β Interaction Term (95% CI)
rs1387153 [a]	T/C	0.30	0.04 (−0.02, 0.10)	0.01 (−0.01, 0.02)	0.66 (0.17, 1.15) *	0.92 (0.45, 1.39) †
rs1387153 [b]	T/C	0.30	0.04 (−0.02, 0.10)	0.01 (−0.01, 0.02)	0.67 (0.17, 1.18) *	0.92 (0.44, 1.40) †
rs10830962 [a]	G/C	0.42	0.03 (−0.02, 0.09)	−0.003 (−0.02, 0.01)	0.56 (0.18, 0.95) *	0.83 (0.26, 1.39) *
rs10830962 [b]	G/C	0.42	0.03 (−0.02, 0.09)	−0.003 (−0.02, 0.01)	0.57 (0.18, 0.96) *	0.83 (0.21, 1.46) *
rs4753426 [a]	C/T	0.49	0.01 (−0.05, 0.07)	−0.01 (−0.04, 0.02)	0.30 (−0.19, 0.80)	0.42 (−0.01, 0.85) ‡
rs4753426 [b]	C/T	0.49	0.01 (−0.05, 0.07)	−0.01 (−0.04, 0.02)	0.31 (−0.18, 0.80)	0.42 (0.01, 0.83) *
rs8192552 [a]	G/A	0.93	−0.001 (−0.06, 0.05)	−0.02 (−0.05, 0.02)	0.31 (−0.52, 1.15)	0.28 (−0.41, 0.98)
rs8192552 [b]	G/A	0.93	−0.001 (−0.05, 0.05)	−0.02 (−0.05, 0.02)	0.31 (−0.52, 1.14)	0.28 (−0.45, 1.01)
rs10830963 [a]	G/C	0.29	0.04 (−0.03, 0.10)	−0.004 (−0.02, 0.02)	0.80 (0.16, 1.45) *	1.21 (0.62, 1.78) †
rs10830963 [b]	G/C	0.29	0.04 (−0.03, 0.10)	−0.004 (−0.02, 0.02)	0.80 (0.14, 1.47) *	1.21 (0.59, 1.83) †
rs3781638 [a]	T/G	0.55	0.01 (−0.06, 0.07)	−0.002 (−0.04, 0.03)	0.18 (−0.17, 0.54)	0.32 (−0.08, 0.71) ‡
rs3781638 [b]	T/G	0.55	0.01 (−0.05, 0.07)	−0.002 (−0.04, 0.03)	0.18 (−0.17, 0.54)	0.31 (−0.06, 0.69) ‡
MTNR1B score [a]			0.01 (−0.01, 0.02)	−0.001 (−0.01, 0.004)	0.15 (0.01, 0.30) *	0.23 (0.10, 0.36) *
MTNR1B score [b]			0.01 (−0.01, 0.02)	−0.001 (−0.01, 0.004)	0.15 (0.01, 0.30) *	0.23 (0.10, 0.37) *
MTNR1B score >6 vs. ≤6 § [a]			0.07 (−0.01, 0.15) ‡	0.01 (−0.01, 0.03)	0.87 (0.38, 1.36) †	1.25 (0.70, 1.80) †
MTNR1B score >6 vs. ≤6 § [b]			0.07 (−0.01, 0.15) ‡	0.01 (−0.02, 0.04)	0.87 (0.38, 1.37) †	1.25 (0.68, 1.83) †
§ + Short Form-36 mental health ¦ [a]			0.07 (−0.01, 0.15) ‡	0.02 (−0.02, 0.06)	0.82 (0.36, 1.28) †	1.11 (0.76, 1.46) †
§ + Short Form-36 mental health ¦ [b]			0.07 (−0.01, 0.15) ‡	0.02 (−0.02, 0.06)	0.81 (0.36, 1.27) †	1.12 (0.74, 1.49) †

[a] Adjusted model without BMI; [b] Adjusted model with BMI; § *MTNR1B* score > 6 vs. ≤6 model; * $p < 0.05$; † $p < 0.001$; ‡ $p < 0.1$. ¦ N (all) = 1865; N (no diabetes) = 1711; N (diabetes) = 152; N (diabetes on medication) = 99.

Although we observed statistically significant associations between RTN and mean ΔHbA1c among non-diabetic participants with high MGRS, the difference in associations between non-diabetic participants with high and low MGRS was non-significant ($P_{\text{interaction}}$ = 0.39) (Figure 2).

All results are presented as increase or decrease in mean change in HbA1c per 10 dB of night-time road traffic noise. *MTNR1B* genetic risk score represents the sum of risk alleles across six included *MTNR1B* variants. All models were adjusted for age, sex, education, neighborhood socio-economic index, smoking status, passive smoking, alcohol consumption, body mass index, green space within a 2 km residential buffer, residential levels of nitrogen dioxide, night-time railway, aircraft noise and their truncation indicators. All models include random intercepts at the level of the study areas, and were corrected for potential selection bias by applying the probability of participation in present analyses as weights derived from a logistic regression with predictors from the baseline study in 1991.

The interactions between RTN and MGRS did not significantly differ by sex in both diabetic and non-diabetic participants ($P_{\text{interaction}} \geq 0.30$). Similar to the main models, interaction models were also very stable to adjustments for BMI (Table 3). Even though the direction of interaction terms in models including all participants was generally similar to those in non-movers, the magnitude of interactions was smaller, and the interaction terms in those models were not statistically significant (Table S3).

3.5. Sensitivity Analyses

Our results were robust to sensitivity analyses, including complete case analyses without correction for potential selection bias, as well as RTN–MGRS interaction analyses excluding asthmatic participants Sensitivity analyses limited to participants whose bedrooms were oriented towards the street also showed very similar results. Interactions with self-reported sleep insufficiency were consistent in diabetic participants, reaching significance among those on medication. Unlike with MGRS, we observed significant interaction with self-reported sleep sufficiency in non-diabetic individuals (Table 4).

MTNR1B score represents the sum of risk alleles across six included variants. All estimates represent an increase (+) or decrease (−) in mean change in HbA1c per 10 dB of night-time road traffic noise. Adjusted models included age, sex, education, neighborhood socio-economic index, smoking status, passive smoking, alcohol consumption, body mass index, green space within a 2 km residential buffer, residential levels of nitrogen dioxide, night-time railway, aircraft noise, and their truncation indicators. All models include random intercepts at the level of the study areas, and were adjusted for potential selection bias by applying the probability of participation in present analyses as weights derived from a logistic regression with predictors from the baseline study in 1991. NA: not applicable.

Table 4. Sensitivity analyses on the association between night-time road traffic noise and change in HbA1c, in non-movers.

Model	Categories	All N = 2142	No Diabetes N = 1960	Diabetes N = 179	Diabetes on Medication N = 117
		β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)
Main model corrected for potential selection bias	NA	0.02 (0.003, 0.04) *	0.01 (0, 0.03) *	0.03 (−0.14, 0.19)	0.10 (−0.09, 0.29)
Complete case analyses: Adjusted model	NA	0.02 (−0.002, 0.05) †	0.02 (0, 0.04) *	0.03 (−0.21, 0.28)	0.05 (−0.28, 0.37)
Complete case analyses: Stratification by *MTNR1B* score	≤6	−0.01 (−0.04, 0.03)	0.01 (−0.01, 0.04)	−0.34 (−0.64, −0.04) *	−0.40 (−0.76, −0.05) *
	>6	0.06 (0.02, 0.09) *	0.03 (0.001, 0.05) *	0.49 (0.16, 0.82) *	0.81 (0.37, 1.26) ‡
	P-value of interaction	*<0.01*	*0.48*	*<0.001*	*<0.001*
Main model excluding participants whose bedrooms did not face the street: Stratification by *MTNR1B* score	≤6	0.01 (−0.03, 0.04)	0.02 (−0.01, 0.05)	−0.19 (−0.39, −0.003) *	−0.02 (−0.38, 0.34)
	>6	0.07 (0.01, 0.13) *	0.03 (−0.001, 0.06)	0.49 (0.08, 0.90) *	0.85 (0.23, 1.47) ‡
	P-value of interaction	*0.14*	*0.42*	*<0.001*	*<0.001*
Main model excluding asthma cases: Stratification by *MTNR1B* score	≤6	−0.01 (−0.03, 0.02)	0.004 (−0.02, 0.03)	−0.26 (−0.48, −0.03) *	−0.30 (−0.53, −0.07) *
	>6	0.06 (−0.003, 0.11) *	0.02 (−0.004, 0.05) *	0.50 (0.07, 0.92) *	0.98 (0.35, 1.61) *
	P-value of interaction	*0.07*	*0.39*	*0.001*	*<0.001*
Main model: Stratification by self-reported insufficient sleep	No	−0.002 (−0.03, 0.02)	−0.001 (−0.02, 0.02)	−0.08 (−0.28, 0.15)	−0.11 (−0.36, 0.14)
	Yes	0.07 (0.02, 0.14) *	0.05 (0.004, 0.09) *	0.28 (−0.13, 0.69)	0.67 (0.11, 1.23) *
	P-value of interaction	*0.092*	*0.059*	*0.2*	*0.029*

* $p < 0.05$; † $p < 0.1$; ‡ $p < 0.001$.

4. Discussion

We found positive associations between exposure to RTN and eight-year change in HbA1c in non-movers, which were significantly stronger among diabetic individuals at genetic risk of circadian rhythm disturbances. We also found positive associations with aircraft noise that were again stronger in diabetics. Railway noise was not associated with change in HbA1c, supporting findings from previous studies where road traffic and aircraft noise, but not railway noise [12,13] was associated with diabetes risk.

Melatonin is involved in the regulation of human circadian rhythms through its role in thermoregulation and sleep induction [17]. The secretion of melatonin occurs during the biological night, in a fasting state, when insulin secretion is low [52]. Therefore, the melatonin pathway may play a role in noise susceptibility if noise exposure causes early awakenings and potential early meals, which could stimulate insulin secretion during high melatonin levels, leading to impaired glucose tolerance [22,53]. Diabetics with high MGRS had higher mean BMI (31 kg/m^2) compared to those with low MGRS (29 kg/m^2). Noise could also delay sleep onset [54], possibly leading to later chronotype which was associated with metabolic disturbances and diabetes [23,55]. We found supportive evidence for this hypothesis only in diabetic participants. One explanation for this may be that individuals with diabetes usually have comorbidities and worse homeostatic mechanisms, and are more prone to environmental stressors [56,57]. Exposure to high traffic volume was shown to have a stronger impact among diabetic individuals on insulin, which could imply more complicated diabetes, expressing higher inflammatory profiles compared to those on oral hypoglycemic agents [58]. Participants using diabetic medication in this study were more likely to be overweight, to have cardiovascular diseases and higher C-reactive protein levels compared to the other participants. Diabetic individuals with a high genetic risk for melatonin profile dysregulation may be particularly sensitive to poor glucose homeostasis. Diabetic individuals have also been shown to have sleep–wake cycle irregularities compared to non-diabetics [59,60]. Our finding of a significant negative effect of noise on glycemia among diabetics with low MGRS is surprising. Although this could be a chance finding, diabetics with low MGRS had lower average night-time noise exposure (46 dB), and were less noise sensitive (45%) compared to those with high MGRS (48 dB noise level and 51% noise sensitivity). Diabetic participants with low MGRS potentially have better melatonin profiles, and could be at lower risk for noise disturbances. The complexity of the melatonin system in influencing several physiological processes [24] calls for more research to better understand this finding.

Our observation of interaction of RTN with self-reported sleep insufficiency, but not with genetic risk for melatonin profile dysregulation on ΔHbA_{1c} in non-diabetic individuals, suggests that the melatonin pathway may not be relevant in this group. Interestingly, noise-induced sleep disruption was reported to impair glucose homeostasis through non-melatonin pathways, including the activation of sympathetic nervous system and release of stress-related hormones [61]. As GWAS continues to identify more circadian-related variants, future studies should consider variants covering the entire circadian pathways, and also incorporate objective sleep and stress-related parameters to improve understanding of the cardiometabolic effects of noise.

Since night-time noise levels were correlated to Lden, we may not exclude the contributions of day-time noise exposure (potentially via the stress pathway) to our observations. Noise, through stress/anxiety, may reduce adherence to medication in diabetic individuals, worsening their glucose control [62,63]. Exploratory analyses among the respondents to the SF-36 mental health survey [32] showed a reduction in the magnitude of the observed interactions (in diabetic participants) following adjustment for their mental health scores (Table 3).

The strengths of this study derive from its novelty in applying gene–environment interactions to better understand the impact of noise on glucose control in a longitudinal design, and the availability of detailed phenotypic and genotypic information in the SAPALDIA study. Information on medication use allowed the exploration of different diabetes phenotypes. We could test the hypothesis covering a potential pathway of glycemic effects of noise exposure, and used validated models with high spatial resolutions to assign individual estimates of noise and air pollution. The availability of information on change of residence allowed focusing analyses on non-movers, allowing the use of baseline noise levels as long-term exposures towards our health outcome.

Although our study is limited by sample size which calls for cautionary interpretation, we controlled for potential selection bias and made salient findings which could be generalized to all non-movers in the study period. Potentially, post-transcriptional/translational modifications may have affected melatonin profile, hence the absence of a risk variant may not imply normal melatonin profiles, and vice versa. We also lack information on melatonin drug use by the participants. However, the lack of correlation (R = 0.01) between self-reported sleep insufficiency and MGRS in our study corroborates the findings of previous studies where the lead *MTNR1B* variant was not associated with sleep duration [22,64], validating our application of MGRS in this study. We did not have adequate nutrient intake information e.g., antioxidant or fiber intake. Adjustment for nutrients may be considered an over-adjustment, if part of the noise–HbA1c association is mediated by impact of noise on food preference. Although our noise estimates were from validated models, some degree of misclassification will have occurred due to errors in input data. We did not have information on participants' work shifts in our noise exposure models. However, the resulting bias is more likely non-differential, and to bias effect estimates towards null. We did not consider window opening habits which may be related to exposure level, and could have also biased our estimates towards null. Since we found significant associations in non-movers, we cannot generalize our findings to all SAPALDIA participants. Noise exposure metrics were significantly different between movers and non-movers, with a tendency for movers to move to areas with lower RTN (Table S6). This tendency per se, might have led to an over-estimation of the noise effects among movers if there had been no random misclassification. However, the total bias is the sum of any bias associated with differential misclassification, and the attenuation bias associated with random misclassification. Although we cannot prove it, we are quite convinced that the attenuation bias had a stronger impact on our results than any potential bias due to differential misclassification.

5. Conclusions

Our findings raise the hypothesis that genetic risk for melatonin profile dysregulation, in combination with long-term road traffic noise, better captured in non-movers, may increase the risk for poor glucose homeostasis, particularly in diabetic patients. While our findings need replication and confirmation in other independent cohorts, experimental interventional studies should test the hypothesis to determine if counseling for noise control should be added to diabetes care.

Supplementary Materials: The following are available online at www.mdpi.com/1660-4601/14/12/1492/s1, Table S1: Characteristics of common *MTNR1B* variants included in the study; Table S2: Characteristics of study participants included and excluded in the study; Table S3: Interaction between *MTNR1B* variants (single/score) and night-time road traffic noise (10 dB difference) on change in HbA1c, in all participants; Table S4: Association between *MTNR1B* variants and measures of HbA1c (%); Table S5: Interaction between *MTNR1B* genetic risk score and 10 dB difference in night-time road, railway and aircraft noise, on change in HbA1c; Table S6: Characteristics of included participants stratified by moving status; Figure S1: Selection of participants for the present study; Figure S2: Linkage disequilibrium matrix for seven common *MTNR1B* variants; Figure S3: Distribution of night-time road traffic, aircraft and railway noise in the included participants.

Acknowledgments: This work was supported by the Swiss National Science Foundation, SNF-SAPALDIA (grants numbers 33CS30-148470/1, 33CSCO-134276/1, 33CSCO-108796, 324730_135673, 3247BO-104283, 3247BO-104288, 3247BO-104284, 3247-065896, 3100-059302, 3200-052720, 3200-042532, 4026-028099, PMPDP3_129021/1, PMPDP3_141671/1); SNF-SiRENE (grant number CRSII3_147635) and the Swiss Federal Office for the Environment. SAPALDIA is also supported by the Federal Office of Public Health, the Federal Office of Roads and Transport, the canton's government of Aargau, Basel-Stadt, Basel-Land, Geneva, Luzern, Ticino, Valais, and Zürich, the Swiss Lung League, the canton's Lung League of Basel Stadt/Basel Landschaft, Geneva, Ticino, Valais, Graubünden and Zurich, Stiftung ehemals Bündner Heilstätten, SUVA, Freiwillige Akademische Gesellschaft, UBS Wealth Foundation, Talecris Biotherapeutics GmbH, Abbott Diagnostics, European Commission 018996 (GABRIEL), Wellcome Trust WT 084703MA.

Author Contributions: N.P.-H and I.C.E. conceived and designed the experiments; M.I., A.K., D.V., R.P., A.V., M.B., J.-M.W., C.C., M.R. and N.P.-H. performed the measurements; I.C.E., M.I., M.F., E.S., A.K., D.V., H.H., R.P., A.V., C.S., M.B., J.-M.W., C.C., M.R. and N.P.-H. analyzed the data; I.C.E., M.I., M.F., E.S., A.K., D.V., H.H., F.R., L.T., R.P., A.V., C.S., M.B., J.-M.W., C.C., M.R. and N.P.-H. wrote the paper.

Conflicts of Interest: The authors declare no competing interests. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Derr, R.; Garrett, E.; Stacy, G.A.; Saudek, C.D. Is hba(1c) affected by glycemic instability? *Diabetes Care* **2003**, *26*, 2728–2733. [CrossRef] [PubMed]
2. Moss, S.E.; Klein, R.; Klein, B.E.; Meuer, S.M. The association of glycemia and cause-specific mortality in a diabetic population. *Arch. Intern. Med.* **1994**, *154*, 2473–2479. [CrossRef] [PubMed]
3. Florkowski, C. Hba1c as a diagnostic test for diabetes mellitus—Reviewing the evidence. *Clin. Biochem. Rev.* **2013**, *34*, 75–83. [PubMed]
4. Lipska, K.J.; Yao, X.; Herrin, J.; McCoy, R.G.; Ross, J.S.; Steinman, M.A.; Inzucchi, S.E.; Gill, T.M.; Krumholz, H.M.; Shah, N.D. Trends in drug utilization, glycemic control, and rates of severe hypoglycemia, 2006–2013. *Diabetes Care* **2016**, *40*, 468–475. [CrossRef] [PubMed]
5. Munzel, T.; Sorensen, M.; Gori, T.; Schmidt, F.P.; Rao, X.; Brook, F.R.; Chen, L.C.; Brook, R.D.; Rajagopalan, S. Environmental stressors and cardio-metabolic disease: Part II-mechanistic insights. *Eur. Heart J.* **2016**, *38*, 557–564. [CrossRef] [PubMed]
6. Snell-Bergeon, J.K.; Wadwa, R.P. Hypoglycemia, diabetes, and cardiovascular disease. *Diabetes Technol. Ther.* **2012**, *14*. [CrossRef] [PubMed]
7. Davidson, J.A.; Parkin, C.G. Is hyperglycemia a causal factor in cardiovascular disease? Does proving this relationship really matter? Yes. *Diabetes Care* **2009**, *32*, S331–S333. [CrossRef] [PubMed]
8. Vitelli, L.L.; Shahar, E.; Heiss, G.; McGovern, P.G.; Brancati, F.L.; Eckfeldt, J.H.; Folsom, A.R. Glycosylated hemoglobin level and carotid intimal-medial thickening in nondiabetic individuals. The atherosclerosis risk in communities study. *Diabetes Care* **1997**, *20*, 1454–1458. [CrossRef] [PubMed]
9. Hanninen, O.; Knol, A.B.; Jantunen, M.; Lim, T.A.; Conrad, A.; Rappolder, M.; Carrer, P.; Fanetti, A.C.; Kim, R.; Buekers, J.; et al. Environmental burden of disease in europe: Assessing nine risk factors in six countries. *Environ. Health Perspect.* **2014**, *122*, 439–446. [CrossRef] [PubMed]
10. Dzhambov, A.M. Long-term noise exposure and the risk for type 2 diabetes: A meta-analysis. *Noise Health* **2015**, *17*, 23–33. [CrossRef] [PubMed]
11. Sorensen, M.; Andersen, Z.J.; Nordsborg, R.B.; Becker, T.; Tjonneland, A.; Overvad, K.; Raaschou-Nielsen, O. Long-term exposure to road traffic noise and incident diabetes: A cohort study. *Environ. Health Perspect.* **2013**, *121*, 217–222. [PubMed]
12. Eze, I.C.; Foraster, M.; Schaffner, E.; Vienneau, D.; Heritier, H.; Rudzik, F.; Thiesse, L.; Pieren, R.; Imboden, M.; von Eckardstein, A.; et al. Long-term exposure to transportation noise and air pollution in relation to incident diabetes in the sapaldia study. *Int. J. Epidemiol.* **2017**, *46*, 1115–1125. [CrossRef] [PubMed]
13. Roswall, N.; Raaschou-Nielsen, O.; Jensen, S.S.; Tjonneland, A.; Sorensen, M. Long-term exposure to residential railway and road traffic noise and risk for diabetes in a danish cohort. *Environ. Res.* **2017**, *160*, 292–297. [CrossRef] [PubMed]
14. Clark, C.; Sbihi, H.; Tamburic, L.; Brauer, M.; Frank, L.D.; Davies, H.W. Association of long-term exposure to transportation noise and traffic-related air pollution with the incidence of diabetes: A prospective cohort study. *Environ. Health Perspect.* **2017**, *125*. [CrossRef] [PubMed]
15. Barcelo, M.A.; Varga, D.; Tobias, A.; Diaz, J.; Linares, C.; Saez, M. Long term effects of traffic noise on mortality in the city of Barcelona, 2004–2007. *Environ. Res.* **2016**, *147*, 193–206. [CrossRef] [PubMed]
16. Cui, B.; Gai, Z.; She, X.; Wang, R.; Xi, Z. Effects of chronic noise on glucose metabolism and gut microbiota-host inflammatory homeostasis in rats. *Sci. Rep.* **2016**, *6*, 36693. [CrossRef] [PubMed]
17. Cajochen, C.; Krauchi, K.; Wirz-Justice, A. Role of melatonin in the regulation of human circadian rhythms and sleep. *J. Neuroendocrinol.* **2003**, *15*, 432–437. [CrossRef] [PubMed]
18. Boden, G.; Ruiz, J.; Urbain, J.L.; Chen, X. Evidence for a circadian rhythm of insulin secretion. *Am. J. Physiol.* **1996**, *271*, E246–E252. [PubMed]

19. Cajochen, C.; Krauchi, K.; Danilenko, K.V.; Wirz-Justice, A. Evening administration of melatonin and bright light: Interactions on the eeg during sleep and wakefulness. *J. Sleep Res.* **1998**, *7*, 145–157. [CrossRef] [PubMed]
20. Rubio-Sastre, P.; Scheer, F.A.; Gomez-Abellan, P.; Madrid, J.A.; Garaulet, M. Acute melatonin administration in humans impairs glucose tolerance in both the morning and evening. *Sleep* **2014**, *37*, 1715–1719. [CrossRef] [PubMed]
21. Mussig, K.; Staiger, H.; Machicao, F.; Haring, H.U.; Fritsche, A. Genetic variants in MTNR1b affecting insulin secretion. *Ann. Med.* **2010**, *42*, 387–393. [CrossRef] [PubMed]
22. Lane, J.M.; Chang, A.M.; Bjonnes, A.C.; Aeschbach, D.; Anderson, C.; Cade, B.E.; Cain, S.W.; Czeisler, C.A.; Gharib, S.A.; Gooley, J.J.; et al. Impact of common diabetes risk variant in mtnr1b on sleep, circadian, and melatonin physiology. *Diabetes* **2016**, *65*, 1741–1751. [CrossRef] [PubMed]
23. Anothaisintawee, T.; Lertrattananon, D.; Thamakaison, S.; Knutson, K.L.; Thakkinstian, A.; Reutrakul, S. Later chronotype is associated with higher hemoglobin a1c in prediabetes patients. *Chronobiol. Int.* **2017**, *34*, 393–402. [CrossRef] [PubMed]
24. Bonnefond, A.; Froguel, P. Disentangling the role of melatonin and its receptor mtnr1b in type 2 diabetes: Still a long way to go? *Curr. Diabetes Rep.* **2017**, *17*. [CrossRef] [PubMed]
25. Eze, I.C.; Imboden, M.; Kumar, A.; von Eckardstein, A.; Stolz, D.; Gerbase, M.W.; Kunzli, N.; Pons, M.; Kronenberg, F.; Schindler, C.; et al. Air pollution and diabetes association: Modification by type 2 diabetes genetic risk score. *Environ. Int.* **2016**, *94*, 263–271. [CrossRef] [PubMed]
26. Talmud, P.J.; Cooper, J.A.; Morris, R.W.; Dudbridge, F.; Shah, T.; Engmann, J.; Dale, C.; White, J.; McLachlan, S.; Zabaneh, D.; et al. Sixty-five common genetic variants and prediction of type 2 diabetes. *Diabetes* **2015**, *64*, 1830–1840. [CrossRef] [PubMed]
27. Martin, B.W.; Ackermann-Liebrich, U.; Leuenberger, P.; Kunzli, N.; Stutz, E.Z.; Keller, R.; Zellweger, J.P.; Wuthrich, B.; Monn, C.; Blaser, K.; et al. Sapaldia: Methods and participation in the cross-sectional part of the swiss study on air pollution and lung diseases in adults. *Soz. Praventivmed.* **1997**, *42*, 67–84. [CrossRef] [PubMed]
28. Weykamp, C.; Visser-Dekkers, W.; Kemna, E.; Siebelder, C. Effects of hemoglobin d and e on the measurement of hba1c with the modified menarini/arkray adams a1c ha-8180v analyser. *Clin. Chim. Acta* **2012**, *414*, 44–45. [CrossRef] [PubMed]
29. Sacks, D.B. Global harmonization of hemoglobin a1c. *Clin. Chem.* **2005**, *51*, 681–683. [CrossRef] [PubMed]
30. Karipidis, I.; Vienneau, D.; Habermacher, M.; Köpflii, M.; Brink, M.; Probst-Hensch, N.; Röösli, M.; Wunderli, J.-M. Reconstruction of historical noise exposure data for environmental epidemiology in switzerland within the sirene project. *Noise Map.* **2014**, *1*. [CrossRef]
31. Panczak, R.; Galobardes, B.; Voorpostel, M.; Spoerri, A.; Zwahlen, M.; Egger, M. A swiss neighbourhood index of socioeconomic position: Development and association with mortality. *J. Epidemiol. Community Health* **2012**, *66*, 1129–1136. [CrossRef] [PubMed]
32. Ware, J.E.; Kosinski, M. *Sf-36 Physical & Mental Health Summary Scales: A Manual for Users of Version 1*; Quality Metric: Montreal, QC, Canada, 2001.
33. Tetreault, L.F.; Perron, S.; Smargiassi, A. Cardiovascular health, traffic-related air pollution and noise: Are associations mutually confounded? A systematic review. *Int. J. Public Health* **2013**, *58*, 649–666. [CrossRef] [PubMed]
34. Liu, L.-J.S.; Tsai, M.-Y.; Keidel, D.; Gemperli, A.; Ineichen, A.; Hazenkamp-von Arx, M.; Bayer-Oglesby, L.; Rochat, T.; Künzli, N.; Ackermann-Liebrich, U. Long-term exposure models for traffic related NO_2 across geographically diverse areas over separate years. *Atmos. Environ.* **2012**, *46*, 460–471.
35. Moffatt, M.F.; Gut, I.G.; Demenais, F.; Strachan, D.P.; Bouzigon, E.; Heath, S.; von Mutius, E.; Farrall, M.; Lathrop, M.; Cookson, W.O. A large-scale, consortium-based genomewide association study of asthma. *N. Engl. J. Med.* **2010**, *363*, 1211–1221. [CrossRef] [PubMed]
36. Delaneau, O.; Zagury, J.F.; Marchini, J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat. Methods* **2013**, *10*, 5–6. [CrossRef] [PubMed]
37. Fuchsberger, C.; Abecasis, G.R.; Hinds, D.A. Minimac2: Faster genotype imputation. *Bioinformatics* **2015**, *31*, 782–784. [CrossRef] [PubMed]

38. Bouatia-Naji, N.; Bonnefond, A.; Cavalcanti-Proenca, C.; Sparso, T.; Holmkvist, J.; Marchand, M.; Delplanque, J.; Lobbens, S.; Rocheleau, G.; Durand, E.; et al. A variant near mtnr1b is associated with increased fasting plasma glucose levels and type 2 diabetes risk. *Nat. Genet.* **2009**, *41*, 89–94. [CrossRef] [PubMed]
39. Staiger, H.; Machicao, F.; Schafer, S.A.; Kirchhoff, K.; Kantartzis, K.; Guthoff, M.; Silbernagel, G.; Stefan, N.; Haring, H.U.; Fritsche, A. Polymorphisms within the novel type 2 diabetes risk locus mtnr1b determine beta-cell function. *PLoS ONE* **2008**, *3*, e3962. [CrossRef] [PubMed]
40. Andersson, E.A.; Holst, B.; Sparso, T.; Grarup, N.; Banasik, K.; Holmkvist, J.; Jorgensen, T.; Borch-Johnsen, K.; Egerod, K.L.; Lauritzen, T.; et al. Mtnr1b g24e variant associates with bmi and fasting plasma glucose in the general population in studies of 22,142 europeans. *Diabetes* **2010**, *59*, 1539–1548. [CrossRef] [PubMed]
41. Prokopenko, I.; Langenberg, C.; Florez, J.C.; Saxena, R.; Soranzo, N.; Thorleifsson, G.; Loos, R.J.; Manning, A.K.; Jackson, A.U.; Aulchenko, Y.; et al. Variants in mtnr1b influence fasting glucose levels. *Nat. Genet.* **2009**, *41*, 77–81. [CrossRef] [PubMed]
42. Lyssenko, V.; Nagorny, C.L.; Erdos, M.R.; Wierup, N.; Jonsson, A.; Spegel, P.; Bugliani, M.; Saxena, R.; Fex, M.; Pulizzi, N.; et al. Common variant in mtnr1b associated with increased risk of type 2 diabetes and impaired early insulin secretion. *Nat. Genet.* **2009**, *41*, 82–88. [CrossRef] [PubMed]
43. Ronn, T.; Wen, J.; Yang, Z.; Lu, B.; Du, Y.; Groop, L.; Hu, R.; Ling, C. A common variant in mtnr1b, encoding melatonin receptor 1b, is associated with type 2 diabetes and fasting plasma glucose in han chinese individuals. *Diabetologia* **2009**, *52*, 830–833. [CrossRef] [PubMed]
44. Sparso, T.; Bonnefond, A.; Andersson, E.; Bouatia-Naji, N.; Holmkvist, J.; Wegner, L.; Grarup, N.; Gjesing, A.P.; Banasik, K.; Cavalcanti-Proenca, C.; et al. G-allele of intronic rs10830963 in mtnr1b confers increased risk of impaired fasting glycemia and type 2 diabetes through an impaired glucose-stimulated insulin release: Studies involving 19,605 europeans. *Diabetes* **2009**, *58*, 1450–1456. [CrossRef] [PubMed]
45. Langenberg, C.; Pascoe, L.; Mari, A.; Tura, A.; Laakso, M.; Frayling, T.M.; Barroso, I.; Loos, R.J.; Wareham, N.J.; Walker, M. Common genetic variation in the melatonin receptor 1b gene (mtnr1b) is associated with decreased early-phase insulin response. *Diabetologia* **2009**, *52*, 1537–1542. [CrossRef] [PubMed]
46. Ingelsson, E.; Langenberg, C.; Hivert, M.F.; Prokopenko, I.; Lyssenko, V.; Dupuis, J.; Magi, R.; Sharp, S.; Jackson, A.U.; Assimes, T.L.; et al. Detailed physiologic characterization reveals diverse mechanisms for novel genetic loci regulating glucose and insulin metabolism in humans. *Diabetes* **2010**, *59*, 1266–1275. [CrossRef] [PubMed]
47. Reiling, E.; van't Riet, E.; Groenewoud, M.J.; Welschen, L.M.; van Hove, E.C.; Nijpels, G.; Maassen, J.A.; Dekker, J.M.; 't Hart, L.M. Combined effects of single-nucleotide polymorphisms in GCK, GCKR, G6PC2 and MTNR1B On fasting plasma glucose and type 2 diabetes risk. *Diabetologia* **2009**, *52*, 1866–1870. [CrossRef] [PubMed]
48. Dimas, A.S.; Lagou, V.; Barker, A.; Knowles, J.W.; Magi, R.; Hivert, M.F.; Benazzo, A.; Rybin, D.; Jackson, A.U.; Stringham, H.M.; et al. Impact of type 2 diabetes susceptibility variants on quantitative glycemic traits reveals mechanistic heterogeneity. *Diabetes* **2014**, *63*, 2158–2171. [CrossRef] [PubMed]
49. Scott, R.A.; Lagou, V.; Welch, R.P.; Wheeler, E.; Montasser, M.E.; Luan, J.; Magi, R.; Strawbridge, R.J.; Rehnberg, E.; Gustafsson, S.; et al. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat. Genet.* **2012**, *44*, 991–1005. [CrossRef] [PubMed]
50. Ward, L.D.; Kellis, M. Haploreg: A resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* **2012**, *40*. [CrossRef] [PubMed]
51. Walther, D.; Curjuric, I.; Dratva, J.; Schaffner, E.; Quinto, C.; Schmidt-Trucksass, A.; Eze, I.C.; Burdet, L.; Pons, M.; Gerbase, M.W.; et al. Hypertension, diabetes and lifestyle in the long-term—Results from a swiss population-based cohort. *Prev. Med.* **2017**, *97*, 56–61. [CrossRef] [PubMed]
52. Laposky, A.D.; Bass, J.; Kohsaka, A.; Turek, F.W. Sleep and circadian rhythms: Key components in the regulation of energy metabolism. *FEBS Lett.* **2008**, *582*, 142–151. [CrossRef] [PubMed]
53. Sartori, C.; Dessen, P.; Mathieu, C.; Monney, A.; Bloch, J.; Nicod, P.; Scherrer, U.; Duplain, H. Melatonin improves glucose homeostasis and endothelial vascular function in high-fat diet-fed insulin-resistant mice. *Endocrinology* **2009**, *150*, 5311–5317. [CrossRef] [PubMed]
54. Halperin, D. Environmental noise and sleep disturbances: A threat to health? *Sleep Sci.* **2014**, *7*, 209–212. [CrossRef] [PubMed]

55. Reutrakul, S.; Hood, M.M.; Crowley, S.J.; Morgan, M.K.; Teodori, M.; Knutson, K.L.; Van Cauter, E. Chronotype is independently associated with glycemic control in type 2 diabetes. *Diabetes Care* **2013**, *36*, 2523–2529. [CrossRef] [PubMed]
56. O'Neill, M.S.; Veves, A.; Zanobetti, A.; Sarnat, J.A.; Gold, D.R.; Economides, P.A.; Horton, E.S.; Schwartz, J. Diabetes enhances vulnerability to particulate air pollution-associated impairment in vascular reactivity and endothelial function. *Circulation* **2005**, *111*, 2913–2920. [CrossRef] [PubMed]
57. Raaschou-Nielsen, O.; Sorensen, M.; Ketzel, M.; Hertel, O.; Loft, S.; Tjonneland, A.; Overvad, K.; Andersen, Z.J. Long-term exposure to traffic-related air pollution and diabetes-associated mortality: A cohort study. *Diabetologia* **2013**, *56*, 36–46. [CrossRef] [PubMed]
58. Rioux, C.L.; Tucker, K.L.; Brugge, D.; Mwamburi, M. Medication type modifies inflammatory response to traffic exposure in a population with type 2 diabetes. *Environ. Pollut.* **2015**, *202*, 58–65. [CrossRef] [PubMed]
59. Nakanishi-Minami, T.; Kishida, K.; Funahashi, T.; Shimomura, I. Sleep-wake cycle irregularities in type 2 diabetics. *Diabetol. Metabol. Syndr.* **2012**, *4*, 18. [CrossRef] [PubMed]
60. Khandelwal, D.; Dutta, D.; Chittawar, S.; Kalra, S. Sleep disorders in type 2 diabetes. *Ind. J. Endocrinol. Metab.* **2017**, *21*, 758–761.
61. Babisch, W. Stress hormones in the research on cardiovascular effects of noise. *Noise Health* **2003**, *5*, 1–11. [PubMed]
62. Roohafza, H.; Kabir, A.; Sadeghi, M.; Shokouh, P.; Ahmadzad-Asl, M.; Khadem-Maboudi, A.A.; Sarrafzadegan, N. Stress as a risk factor for noncompliance with treatment regimens in patients with diabetes and hypertension. *ARYA Atheroscler.* **2016**, *12*, 166–171. [PubMed]
63. Gonzalez, J.S.; Kane, N.S.; Binko, D.H.; Shapira, A.; Hoogendoorn, C.J. Tangled up in blue: Unraveling the links between emotional distress and treatment adherence in type 2 diabetes. *Diabetes Care* **2016**, *39*, 2182–2189. [CrossRef] [PubMed]
64. Liu, C.; Wu, Y.; Li, H.; Qi, Q.; Langenberg, C.; Loos, R.J.; Lin, X. Mtnr1b rs10830963 is associated with fasting plasma glucose, hba1c and impaired beta-cell function in chinese hans from shanghai. *BMC Med. Genet.* **2010**, *11*, 59. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Review

Sepiolite-Based Adsorbents for the Removal of Potentially Toxic Elements from Water: A Strategic Review for the Case of Environmental Contamination in Hunan, China

Zhenghua Wang [1,2], Lina Liao [2], Andrew Hursthouse [1,3,*], Na Song [3] and Bozhi Ren [1,2]

1 Hunan Provincial Key Laboratory of Shale Gas Resource Exploitation, Hunan University of Science & Technology, Xiangtan 411201, China; wzh@hnust.edu.cn (Z.W.); xtrbz@sina.com (B.R.)
2 School of Civil Engineering, Hunan University of Science & Technology, Xiangtan 411201, China; liaolina123@foxmail.com
3 School of Computing, Engineering & Physical Sciences, University of the West of Scotland, Paisley PA1 2BE, UK; na.song@uws.ac.uk
* Correspondence: andrew.hursthouse@uws.ac.uk

Received: 4 July 2018; Accepted: 2 August 2018; Published: 3 August 2018

Abstract: The last few decades have seen rapid industrialization and urban development in many regions globally; with associated pollution by potentially toxic elements; which have become a threat to human health and the food chain. This is particularly prevalent in a number of regions in China that host multiple mineral resources and are important agricultural locations. Solutions to protect contamination of the food chain are more effective and sustainable if locally sourced materials are available; and in this context; we review the potential of local (sepiolite) mineral deposits to treat water contamination in the Hunan Municipality; central south China; widely recognized for significant environmental pollution issues (particularly by Hg; Cd; Pb; and Cr) and the high agricultural productivity of the region. Sepiolite is an abundant fibrous clay mineral with modest to good adsorption properties and extensive industrial process applications. It shows reasonable performance as an adsorbent for element removal. In addition; a number of surface modification strategies are available that improve this capability. We review these studies; focused on sorption reaction mechanisms and regeneration potential; with a view to present options for a localized and effective economic strategy for future application.

Keywords: sepiolite; adsorption; potentially toxic elements; modification; regeneration; Xiangjiang River

---

1. Introduction

With the international scale of rapid industrialization and urbanization, many developing and emerging economies have exploited local natural resources. These activities are energy-intensive, associated with significant interventions in the natural ecosystem including the water balance, which leads to emissions of pollutants to water, soil, and air. Metal pollution is one of the most serious and frequently encountered problems. For example, the River Ganga in India, which is considered sacred by Indian society, has serious pollution from Mn, Cr, Pb, Cd, and other potentially toxic elements (PTEs) [1]. Pollutants, such as Cd, Cu and Pb, in the Red River in Vietnam, are also much higher than the local discharge standards, and the maximum enrichment factor for Cd is 19.3 [2]. China's rapid industrialization has also led to a severe deterioration in water quality in the country's lakes and rivers. More than 80 percent of Chinese rivers and lakes, including seven key river systems, are contaminated with different types and to different degrees of PTEs with As, Cd, Hg, and Pb being the most frequently detected in these rivers [3–5]. Soils in China also suffer from high degrees of contamination. A report of the national survey of soil contamination of China, which was published in 2014, showed that the exceedance of environmental standards for Cd, Hg, As, Cu, Pb, Cr, Zn, and Ni in soil samples reached 7.0%, 1.6%, 2.7%, 2.1%, 1.5%, 1.1%, 0.9%, and 4.8%, respectively [6].

The persistence of PTE pollution has the potential to impact on both the human and wider ecological environment because of their long residence time and the potentially toxic impact through biological amplification [7,8]. PTE enrichment in the human body through the food chain eventually destroys the normal function of proteins and enzymes in the body, and high concentrations can form more toxic compounds, which do great harm to organisms [9,10]. Acute toxicity mainly affects the normal function of a particular organ and can damage or destroy the reproductive organs with intergenerational influence on child health [4,11].

Many techniques including adsorption, precipitation, biological treatment, ion exchange, and membrane separation are used to deal directly with PTE pollution [12–14]. Among them, adsorption, which refers to the adsorption of metals by means of intermolecular force or electron transfer and electron pair bonding [15], is widely used in pollution control because it is cheap and easy to apply and operate and systems are often reusable. Adsorption processes often refer generically to removal of target pollutants, which may also include other mechanisms such as ion exchange (the exchange of aqueous pollutant ions with available surface ions on the solid phase) and precipitation (where solution conditions exceed solubility conditions for specific species). Many kinds of adsorbent materials have been applied for the removal of soluble pollutant metals such as activated carbon, and modified complex materials such as metal ferrite doped carbon [16], and metal organic framework systems [17]. Capacities can be quite high, for example, 200–300 mg/g for Hg on modified carbon and similar range for, for example, U and Th on metal organic framework materials. However, applications are often limited due to its relatively high synthesis costs. Naturally derived clay minerals like kaolin, zeolite, sepiolite, bentonite, and perlite have also been utilized as alternative low-cost adsorbents for remediation of metal polluted environments. Whilst adsorption capacities for, for example, sepiolite may be an order of magnitude lower than synthetic systems [18] (there are many examples of studies of these materials for the treatment of metals by adsorption that show useful level of performance [19,20]. The focus for the future development of adsorption based systems should be on identifying adsorbent materials that are cheap and effective adsorbents in the context of the treatment scenarios (for example, at point source or to deal with diffuse pollution), to improve the effectiveness of these treatments and to ensure that no secondary pollution is produced [21–24]. Secondary tasks include the need for good solid/water separation and the regeneration of adsorbent. In this review, we focus on the potential of one mineral system in the context of a local demand for treatment, potentially supported by locally derived materials, which fits with the circular economy principles of resource use and efficiency. There are a number of serious regional pollution problems, within the Central Southern Chinese province of Hunan. It is rich in mineral resources, which have been extensively exploited over recent decades, compromising its significant contribution to food production from a strong agricultural sector. In addition to extensive base metal deposits, one of the regions' other significant mineral resources is sepiolite, a clay mineral with already widespread industrial and manufacturing applications, located in Xiangtan City, Hunan Province. We focus our review to consider work describing mechanisms for the removal of locally relevant metal pollutants, the modification and regeneration steps from the view of economic efficiency and resource sustainability. We identified a number of published studies of this topic, presented in Chinese literature and academic repositories. Our review, therefore, also provides wider international access to relevant data, which will be of benefit to similar locations worldwide.

2. Metal Pollution in Hunan, China

It is well known that there are abundant reserves of non-ferrous metals in Hunan Province, and most ores for mining, mineral processing, and smelting of non-ferrous and rare earth elements are located in the Xiangjiang Valley. The Xiangjiang River, which cuts across Hunan Province, is a main water resource for drinking water, process water, and the irrigation of crops. Because of prolonged mining and smelting activities for non-ferrous metals, wastewater has been discharged to the surrounding environment and the Xiangjiang Valley is the most infamous polluted area in central China [25]. Water, soils, and crops in Xiangjiang River basin are heavily contaminated by Cd, Hg, Pb, as well as As [26,27]. The "12th Five-Year" Plan for Comprehensive Prevention and Control of Heavy Metal Pollution indicates that the five major PTE pollutants in China are identified as Pb, Hg, Cd, Cr, and As [28]. In 2015, the state statistics for "the discharge of major pollutants in regional wastewater", Hg, Cd, and As in Hunan province account for 20.3%, 37.9%, and 32.6% of the total emissions, respectively [29]. The "11th Five-Year Plan" Xiangjiang River basin water pollution prevention and control plan of Hunan province reported that the pollution in Xiangjiang is predominantly caused by Hg, Cd, Pb, As, Zn, and others [30]. The monitoring data for the main pollution indicators in the Xiangjiang are shown in Table 1 with Figure 1 summarizing the geographical distribution of these sources/effects.

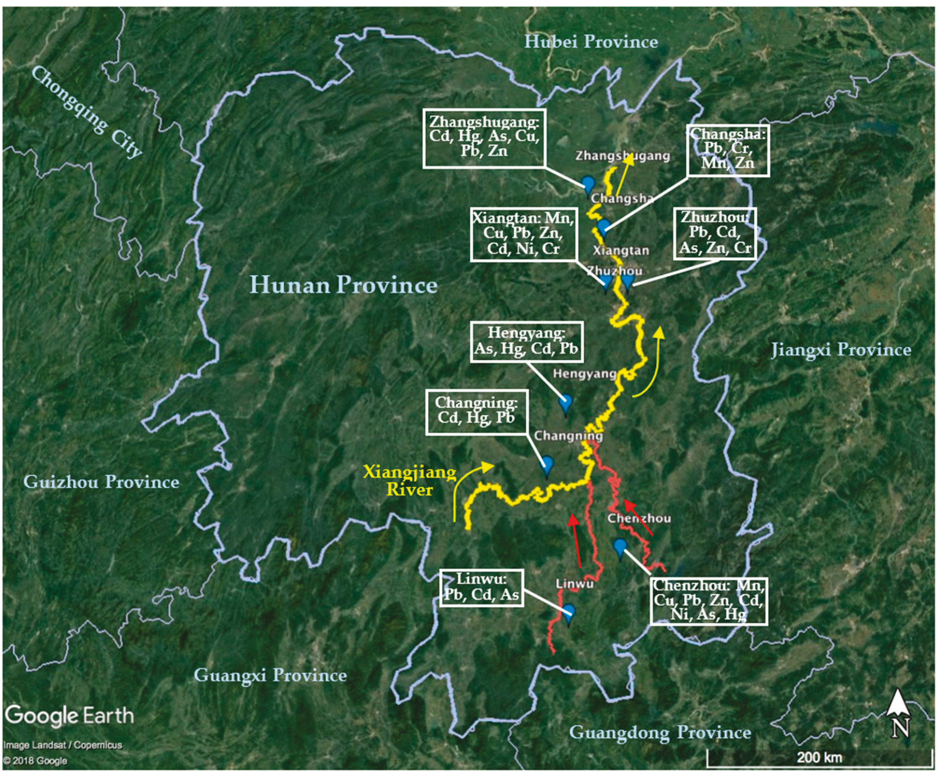


Figure 1. Geographical distribution of the main PTE pollution sources in Xiangjiang Valley. Data from references in Table 1.

Table 1. A summary of monitoring data for pollution indicators in Xiangjiang River.

Monitoring Locations in Xiangjiang River	PTE (s)	Pollution Condition	Reference
Upstream, Chenzhou Reach	Mn, Cu, Pb, Zn, Cd, Ni, As, Hg	Most of the elements exceed the standard e.g., Cd exceeds the discharge standard by about 1.2–9 times.	[31]
Gan river, Linwu (A tributary of Xiangjiang River)	Pb, Cd, As	Compared with the limit value of Environmental quality standards for surface water, Pb exceeds 109 times, Cd exceeds 242 times, As exceeds 75.8 times.	[32]
Changning Reach	Cd, Hg, Pb	Cd, Hg, and Pb were above the limit of the emission standard 265, 104.2, and 13.8 times, respectively.	[31]
Hengyang Reach	As, Hg, Pb, Cd	Compared with the limit value of Environmental quality standards for surface water, As, Hg, Pb, and Cd exceed the standard; 13.58%, 8.94%, 2.32%, and 27.16%, respectively.	[33]
Hengyang–Changsha Reach	As, Cd, Pb, Zn, Cu	Compared with the China Environmental Quality Standard for Soil Metals (GB15618-1995, Grade II), in the sample sites, the ratio of exceedance for As, Cd, and Pb was 13.2%, 68.5%, and 8.7% of soil samples, respectively. Cd, Pb, Zn, and Cu were much higher than their respective background values in the soil of Hunan Province, being 83.1–1178.7, 4.46–15.9, 2.88–16.1, and 3.35–6.22 times as high, respectively.	[34,35]
Changsha–Zhuzhou–Xiangtan Reach	Mn, Zn, Pb, Cu, Cr, Ni	Serious pollution of Pb and Zn, mild or moderate pollution of Mn and Cu in Xiawan, Zhuzhou Reach; moderate pollution of Pb and Zn, mild or moderate pollution of Cu, Mn, Ni, and Cr in Xiangtan and Changsha Reach.	[36]
Zhuzhou Reach	Cd, As, Pb, Zn, Cr	Cd in tobacco leaves is 6.98–37 mg/kg, Cd in cabbage is 15.4–18.3 mg/kg, Cd in rice is 1.03–1.78 mg/kg, Cd in amaranth is 6.03 mg/kg. As levels in vegetables are five times higher than normal.	[31,36]
Xiangtan Reach (manganese mine)	Mn, Cu, Pb, Zn, Cd, Ni	The average contents of Mn, Cu, Pb, Zn, Cd, and Ni are as follows: 7990.21 mg/kg, 66.38 mg/kg, 401.15 mg/kg, 640.32 mg/kg, 13.15 mg/kg, and 91.33 mg/kg. Their content is more than the national average worth several times or even dozens of times.	[37]
Entrance of Dongting Lake	Cd, Hg, As, Cu, Pb, Zn	Cd, Hg, As, Cu, Pb, and Zn in the sediments were 3.27, 0.190, 27.10, 39.8, 38.0, and 157.8 mg /kg, respectively.	[38]

3. Natural Sepiolite

The mineral sepiolite was discovered by German scholar Woemer in 1789, and the original name of sepiolite *Meerschaum* means "foam of the sea" in German. In 1847, it was officially renamed sepiolite. It is a clay mineral with light color and low density, and has the chemical formula $Mg_8(OH_2)_4[Si_6O_{15}]_2(OH)_4{\cdot}8H_2O$ [39]. In 1947, sepiolite was first discovered in Jiangxi, China by Chinese geologists Zhang Renjun. By genetic classification, we can divide the sepiolite into two types: sedimentary and hydrothermal sepiolite. The world has proven reserves of around 80 million tons, the main production of raw sepiolite is from deposits in Spain, followed by China, the United States, and Turkey [40]. China has about 30% of the world's sepiolite reserves. Among Chinese sepiolite reserves, 70% of sepiolite comes from Hunan Province. The city of Xiangtan, in Hunan province, hosts more than 20 million tonnes of sepiolite reserves.

Sepiolite is a clay mineral with a hydrous magnesium silicate, it is a member of the orthorhombic crystal system. It presents a structure of needle-like particles and has talc-like layers that consist of two layers of tetrahedral silica and a central octahedral magnesium layer [41]. As a result of its particular crystal structure, sepiolite has great sorptive, rheological, and catalytic properties, and it is also widely used in a variety of industrial and commercial applications.

The structure of sepiolite is a fibrous needle form, with a hollow channel in the direction of the fiber, which gives special rheological properties. The flow properties of sepiolite means that it is used in drilling muds as a thickener and suspension agent [42,43]. The acidity and alkalinity of sepiolite itself makes its catalytic activity more versatile and widespread. As a negative charge carrier, sepiolite can be utilized to remove pollutant cations [44]. It has a large specific surface area, which can reach 800–900 m^2/g, and with its porous properties, provides its good access to adsorption sites. This performance can play a part in applications for bleaching, cleaning agent, and other sorption functions. We focus here on those applications relevant to metal adsorption in the context of local environmental contamination.

4. Modification of Sepiolite

Despite the wide application of sepiolite in a variety of industrial processes when compared with other sorbent systems, it has relatively low surface acidity, narrow channels, low surface area, and poor thermal stability. This limits some applications of natural sepiolite [44]. The adsorption performance of the "modified" sepiolite can be much better than that of natural sepiolite, and studies have shown that the specific surface area can be increased significantly from 29–87 m^2/g [45]. The adsorption and removal capability of magnetic modified sepiolite for the heavy metal Cr (VI) is 10 times that of natural sepiolite [46] and of similar magnitude in the case of Hg^{2+} for surfactant modified sepiolite [18], but is still 0.5 to 0.3 of modified carbon and synthetic metal organic systems highlighted in the introduction above. When comparing natural sepiolite with a number of modified sepiolite systems to remove Pb^{2+}, it was found that the order of adsorption capacity is as follows: H_2O_2 modified sepiolite > KNO_3 modified sepiolite > natural sepiolite. When the initial Pb^{2+} concentration is 2.5 mg/L, the adsorption capacity of H_2O_2 modified sepiolite is twice as much as natural sepiolite [47]. According to other studies, the adsorption of Cr (VI) by activated sepiolite follows the following order: acid activated-mercapto silane organic modified sepiolite > sulphur silane modified sepiolite > acid activated modified sepiolite > natural sepiolite [48]. For the adsorption of Pb^{2+} thermally modified sepiolite > natural sepiolite > and for the adsorption of Cd^{2+} thermally modified sepiolite > natural sepiolite [49]. This provides a range of strategies for modification and sources of reactants to enable optimization of technological approach.

4.1. Acid Treatment

In acid treatment the reaction with carbonate in the sepiolite dissolves these impurities and clears any surface obstruction. The H^+ from the acid will replace the Ca^{2+}, Mg^{2+}, Na^+, and K^+ in the sepiolite interlayer, and improves the access to the surface and cavities in the sepiolite and increases the surface area and microporosity to provide improved adsorption performance [48,50,51].

4.2. Thermal Treatment

Thermal modification of sepiolite is the process of calcining natural sepiolite at different temperatures. At the different temperatures, the associated water in hygroscopic, zeolitic, and coordinated and structural octahedral hydroxyls groups will be removed to reduce the water film resistance, increasing the porosity and in doing so, improving the adsorption performance of sepiolite [52–54].

In calcination temperatures around 500 °C, the adsorbed water disappears from the sepiolite structure and apertures expand, with magnesium loss from the mineral structure. This increases the metal ion space, which is beneficial for the removal of Cr^{3+}, at the same time. The removal ability of heavy metals from wastewater was improved because of the change of internal cavity structure of the thermally treated sepiolite [55].

4.3. Magnetic Modification

After treatment of heavy metal contaminated wastewater by sepiolite, it is difficult to achieve good separation of mineral/water mixtures, which leads to the difficulty of secondary reuse of sepiolite. The magnetization of sepiolite provides an effective way to facilitate separation and allows further treatment and/or reuse. In addition to using iron-based systems, the Fe^{3+} present in the modified sepiolite has oxidative properties, and the dissolution of Fe^{3+} adds acidity, which is beneficial to the removal of heavy metal ions [46,56–58]. In the process of adsorption, the phase structure of magnetic sepiolite did not show any obvious change. The adsorption mechanism is an ion exchange process between heavy metal ions and zeolite in the magnetised sepiolite crystal and coupled adsorption occurs between the heavy metal ions and the hydroxyl groups (Fe–OH and Si–OH) on the surface of magnetised sepiolite [59].

4.4. Organic Modification

Organic modification uses a range of molecules such as surfactants, polymerised organic matter, or microorganisms to load or graft copolymerization to the surface or in a cavity of the sepiolite and modify its structure and the properties of the material surface. This is a relatively new area of research for metal removal from wastewater [48,60–63].

4.5. Acid Thermal Treatment

Acid thermal modification combines acid treatment with thermal modification of sepiolite. This approach to removing metal ions from water is widely used, and the adsorption using this acid thermal modified sepiolite is much better than for natural sepiolite, or separate acid modified and thermal treatment sepiolite. The sepiolite was treated with hydrochloric acid solution, and then calcinated at 450 °C, and subsequently used to treat Zn^{2+}, Pb^{2+}, Cu^{2+}, and Cd^{2+} in solution through an ion exchange and surface complex adsorption process [48].

5. Examples of the Application of Sepiolite to Potentially Toxic Element Removals from Aqueous Environmental Systems

It has been reported that sepiolite-based materials can be used to remove a wide range of pollutant elements from water and soils. We focus on reports relating to the key pollution metal elements in Hunan province (Hg, Cd, Pb, and Cr), where most of the research is still at the bench-scale phase.

5.1. The Removal Hg^{2+} from Wastewater

There are few studies on the removal of mercury ions using sepiolite and modified sepiolite. Those that are available show that modified sepiolite has a significantly improved effect on Hg^{2+} removal. As shown in Figure 2, the removal of Hg^{2+} was more than 90% after acid modification, acid thermal modification, and organic modification, compared with the natural sepiolite (about 50%).

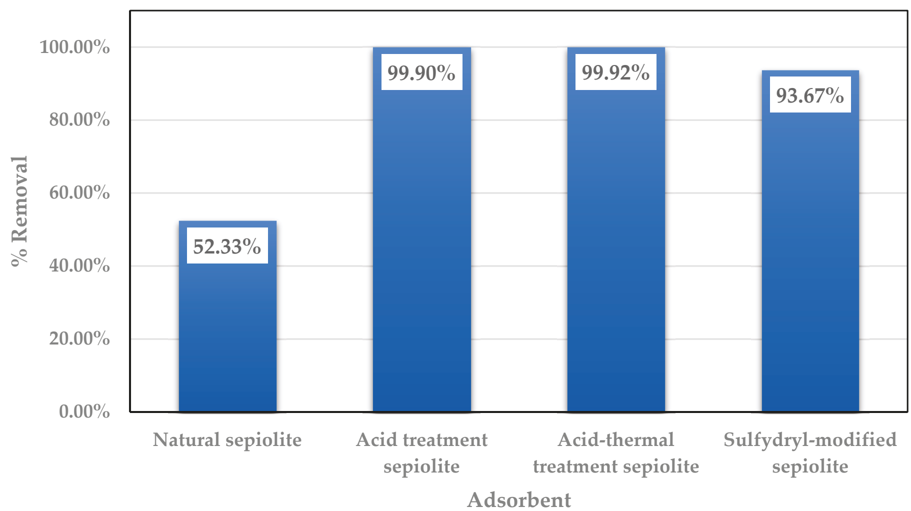


Figure 2. The removal of Hg by sepiolite and various modified sepiolite products [64–66].

The mechanism of removing Hg^{2+} using sepiolite and different modified conditions is also different. The reaction mechanism for acid modified sepiolite is the dissolution of the impurities in the sepiolite by the acid, at the same time, the pore area is enlarged, and the acidic hydroxyl group is exposed to contact with the Hg ions [65]. In the case of acid thermal treatment, the mechanism of the reaction with adsorbing ions is mainly due to the combination the approach for both acid and thermal activation. The process of dissolving the impurities in the sepiolite results in improved thermal removal of the structural water in the sepiolite and contact resistance is reduced. The combined effect is better than that of pure acid modification or thermal modification [66]. The reaction of sulfhydryl modified sepiolite for Hg^{2+} conforms to the pseudo-second-order kinetics equation, and adsorption of Hg^{2+} onto sepiolite fits Langmuir and Freundlich isotherm models. The regression coefficient (R^2 = 0.994) suggests that Hg^{2+} adsorption on sepiolite more closely followed the Langmuir model. The sulfhydryl modified sepiolite has a smoother surface, and its internal pores are enlarged, and the increased negative charge is conducive for reaction with metal ions, so as to more effectively remove the Hg^{2+} [64].

5.2. The Removal of Cd from Wastewater

Studies found that the adsorption reaction of sepiolite with Cd^{2+} conforms to a pseudo-second-order kinetics model, and its R^2 is 0.999. It also satisfies the isothermal adsorption model of both Langmuir and Freundlich, and the degree of fit to the Langmuir isothermal model is high (R^2 = 0.999). The saturated adsorption capacity of sepiolite for Cd^{2+} was 11.48 mg/g, and the saturated adsorption capacity of the acid modified sepiolite for Cd^{2+} was 13.62 mg/g [67,68]. The study on the adsorption of heavy metal Cd by sepiolite on acid and thermal treatment found that the adsorption of Cd^{2+} was increased by calcining the sepiolite, the main reason is that the CaO produced by the high temperature roasting makes the liquid alkaline, so that the Cd^{2+} was removed by precipitation reactions, in the process of acid treatment, the treatment effect of sulfuric acid is better than that of nitric acid and hydrochloric acid, which is due to the precipitation reaction between H_2SO_4 and Cd^{2+} [69]. The reaction mechanism of magnetic modified sepiolite to treat Cd^{2+} showed that the degree of fit to the Langmuir isothermal model was higher than Freundlich model, which indicates that the adsorption reaction of magnetic modified sepiolite and Cd^{2+} was based on single ion layer surface coverage [70]. Figure 3 shows that the adsorption capacity of Cd^{2+} was increased 3–17 times after modification, especially for the combined acid thermal modification.

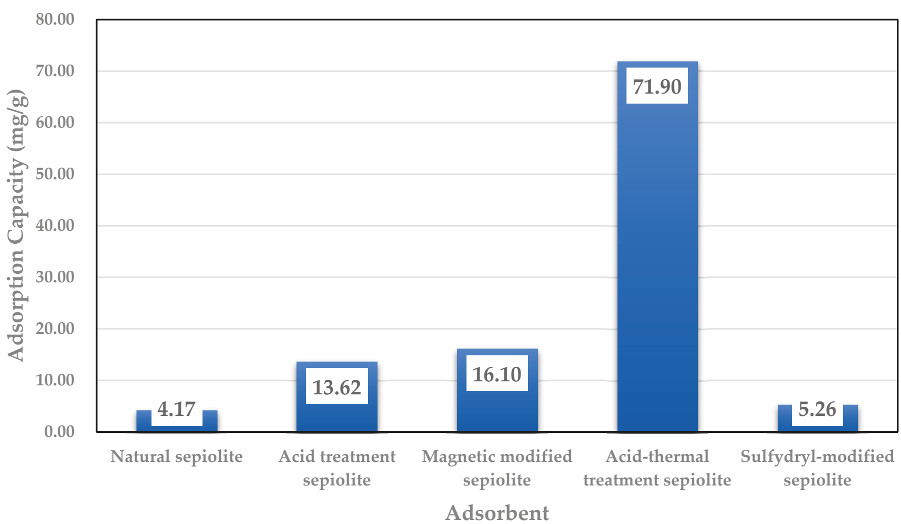


Figure 3. Adsorption capacity of Cd by sepiolite from solutions [66,68,70,71].

5.3. The Removal of Pb from Wastewater

The adsorption of Pb^{2+} with sepiolite and modified sepiolite departs from the Langmuir isothermal model, as a result of the precipitation caused by the reaction. The reaction mechanism is not only due in part to the complexation at the ion exchange surface, but also that the Pb^{2+} will precipitate during the reaction process [68]. The reaction mechanism of sulfhydryl modified sepiolite and Pb^{2+} fits well with the Freundlich isothermal model, and it also conforms to the pseudo-second-order kinetics, for which the R^2 = 0.9976, and its maximum adsorption capacity for Pb is 97 mg/g [63]. The adsorption of Pb by natural sepiolite and iron oxide-coated sepiolite was found that both fit well with the Freundlich and Langmuir isothermal models, with the degree of fit for the Langmuir model was the best (R^2 = 0.990). It can be seen that the reaction mechanism of sulfhydryl modified sepiolite and natural sepiolite and iron oxide coated sepiolite differ. The introduction of sulfhydryl group and the stable coordination bond with the heavy metal ions in the sulfhydryl group have a good influence on adsorption, the adsorption effect of iron coating sepiolite on pollutant metals is higher than that for natural sepiolite, which may be due to replacement of the zeolite water in the sepiolite and increase in adsorption sites [72,73]. In addition, the effect of organic modification (dodecyl benzene sulfonic acid sodium and sodium chloride modified sepiolite) on the adsorption of Pb^{2+} is greater than that of unmodified sepiolite and its mechanism is related to the organic modification providing surface enrichment of macromolecule groups for the adsorption of metal ions. Adsorption isotherm has a good fit to the Langmuir model and pseudo-second-order kinetic equation, the quasi secondary maximum adsorption capacity is 226.8 mg/g [74]. The sepiolite has a good performance for the adsorption of Pb^{2+} after acid soaking and high temperature calcination, and its removal at 50 mg/L Pb^{2+} is 80.9%. The reaction has a best fit to the Freundlich adsorption isotherm [75].

5.4. The Application of Remove Cr in the Wastewater by Sepiolite

The mechanism for sepiolite adsorption of Cr (VI) occurs in two stages. Firstly, to remove part of the contamination by surface of physical adsorption, followed by the likely reduction of Cr (VI) to Cr (III) [76,77].

At pH = 2, it was found that amine functionalized natural and acid-activated sepiolites [78] had the best adsorption effect on Cr (VI), and the adsorption capacity was 37 mg/g and 60 mg/g, respectively. The system showed a good fit to the Freundlich isothermal model compared with Langmuir and D–R isothermal models. It shows that in the process of adsorption should be simultaneously to multiple sites. By analyzing the R^2 value for the pseudo-second-order kinetics, the cause of limited adsorption efficiency in chemical adsorption process can be analyzed. In the study of the adsorption treatment of Cr (VI) with sepiolite-supported nanoscale zero-based iron (S-NZVI), the maximum adsorption capacity of S-NZVI for Cr (VI) was 43.86 mg/g, with its response fitting well to the Freundlich isothermal model, the R^2 is greater than for Langmuir isothermal model, which suggested that that is due to the S-NZVI surface heterogeneity the effects on the Cr (VI) removal, the linear relationship between the removal of Cr (VI) and the input of S-NZVI fits to pseudo-first-order kinetics [77]. In the study of the adsorption of Cr on magnetic modified sepiolite, it was found that the reaction is also strongly related to the Freundlich isothermal model. This shows that the magnetic surface of modified sepiolite exhibits heterogeneity, with adsorption between monolayer and multilayer adsorption mechanisms. The reaction has a good fit to pseudo-second-order dynamics ($R^2 > 0.99$) and the maximum adsorption capacity was 3.6 mg/g. Compared with natural sepiolite, the removal of Cr (VI) by modified sepiolite is much higher than for natural sepiolite [46].

Other studies of natural sepiolite adsorption of Cr^{3+}, Cd^{2+}, and Mn^{2+} showed the best adsorption effect for Cr^{3+} was on natural sepiolite, with good fit to the Langmuir isothermal model. The adsorption process is not only the ion exchange, but also the formation of complex and surface adsorption [79].

6. Regeneration of Sepiolite

The process of regeneration of modified sepiolite treated with wastewater is to restore most of the adsorption capacity of sepiolite, so that the material can be reused, reducing operational costs and preserving resources, which fitswith green production process philosophy. At present, there are few reports on the adsorption and regeneration of sepiolite, and it is of great significance to find an economical and efficient method for the treatment of waste water [80,81].

The method of sepiolite regeneration includes: acid regeneration, alkaline regeneration, and salt regeneration. Jia et al. used salt regeneration to regenerate the sepiolite that had adsorbed Zn^{2+}, it shows that the capacity adsorption of regenerated sepiolite is still high although some minor reduction occurs [82]. The adsorption ability of sepiolite pre- and post-regeneration is shown in Table 2.

Table 2. The adsorption ability of sepiolite before and after regeneration.

Regeneration Methods	Metal (s)	Before Regeneration	After Regeneration	Reference
HCl	Ga^{3+}	The removal is 98.8%	The removal is 94.4% after the fourth cycle of adsorption–desorption	[83]
	Pb^{2+}	equilibrium absorption capacity is 638.9 mg/g	The equilibrium absorption capacity is 489.2 mg/g after fifth cycle of adsorption–desorption	[84]
HNO_3	Pb^{2+}	The saturated adsorption capacity is 114.2 mg/g	The saturated adsorption capacity is 97.6 mg/g	[66]
	Hg^{2+}	The saturated adsorption capacity is 84.6 mg/g	The saturated adsorption capacity is 64.1 mg/g	[66]
	Cd^{2+}	The saturated adsorption capacity is 71.9 mg/g	The saturated adsorption capacity is 52.5 mg/g	[66]
	Fe^{3+}	/	Removal of iron ions was decreased less than 5% after the fourth cycle of adsorption-desorption	[85]
NaOH	Co^{2+}	The sorption capacities of Co^{2+} is 16.02 mg/g,	The sorption capacities of Co^{2+} is 14.50 mg/g after the sixth cycle of adsorption-desorption	[86]
	Cd^{2+}	The sorption capacities of Cd^{2+} is 12.38 mg/g	The sorption capacities of Cd^{2+} is 10.99 mg/g after the sixth cycle of adsorption–desorption	[86]
	Sb	/	Removal efficiency was decreased less than 7% after the fifth cycle of adsorption–desorption	[87]
NaCl	Pb, Zn	/	Removal efficiency was decreased by 21.27% after the fifth cycle of adsorption–desorption	[88]

From Table 2, we can see that the regenerated sepiolite maintains good adsorption properties for metals ions. Li et al. [87] used hydrochloric acid, sodium chloride, and sodium hydroxide in three ways to regenerate the sepiolite. It was found that NaOH had the best treatment effect on the regeneration of sepiolite, and the treatment effect was similar with that of water and sodium chloride, with the use of hydrochloric acid being the poorest. Li et al. [88] used two different kinds of acid, salt and alkaline, to study sepiolite regeneration, and showed that the NaOH and nitric acid had the best effect on sepiolite regeneration, but hydrochloric acid and NaOH treatments were not ideal for sepiolite re-use. Yan showed several different methods for the regeneration of sepiolite: two acid regeneration (HNO_3, HCl), alkaline regeneration, and salt regeneration. It shows that the removal of metal ions was reduced by 19.72% when the regenerating solution is HNO_3, the removal of metal ions was reduced by 16.28% when the regenerating solution is HCl, for NaCl, it was reduced by 5.32% and after five steps, for NaOH, the removal of metal ions was reduced by 6.24% [89]. The results of these trials show that regeneration process is metal specific, which must be considered when developing protocols for full trials and field applications.

7. Conclusions

Pollution from mining and industry continues to be serious environmental problem in Xiangjiang Valley, which is the key rice production area in Hunan province. We need to identify approaches to deal with the protection of the food chain over widespread areas. Adsorption is still an effective technology for removing metals from water, but requires careful consideration to deal with multiple metals, sources, and regional scale contamination. In this review, sepiolite based materials were selected as target adsorbents for heavy metals removal because of their local abundance and potential cost effective application. The characteristics and function of natural and modified sepiolites are reviewed and compared. The sorption performance of modified sepiolite obtained from acid, magnetic, organic, and thermal treatment is significantly improved over natural materials and can be re-used through regeneration. These methods include acid, alkaline, and salt regeneration. The sorption performance of sepiolite after regeneration is greater than 70% of its original performance.

We identified that the relatively low efficiency for heavy metal removal by natural sepiolite was because of the scarcity of exchange sites for contaminant metal ions and capacities are much lower than other synthetic sorbent materials. However, the modification of sepiolite is viable and provides potentially useful adsorption capacity. However, the process of modification also increases the cost of any potential application. In our review, few reports have considered the cost–benefit for modification of sepiolite and application of modified sepiolite on removing metals. Most studies still focus on the mechanism of modification and regeneration of sepiolite. In order to fully establish the potential of sepiolite, as a low-cost and effective adsorbent, further field scale research involving a product life cycle approach is required. This will identify the full potential of local resources to treat local pollution and meet with recent national (China) international (EU–China) agreements on resource conservation and environmental protection.

Author Contributions: Drafting the manuscript, Z.W. & L.L.; Data collection, Z.W., L.L., & N.S.; Critical review, Z.W., & A.H.; Writing–review & editing, A.H., Z.W., & B.R.; Funding Acquisition, A.H.; All authors have given approval of the version to be published.

Funding: This research was part funded by the Hunan Provincial Hundred Talents Program.

Acknowledgments: Zhenghua Wang was supported by the China Scholarship Council for one year study at the University of the West of Scotland.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Paul, D. Research on heavy metal pollution of river Ganga: A review. *Ann. Agrar. Sci.* **2017**, *15*, 278–286. [CrossRef]
2. Nguyen, T.T.; Zhang, W.; Li, Z.; Li, J.; Ge, C.; Liu, J.; Bai, X.; Feng, H.; Yu, L. Assessment of heavy metal pollution in Red River surface sediments, Vietnam. *Mar. Pollut. Bull.* **2016**, *113*, 513–519. [CrossRef] [PubMed]
3. Sun, W.; Xiao, D. Status and control technology of heavy metal pollution. *Energy Energy Conserv.* **2012**, *2*, 49–50. (In Chinese)
4. Wang, H.; Fang, F.; Xie, H. Research situation and outlook on heavy metal pollution in water environment of China. *Guangdong Trace Elements Sci.* **2010**, *1*, 14–18. (In Chinese)
5. Yue, X.; Liu, K.; Lin, X.; Zhou, Q.; Mao, G.; Zhou, B.; Zhao, J. Current status of heavy metal pollution in seven major water systems in China. *Prev. Med. Trib.* **2014**, *3*, 209–213. (In Chinese)
6. Duan, Q.; Lee, J.; Liu, Y.; Chen, H.; Hu, H. Distribution of heavy metal pollution in surface soil samples in China: A graphical review. *Bull. Environ. Contam. Toxicol.* **2016**, *97*, 303–309. [CrossRef] [PubMed]
7. Morton-Bermea, O.; Hernández-Álvarez, E.; González-Hernández, G.; Romero, F.; Lozano, R.; Beramendi-Orosco, L.E. Assessment of heavy metal pollution in urban topsoils from the metropolitan area of Mexico City. *J. Geochem. Explor.* **2009**, *101*, 218–224. [CrossRef]
8. Wang, X.; Zhang, L.; Zhao, Z.; Cai, Y. Heavy metal contamination in surface sediments of representative reservoirs in the hilly area of southern China. *Environ. Sci. Pollut. Res.* **2017**, *24*, 26574–26585. [CrossRef] [PubMed]
9. Paul, D.; Sinha, S.N. Isolation and characterization of a phosphate solubilizing heavy metal tolerant bacterium from River Ganga, West Bengal, India. *Songklanakarin J. Sci. Technol.* **2015**, *37*, 651–657.
10. Rajbanshi, A. Study on heavy metal resistant bacteria in Guheswori Sewage Treatment Plant. *Our Nat.* **2008**, *6*, 52–57. [CrossRef]
11. Wu, W.; Wu, P.; Yang, F.; Sun, D.L.; Zhang, D.X.; Zhou, Y.K. Assessment of heavy metal pollution and human health risks in urban soils around an electronics manufacturing facility. *Sci. Total Environ.* **2018**, *630*, 53–61. [CrossRef] [PubMed]
12. Fang, L.; Li, L.; Qu, Z.; Xu, H.; Xu, J.; Yan, N. A novel method for the sequential removal and separation of multiple heavy metals from wastewater. *J. Hazard. Mater.* **2018**, *342*, 617–642. [CrossRef] [PubMed]

13. Xu, Z.; Gao, G.; Pan, B.; Zhang, W.; Lv, L. A new combined process for efficient removal of Cu (II) organic complexes from wastewater: Fe (III) displacement/UV degradation/alkaline precipitation. *Water Res.* **2015**, *87*, 378–384. [CrossRef] [PubMed]
14. Carro, L.; Barriada, J.L.; Herrero, R.; Vicente, M.E.S.D. Interaction of heavy metals with Ca-pretreated Sargassum muticum algal biomass: Characterization as a cation exchange process. *Chem. Eng. J.* **2015**, *264*, 181–187. [CrossRef]
15. Zhang, W.; Li, J.; Tang, J.; Wang, B.; Huang, F. Advance of the treatment of heavy metal wastewater by adsorption. *Chem. Ind. Eng. Prog.* **2013**, *11*, 2749–2756.
16. Naushad, M.; Ahamad, T.; Al-Maswari, B.; Alqadami, A.A.; Alsheri, S.M. Nickel ferrite bearing nitrogen-doped mesoporous carbon as efficient adsorbent for the removal of highly toxic metal ion from aqueous medium. *Chem. Eng. J.* **2017**, *330*, 1351–1360. [CrossRef]
17. Alqadami, A.A.; Naushad, M.; Alothman, Z.A.; Ghfar, A.A. Novel metal-organic framework (MOF) based composite material for the sequestration of U(VI) and Th (IV) metal ions from aqueous environment. *Appl. Mater. Interfaces* **2017**, *9*, 36026–36037. [CrossRef] [PubMed]
18. Vaizogullar, A.I.; Ugurlu, M.; Kula, I. Comparing adsorption activity of raw sepiolite and CTAB modified sepiolite: Kinetic and adsorption study for removal of Hg^{2+}. *Int. J. Environ.* **2015**, *4*, 19–31. [CrossRef]
19. Renu; Agarwal, M.; Singh, K. Heavy metal removal from wastewater using various adsorbents: A review. *J. Water Reuse Desalin.* **2017**, *7*, 387–419. [CrossRef]
20. Alaba, P.A.; Oladoja, N.A.; Sani, Y.M.; Ayodele, O.B.; Mohammed, I.Y.; Olupinla, S.F.; Daud, W.M.W. Insight into wastewater decontamination using polymeric adsorbents. *J. Environ. Chem. Eng.* **2018**, *6*, 1651–1672. [CrossRef]
21. Malik, D.S.; Jain, C.K.; Yadav, A.K. Removal of heavy metals from emerging cellulosic low-cost adsorbents: A review. *Appl. Water Sci.* **2017**, *7*, 2113–2136. [CrossRef]
22. Ko, D.; Mines, P.D.; Jakobsen, M.H.; Yavuz, C.T.; Hansen, H.C.B.; Andersen, H.R. Disulfide polymer grafted porous carbon composites for heavy metal removal from stormwater runoff. *Chem. Eng. J.* **2018**, *348*, 685–692. [CrossRef]
23. Jlassi, K.; Abidi, R.; Benna, M.; Chehimi, M.M.; Kasak, P.; Krupa, I. Bentonite-decorated calix 4 arene: A new, promising hybrid material for heavy-metal removal. *Appl. Clay Sci.* **2018**, *161*, 15–22. [CrossRef]
24. Zhou, Y.Y.; Wang, Z.H.; Hursthouse, A.; Ren, B.Z. Gemini surfactant-modified activated carbon for remediation of hexavalent chromium from water. *Water* **2018**, *10*. [CrossRef]
25. Xie, X.; Wang, F.; Wang, G.; Mei, R.; Wang, C. Study on heavy metal pollution in surface water in China. *Environ. Sci. Manag.* **2017**, *2*, 31–34. (In Chinese)
26. Cao, C.; Li, Y. Heavy metal pollution in Hunan is shocking. *Environ. Educ.* **2014**, *12*, 27–30. (In Chinese)
27. Tian, B. Interpretation of "Hunan experiment" for heavy metal removal. *People's Friend* **2012**, *8*, 15–16. (In Chinese)
28. Liu, J.; Chen, M. Construction of heavy metal monitoring technical system during twelfth five-year plan in China. *Environ. Sci. Manag.* **2014**, *39*, 125–128. (In Chinese)
29. Qiao, R. Study on the Legal Problems of Water Treatment in Xiangjiang River Basin. Master's Thesis, University of Geosciences, Beijing, China, 2017. (In Chinese)
30. Lei, M.; Qin, P.; Tie, B. The Present Situation and Analysis of Heavy Metal Pollution in Hunan Xiangjiang River Basin. In Proceedings of the Third National Symposium on Agricultural Environmental Science, Tianjing, China, 23–26 October 2009. (In Chinese)
31. Wang, Q.; Wang, S.; Liu, M. Study evaluation on pollution of Xiang River Valley in Hunan Province. *China Water Wastewater* **2004**, *8*, 104–106. (In Chinese)
32. Lei, D. Analysison heavy metals pollution status in Hunan Province and its remediation strategy. *Hunan Nonferrous Met.* **2012**, *1*, 57–60. (In Chinese)
33. Liu, Y.; Gao, L.; Li, Z.; Liu, S.; Huang, K.; Li, J. Analysis on heavy metals pollution status and reasons in Xiangjiang River and discussion on its countermeasures. *Environ. Prot. Sci.* **2010**, *36*, 26–29. (In Chinese) [CrossRef]
34. Wang, L.; Guo, Z.; Xiao, X.; Chen, T.; Liao, X.; Song, J.; Wu, B. Heavy metal pollution of soils and vegetables in the midstream and downstream of the Xiangjiang River, Hunan Province. *J. Geogr. Sci.* **2008**, *18*, 353–362. [CrossRef]

35. Liu, J.; Zhu, Y.; Hu, J.; Peng, G.; Xie, H.; Li, Z.; Peng, G. Characteristics and potential ecological risk assessment of heavy metal pollution in the sediments of middle reaches of Xiangjiang River. *J. Ecol. Rural Environ.* **2017**, *33*, 135–141. (In Chinese)
36. Zhang, K.; Yang, X.; Wu, Y.; Wu, B.; Kuang, X. Pollution characteristics and ecological risk assessment of heavy metals in surface sediments in Changsha-Zhuzhou-Xiangtan Reach, Xiang Jiang River, China. *J. Agric. Resour. Environ.* **2015**, *32*, 60–65. (In Chinese)
37. Fang, X.; Tian, D.; Xie, R. Soil physical and chemical properties of the wasteland in Xiangtan manganese mine. *Xiangtan Manganese Min.* **2006**, *5*, 1494–1501. (In Chinese)
38. Tian, Q.; Zhang, G.; Xie, Y.; Mo, Y. Distribution and ecological risk assessment of heavy metals in surface sediments from main tributary entrances of Dongting Lake. *Asian J. Ecotoxicol.* **2017**, *12*, 191–200. (In Chinese)
39. Zhou, Z. Overview of the resources of sepiolite clay in China and abroad. *Land Resour. Herald* **1985**, *1*, 148–156. (In Chinese)
40. Wang, C.; Wang, G.; Tao, T.; Zhu, M.; Qu, F. Research progress and application status of sepiolite functional green building materials. *Bull. Chin. Ceram. Soc.* **2017**, *10*, 3285–3291. (In Chinese)
41. Zhuang, G.; Gao, J.; Chen, H.; Zhang, Z. A new one-step method for physical purification and organic modification of sepiolite. *Appl. Clay Sci.* **2018**, *153*, 1–8. [CrossRef]
42. Carmona, J.A.; Ramírez, P.; Trujillo-Cayado, L.A.; Caro, A.; Muñoz, J. Rheological and microstructural properties of sepiolite gels. Influence of the addition of ionic surfactants. *J. Ind. Eng. Chem.* **2018**, *59*, 1–7. [CrossRef]
43. Wang, C. Research on Low Density Cement Slurry System for Complex Well Cementing. Master's Thesis, Southwest Petroleum University, Chengdu, China, 2004. (In Chinese)
44. Liang, K. The Mineralogy Research of the Sepiolite and Its Application in Environment Conservation. Ph.D. Thesis, Central South University, Changsha, China, 2008. (In Chinese)
45. Duan, E.; Han, J.; Song, Y.; Guan, Y.; Zhao, W.; Yang, B.; Guo, B. Adsorption of styrene on the hydrothermal-modified sepiolite. *Mater. Lett.* **2013**, *111*, 150–153. [CrossRef]
46. Jia, M.; Dai, Y.; Du, T.; Liu, C. Preparation of magnetically modified sepiolite and adsorption of hexavalent chromium. *Environ. Chem.* **2011**, *30*, 1546–1552. (In Chinese)
47. Liu, C.; Huang, Y.; Yu, F.; Sun, X.; Xu, F.; Wu, W.; Zhong, M. Sepiolite modification on Pb adsorption characteristics. *Environ. Chem.* **2013**, *32*, 2024–2029. (In Chinese)
48. Li, C.; Yu, J. Modification and application of sepiolite in chromium-containing wastewater treatment. *Plat. Finish.* **2013**, *35*, 17–22. (In Chinese)
49. Feng, Y. Study on the Adsorption of Lead and Cadmium by Sepiolite. Master's Thesis, University of South, Hengyang, China, 2007. (In Chinese)
50. Feng, Y.; He, S.; Gao, W.; Li, H. Adsorption of organic matter and heavy metals in wastewater on sepiolite. *Water Purif. Technol.* **2006**, *25*, 63–66. (In Chinese)
51. Shi, T.; Jia, S.; Chen, Y.; Wen, Y.; Chen, J.; Huang, R.; Wang, Z.; Liu, Y. Removal of heavy metal ions from wastewater by modified natural ores as adsorbents: A review. *Technol. Water Treat.* **2009**, *35*, 18–23. (In Chinese)
52. Zhang, Y.; Wang, L.; Wang, F.; Liang, J.; Ran, S.; Sun, J. Phase transformation and morphology evolution of sepiolite fibers during thermal treatment. *Appl. Clay Sci.* **2017**, *143*, 205–211. [CrossRef]
53. Li, H. Study on kinetic and thermodynamics for heating modified sepiolite adsorbing acid blue 62 from aqueous solution. *Pop. Sci. Technol.* **2017**, *19*, 36–38. (In Chinese)
54. Miura, A.; Nakazawa, K.; Takei, T.; Kumada, N.; Kinomura, N.; Ohki, R.; Koshiyama, H. Acid-, base-, and heat-induced degradation behavior of Chinese sepiolite. *Ceram. Int.* **2012**, *38*, 4677–4684. (In Chinese) [CrossRef]
55. Valentin, J.L.; Lopez-Manchado, M.A.; Rodriguez, A.; Posadas, P.; Ibarra, L. Novel anhydrous unfolded structure by heating of acid pre-treated sepiolite. *Appl. Clay Sci.* **2007**, *36*, 245–255. [CrossRef]
56. Chen, W.; Ma, L.; Liu, H. Adsorption capacity of magnetic modified sepiolite humic acid in source water. *J. Hohai Univ. (Nat. Sci.)* **2017**, *45*, 109–115. (In Chinese)
57. Li, C.; Xia, Q.; Dai, B. Research on and modification and application of sepiolite in Lead and Zinc contained wastewater treatment. *Plat. Finish.* **2015**, *37*, 19–26. (In Chinese)
58. Li, C.; Xia, Q.; Cao, Y.; Dai, B.; Song, F.; Liu, Z.; Ge, H. Treatment of nickel-containing wastewater with magnetic modified sepiolite. *Electroplat. Finish.* **2015**, *1*, 47–52. (In Chinese)

59. Yu, T.; Dai, Y.; Wang, W.; Jia, M.; Li, F.; Gong, M. Adsorption characteristic and mechanism of heavy metals onto magnetically modified sepiolite. *Environ. Chem.* **2013**, *32*, 1566–1570. (In Chinese)
60. Zhang, Q. Study on Adsorption Characteristic of Sepiolite. Master's Thesis, Hebei University of Technology, Tianjin, China, 2002.
61. Zhou, Q. Synthesis of Organo-functionalized Clay Minerals and Their Adsorption Performance for Heavy Metal Ions. Ph.D. Thesis, China University of Geosciences, Beijing, China, 2016.
62. Liang, X.; Xu, Y.; Sun, G.; Wang, L.; Sun, Y.; Yang, S.; Xu, Q. Preparation and characterization of mercapto functionalized sepiolite and their application for sorption of lead and cadmium. *Chem. Eng. J.* **2011**, *174*, 436–444. [CrossRef]
63. Liang, X.; Xu, Y.; Wang, L.; Sun, Y.; Lin, D.; Sun, Y.; Qin, X.; Wan, Q. Sorption of Pb^{2+} on mercapto functionalized sepiolite. *Chemosphere* **2013**, *90*, 548–555. [CrossRef] [PubMed]
64. Xie, J.; Chen, B.; Zhang, J.; Liu, J. Adsorption of Hg (II) in water by sulfydryl-modified sepiolite. *Environ. Sci.* **2016**, *37*, 2187–2194. (In Chinese)
65. Jin, S.; Yang, W.; Tang, M. Study on surface modification and application of sepiolite. *Non-Met. Mines* **2001**, *4*, 23–24. (In Chinese)
66. Luo, D.; Yi, P.; Chen, A.; Shi, H. Adsorption of modified unchanged meerschaum on Pb^{2+}, Hg^{2+} and Cd^{2+} in wastewater. *Technol. Water Treat.* **2003**, *29*, 89–91. (In Chinese)
67. Huang, J.; Wu, Z.; Chen, L.; Sun, Y. The sorption of Cd (II) and U (VI) on sepiolite: a combined experimental and modeling studies. *J. Mol. Liquids* **2015**, *209*, 706–712. [CrossRef]
68. Xu, Y.; Liang, X.; Sun, G.; Sun, Y.; Qin, X.; Wang, L. Surface chemical characteristics of sepiolites and their adsorption mechanisms of Pb^{2+}, Cd^{2+} and Cu^{2+}. *J. Agro-Environ. Sci.* **2009**, *28*, 2057–2063.
69. Xu, Y.; Liang, X.; Sun, G.; Sun, Y.; Qin, X.; Wang, L.; Dai, X. Effects of acid and heating treatments on the structure of sepiolite and its adsorption of lead and cadmium. *Environ. Sci.* **2010**, *31*, 1560–1567. (In Chinese)
70. Wang, W.; Dai, Y.; Jia, M.; Li, X.; Du, T. Surface point of zero charge of magnetic sepiolite and adsorption characteristics of cadmium. *Environ. Chem.* **2012**, *31*, 1691–1696. (In Chinese)
71. Xie, J. Study on the Adsorption Characteristics of Sulfydryl-Modified Sepiolite to Hg (II) and Cd (II). Master's Thesis, Southwest University, Chongqing, China, 2016. (In Chinese)
72. Eren, E.; Gumus, H. Characterization of the structural properties and Pb (II) adsorption behavior of iron oxide coated sepiolite. *Desalination* **2011**, *273*, 276–284. [CrossRef]
73. Yamauchi, Y.; Suzuki, N.; Sato, K.; Fukata, N.; Murakami, M.; Shimizu, T. Active mercury (II) ion removal: Stoichiometrically controlled thiol-functionalized mesoporous silica by a mass production spray dry system. *Bull. Chem. Soc. Jpn.* **2009**, *82*, 1039–1043. [CrossRef]
74. Chen, B. Adsorption of Pb (II) by Sodium Dodecyl Benzene Sulfonate Modified Sepiolite. Master's Thesis, Hunan University, Changsha, China, 2013. (In Chinese)
75. Liu, Y.; Jiang, M. Treatment of mult-metals wastewater using sepiolite modified with acid and heating. *Guangzhou Chem. Ind.* **2011**, *21*, 137–139. (In Chinese)
76. Marjanović, V.; Lazarević, S.; Janković-Častvan, I.; Potkonjak, B.; Janaćković, Đ.; Petrović, R. Chromium (VI) removal from aqueous solutions using mercaptosilane functionalized sepiolites. *Chem. Eng. J.* **2010**, *166*, 198–206. [CrossRef]
77. Fu, R.; Yang, Y.; Xu, Z.; Zhang, X.; Guo, X.; Bi, D. The removal of chromium (VI) and lead (II) from groundwater using sepiolite-supported nanoscale zero-valent iron (S-NZVI). *Chemosphere* **2015**, *138*, 726–734. [CrossRef] [PubMed]
78. Marjanović, V.; Lazarević, S.; Janković-Častvan, I.; Jokić, B.; Janaćković, D.; Petrović, R. Adsorption of chromium (VI) from aqueous solutions onto amine-functionalized natural and acid-activated sepiolites. *Appl. Clay Sci.* **2013**, *s80–81*, 202–210. [CrossRef]
79. Kocaoba, S. Adsorption of Cd (II), Cr (III) and Mn (II) on natural sepiolite. *Desalination* **2009**, *244*, 24–30. [CrossRef]
80. Fang, J. Preparation of Co-Fe-Sepiolite Sorbent and Its Regenerated by High Temperature Gas Catalytic Oxidation Method. Master's Thesis, Guangxi Normal University, Guilin, China, 2010.
81. Feng, J. Study on the Modification of Sepiolite and Its Regeneration by Catalytic Oxidation Process Using High Temperature Gas Flow. Master's Thesis, Guangxi Normal University, Guilin, China, 2014.
82. Jia, N.; Wang, H.; Huo, J. The study on adsorption of Zn^{2+} by modified meerschaum. *Chinamining Mag.* **2006**, *4*, 70–72. (In Chinese)

83. Liang, K.; Wang, D.; Long, L.; Xi, C. Study on the recovery of gallium using from acid lixivium of zinc modified meerschaum residues. *Acta Mineral. Sin.* **2006**, *26*, 277–280. (In Chinese)
84. Zheng, Y.; Xie, Y.; Wang, A. Adsorption of Pb^{2+} onto chitosan-grafted-poly (acrylic acid)/sepiolite composite. *Environ. Sci.* **2009**, *30*, 2575–2579. (In Chinese)
85. Gaber, S.; Haija, M.A.; Priyabrata, P.; Selvaraj, M.; Banat, F. Removal of iron from industrial lean methyldiethanolamine solvent by adsorption on sepiolite. *Sep. Sci. Technol.* **2018**, *53*, 404–416. [CrossRef]
86. Yu, S.; Zhai, L.; Zhong, S.; Qiu, Y.; Cheng, L.; Ren, X. Synthesis and structural characterization of magnetite/sepiolite composite and its sorptive properties for Co (II) and Cd (II). *J. Taiwan Inst. Chem. Eng.* **2016**, *59*, 221–228. [CrossRef]
87. Li, S.; Dai, Y.; Li, N.; Shi, L.; Tang, W. Preparation of iron modified sepiolite and its adsorption characteristics of antimony. *Technol. Water Treat.* **2009**, *35*, 49–52. (In Chinese)
88. Li, C.; Yu, J. Modification and application of sepiolite inlead and zinc smelter wastewater treatment. *Sci. Technol. Eng.* **2013**, *13*, 9153–9157. (In Chinese)
89. Yan, H. The Modification of Sepiolite and Its Application in Heavy Metal Wastewater Treatment. Master's Thesis, Northwest University, Chongqing, China, 2013. (In Chinese)

International Journal of
Environmental Research and Public Health

MDPI

Article

Using a Clustering Approach to Investigate Socio-Environmental Inequality in Preterm Birth—A Study Conducted at Fine Spatial Scale in Paris (France)

Severine Deguen [1,2,*], Nina Ahlers [1], Morgane Gilles [1], Arlette Danzon [3], Marion Carayol [3], Denis Zmirou-Navier [1,4,5] and Wahida Kihal-Talantikite [6]

1 School of Public Health (EHESP), DSET&GS, 35043 Rennes CEDEX, France; Nina.Ahlers@ucsf.edu (N.A.); morgane.gilles@eleve.ensai.fr (M.G.); denis.zmirou@ehesp.fr (D.Z.-N.)
2 Department of Social Epidemiology, Institut Pierre Louis d'Epidémiologie et de Santé Publique (UMRS 1136), Sorbonne Universités, UPMC Univ Paris 06, INSERM, 75012 Paris, France
3 Service de Protection Maternelle et Infantile, Direction des Familles et de la Petite Enfance, Mairie de Paris, 75196 Paris, France; arlette.danzon@paris.fr (A.D.); marion.carayol@paris.fr (M.C.)
4 School of medicine, Lorraine University, 54000 Nancy, France
5 Inserm, Irset (Institut de recherche en santé, environnement et travail)—UMR-S 1085, F-35000 Rennes, France
6 Laboratoire Image Ville Environnement, LIVE UMR 7362 CNRS, University of Strasbourg, 6700 Strasbourg, France; wahida.kihal@live-cnrs.unistra.fr
* Correspondence: severine.deguen@ehesp.fr; Tel.: +33-2-99-02-2805

Received: 19 July 2018; Accepted: 29 August 2018; Published: 31 August 2018

Abstract: *Background & Objectives*: Today, to support public policies aiming to tackle environmental and health inequality, identification and monitoring of the spatial pattern of adverse birth outcomes are crucial. Spatial identification of the more vulnerable population to air pollution may orient health interventions. In this context, the objective of this study is to investigate the geographical distribution of the risk of preterm birth (PTB, gestational age $\leq$36 weeks) at the census block level in in city of Paris, France. We also aimed to assess the implication of neighborhood characteristics including air pollution and socio-economic deprivation. *Material & Methods*: Newborn health data are available from the first birth certificate registered by the Maternal and Child Care department of Paris. All PTB from January 2008 to December 2011 were geocoded at the mother residential census block. Each census block was assigned a socioeconomic deprivation level and annual average ambient concentrations of NO_2. A spatial clustering approach was used to investigate the spatial distribution of PTB. *Results*: Our results highlight that PTB is non-randomly spatially distributed, with a cluster of high risk in the northeastern area of Paris (RR = 1.15; p = 0.06). After adjustment for socio-economic deprivation and NO_2 concentrations, this cluster becomes not statistically significant or shifts suggesting that these characteristics explain the spatial distribution of PTB; further, their combination shows an interaction in comparison with SES or NO_2 levels alone. *Conclusions*: Our results may inform the decision makers about the areas where public health efforts should be strengthened to tackle the risk of PTB and to choose the most appropriate and specific community-oriented health interventions.

Keywords: air pollution; neighborhood deprivation index; preterm birth; spatial approach

1. Introduction

Adverse birth outcomes are important public health issues including preterm birth (PTB) and low birthweight (LBW). Over the past 20 years, the literature confirmed that, in developed countries, PTB remains a risk factor of adverse health outcomes including neonatal mortality and short- and long-term morbidity [1–8]. PTB is also recognized as a risk factor for LBW, delayed motor and social skills, as well as learning disabilities [9].

Various contextual determinants, characterizing place where people live, have been reported to be associated with births outcomes, including socio-economic, demographic characteristics and environmental factors such as exposure to environmental contaminants. Several studies concluded that prenatal development is a window of high susceptibility to the adverse impact of environmental nuisances, in particular ambient air pollution [10–17] More specifically, studies revealed that maternal air pollution exposure such as particulate matter ≤10 and ≤2.5 μm in diameter (PM_{10} and $PM_{2.5}$) and nitrogen dioxide (NO_2) reduced birth weight and increased the odds of low birth weight and preterm birth, [11,12,18]. While findings for PTB remains inconsistent in the literature (depending, in particular, on the study design, exposure assessment, pregnancy periods and adjustment for confounders [12,18–21]), experimental studies support plausible biological mechanisms explaining, for instance, how air pollution exposure could reduce gestational age via placental inflammation linked to oxidative stress [22].

In addition, some studies suggested that the adverse health effect of maternal environmental exposure may be influenced by other contextual or individual characteristics (such as sex, socioeconomic position and psychological factors [12,23–25]). Many authors concluded that health risk related to environmental exposure may be different according to the socioeconomic level of populations [26–30]. For instance, Yi et al. in 2010 found a three-fold increase in the PTB risk for an increase in PM_{10} concentrations among babies born in low-income groups [27] and Carbajal-Arroyo et al. in 2011 revealed a significant increase in the risk of all-cause mortality only among infants with low and medium SES [31]. These social inequalities in air pollution exposure of pregnant women and newborns are a public health issues. Additional studies are needed in Europe to improve our level of understanding concerning the underlying mechanisms explaining the existence of environmental inequality and to tackle this public health issues [32].

In epidemiological studies, quantifying the strength of the association between risk factors and health outcomes constitute pivotal information to document causality. However, these measures provide limited guidance for effective policies aimed at improving population health and reducing health inequalities. Spatial approaches may bring, in complement, useful information to help policymakers to elaborate on the choice of intervention.

To our knowledge, few epidemiological studies investigated the spatial distribution of PTB. For instance, in Philadelphia, using a descriptive geographic-spatial approach conducted at census tract level, Boch et al. investigated the geographical patterns of the prevalence of PTB and examined its relationships with race, poverty, crime, and natality [33]. Today, the use of geographical information system for mapping adverse birth outcomes and maternal addresses, while more and more popular, is not sufficient to highlight areas that exhibit a higher risk. Additional spatial analyses are required to explore the spatial pattern of adverse birth outcomes and the spatial implication of neighborhood characteristics that may explain it. In Worcester, Ogneva-Himmelberger et al. in 2015 studied the spatial distribution of preterm births by racial groups to identify spatial clusters using mother's residence address such as point location. Using two different spatial clustering methods, they analyzed associations between PTB and neighborhood characteristics including distance to major roads, exposure to hazardous air pollutants from stationary sources, access to vendors of healthy food, and access to green space and parks [34].

To our knowledge, no study has investigated the geographical distribution of PTB and its spatial association with the level of deprivation and the concentrations of air pollution measured at a small spatial scale. Indeed, to assess spatial patterns of health outcomes and its risk factors, fine spatial scale has been recommended in order to increase the homogeneity of specified variables within each area (such socioeconomic characteristics in this present study) and maximized differences between areas [35,36]; it is particularly important, when the study area, as in Paris city, presents a high population density per km^2 with contrasted socioeconomic profiles. In addition, investigations of the spatial distribution of health events and risk factors conducted at the state or county level may not provide useful results for development of local policies or local decisions aiming to tackle social and environmental inequalities [37]. Small- spatial scale analyses appear to be an appropriated statistical unit to identify areas of high risk of PTB for targeted interventions and for reduction of inequalities in PTB.

In our study, spatial approaches appear to be the most appropriated to examine the spatial distribution of health risk and neighborhood characteristics. Using Kulldorf methods, we sought to perform clustering analysis to map the spatial distribution of the relative risks and to investigate the spatial implication of neighborhood characteristics. Unlike more traditional epidemiological studies which implement logistic regression to estimate impact of air pollution on the risk of preterm birth, with our approach we aim to answer to the same objective with an additional constraint related to spatial distribution of the health event. For example, Sabel et al. revealed that the relative risk (RR) of the pneumonia and influenza cluster adjusted for age, sex and deprivation is 1.92 whereas, the relative risk for the age, sex, deprivation and air pollution adjusted cluster is 1.99, respectively. However, these two clusters were not located in the same part of the territory and include different numbers of Census Area Unit (CAUs) while the risks estimated were similar [38]. More recently, Kihal et al. in France, found that the RR of end-stage renal disease (ESRD) incidence adjusted for sex, age and rural typology was 1.5, whereas the RR adjusted for age, sex and socioeconomic deprivation was 1.44. However, even estimated RR were similar; the two clusters were located at different part of the region: the first in the South-western part and the second in the extremely western Bretagne) and contained also different numbers of census blocks [39].

In this context, the localization of small geographical areas that exhibit a high PTB risk and their fine description may facilitate actions closely targeted towards areas most at risk: it is precisely the objective of this study. This work is not intended to reveal any causal pathway between neighborhood characteristics and PTB risk, an objective that requires other study designs [40,41].

2. Materials

2.1. Study Area

The study area is the city of Paris which counts about 2,250,000 inhabitants. We used the smallest census unit area whose aggregate data can be used on a routine basis: the Ilots Regroupés pour l'Information Statistique (IRIS: the French acronym for 'blocks for incorporating statistical information'). The IRIs is a sub-municipal French census block defined by the National Institute of Statistics and Economic Studies (INSEE). This statistical unit averages 2000 inhabitants and is constructed to be as homogenous as possible in terms of socioeconomic and demographic characteristics and land use. Paris is subdivided into 992 census blocks with a mean population of about 2199 inhabitants and a mean area of 0.11 km^2.

2.2. Health Data: Preterm Birth

The preterm birth case has be defined according to the definition of World Health Organization (WHO) [42,43]: it corresponds to a neonate born before 37 weeks of pregnancy (gestational age ≤36 weeks). The preterm birth cases were identified from the first birth certificate information registered over the period 2008–2011 by the Maternal and Child Care department of Paris (named PMI, for Protection Maternelle et Infantile). This certificate is completed by parents and health professional before exit of the maternity and then sent to the PMI unit of the department of residence.

All the postal addresses of mothers' residency were geocoded at the census block level. For confidential concerns, to be in agreement with the ethical authorization provided for this study, it was not possible to keep individual localization of the newborn. The number of cases was aggregated at census blocks level for the statistical analysis.

2.3. Air pollution: Nitrogen Dioxide (NO_2)

Annual NO_2 concentrations were modelled from a grid of 25 × 25 m resolution throughout the study period (2008–2011) by the local air quality monitoring networks corresponding to the Ile de France region (AirParif: http://www.airparif.asso.fr/). The ESMERALDA inter-regional platform for air quality mapping and forecasting (www.esmeralda-web.fr) provided background pollution data, while the STREET dispersion model [44] was used for traffic-related pollution.

AirParif used a deterministic model which integrates various input parameters including linear (main roads), surface (diffuse road sources, residential and tertiary emissions) or industrial point sources and meteorological data (temperature, wind speed and direction, relative humidity, barometric pressure). More than 200 points sources were selected from the regional emission inventory. Emissions for traffic roads were estimated using the regional traffic network and the COPERT III European database for the 2002–2006 period, and COPERT IV for the 2007–2012 period. Concerning meteorological data, the Mesoscale Meteorological model (MM5: www.mmm.ucar.edu/mm5) developed by the Division of the NCAR Earth System Laboratory (NESL) was used. The NO_2 background concentrations were determined by combining monitored NO_2 concentrations from monitoring stations and those modeled at a regional scale from the ESMERALDA. The NO_2 road traffic concentrations estimated from the STREET software model were added to NO_2 background concentrations. Air pollutant concentrations were then aggregated at census block scale in order to obtain annual mean NO_2 concentration for each census block (for more detail see Kihal-Talantikite et al. [45]; Deguen et al. [46]) over the study period.

NO_2 pollutant was chosen for several reasons: while data of PM_{10}, $PM_{2.5}$ and NO_2 were available at the time of the study, we privileged the NO_2 because this pollutant is recognized to be a good tracer of traffic and other combustion sources (major problems in Paris) [47]. It is also well known that the spatial variability of NO_2 concentrations is higher than that of particulate matter [48]; a crucial point especially for spatial analysis; previous studies already revealed that exposure to NO_2 may be related to adverse birth outcomes [11,12,49–51]. NO_2 has been shown to be the best available indicators of local traffic emissions [48]. Finally, the correlation coefficients between NO_2 and, PM10 and $PM_{2.5}$ are very high: $r = 0.95$ and 0.93, respectively (see Figure S1 and Table S1 in Supplementary Materials).

2.4. Socioeconomic Deprivation Index

To characterize the level of socioeconomic deprivation at the census block scale, an index was created (details in another study [52]). The socioeconomic and demographic variables were provided by the 2010 national census at census block level. Briefly, a principal component analysis was used to select 15 variables out of 41 initial socioeconomic and demographic variables. Previous ecological studies have demonstrated this index's ability to capture environment-related socio-spatial inequalities in France [53–55].

3. Methods

3.1. Spatial Methodology

To investigate the spatial distribution of PTB risk at census block level in Paris, we used a spatial scan statistic approach implemented in the SaTScan software [56].

The null hypothesis (H0) tests whether the risk of PTB is equi-distributed throughout the study area. The alternative hypothesis (H1) tests if there is an elevated PTB risk within the cluster in comparison with census blocks outside the cluster.

In our study, the Poisson probability model implemented in the SaTScan software [56] was chosen as *cluster analysis method*. The number of PTB cases (a rare event) in each census block is assumed to follow a Poisson distribution. The input data for the Poisson model are the cases (PTB) and the population at risk (all birth) to determine if there is significant spatial clustering of the cases.

We therefore compute a relative risk (RR) in each census block weighted by the population at risk count in each census block. The RR is estimated as the observed divided by the expected cases within the cluster divided by the observed divided by the expected cases outside the cluster (Equation (1)):

$$RR = \frac{c/E[c]}{(C-c)/(E[C]-E[c]} = \frac{c/E[c]}{(C-c)/(C-E[c])} \quad (1)$$

where c is the number of observed PTB cases within the cluster and C is the total number of PTB cases in the data set. Note that since the analysis is conditioned on the total number of cases observed, $E[C] = C$.

The procedure to identify the most likely cluster is structured as follow. First, a circle of radius, varying from zero up to 50% of the population size [57], is placed at the centroid of every census blocks. Second, the circle moves across the study area to compare the PTB rate within the circle with what would be expected under a random distribution. Therefore, an infinite number of circles were created around each centroid, with the radius anywhere from zero up to a maximum so that at most 50 percent of the population is included.

The scan statistic approach is likelihood based. The most likely cluster can be selected and tested for statistical significance. The likelihood function for the Poisson model is detailed in Equation (2):

$$\left(\frac{c}{E(c)}\right)^{c}\left(\frac{C-c}{C-E(c)}\right)^{C-c} I\,() \quad (2)$$

where C is the total number of PTB cases, c is the observed number of PTB cases within the window and $E[c]$ is the covariate adjusted expected number of PTB cases within the window under the null hypothesis. Note that since the analysis is conditioned on the total number of cases observed, C-$E[c]$ is the expected number of cases outside the window. I () is an indicator function.

The identification of the most-likely clusters is based on a likelihood ratio test [58] with an associated p-value obtained using Monte Carlo replications [59]. The number of Monte Carlo replications was set to 999 to ensure adequate power for defining clusters and considered a 0.05 level of significance (p value derived from 999 replications).

3.2. Analytical Strategy and Results Interpretation

- When a significant most-likely cluster (with $p < 0.05$) was detected, the next step consist in taking into account the neighborhood characteristics to see whether or not the significant cluster can be explained by them. Spatial analyses were structured in four successive steps: A crude (unadjusted) analysis, to identify and localize the most-likely cluster of high risk of PTB.
- An adjusted analysis for NO_2 concentrations
- An adjusted analysis for socioeconomic deprivation index
- One final adjusted analysis for air pollution and socioeconomic deprivation index including interaction between the two variables.

To incorporate covariates in the model, we categorized NO_2 concentrations and socioeconomic deprivation index into five groups according to the quintile of their distribution. Because the SaTScan software does not allow for an interaction term to be accommodated in the model, we created several dummy variables combining the socioeconomic deprivation and the air pollution categories.

At the first step, a statistically significant test means that the risk of PTB is not randomly distributed in the city of Paris: a cluster of census blocks presents a significant increase in PTB risk in comparison with census blocks located outside the cluster [59].

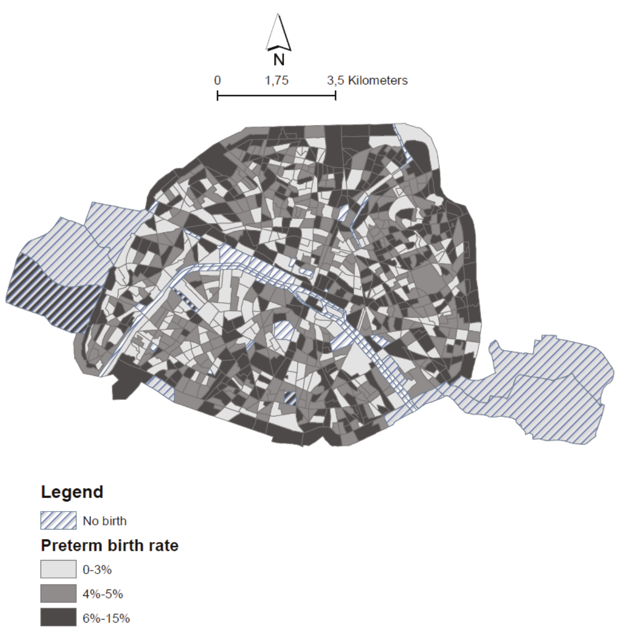
For the three others steps, when the models are adjusted on one or more co-variables, according to the Kulldorff studies [57], several statistical criteria were used to test the H_0 hypothesis: the cluster's localization (the shift or the disappearance of the cluster, or no changes), the level of statistical significance of the cluster and the likelihood ratio value of each model.

According to these criteria, there are three possible results: If, after adjustment, the most likely cluster remains in the same location, (whether or not this cluster is significant) and its likelihood ratio decreases, then it means that the variable(s) incorporated in the model explain partially the excess risk [56]; If the most likely cluster shifts (the centroid of the cluster changes), this suggests that the covariate(s) in the model explain the cluster's excess risk [56] allowing the identification of second cluster. Finally, if the most likely cluster disappears totally, it means that the adjusted PTB risk is now randomly distributed over the study area. To map and visualize the spatial location of the statistically significant most likely clusters, we used ArcGis software (ESRI, Meudon, France).

4. Results

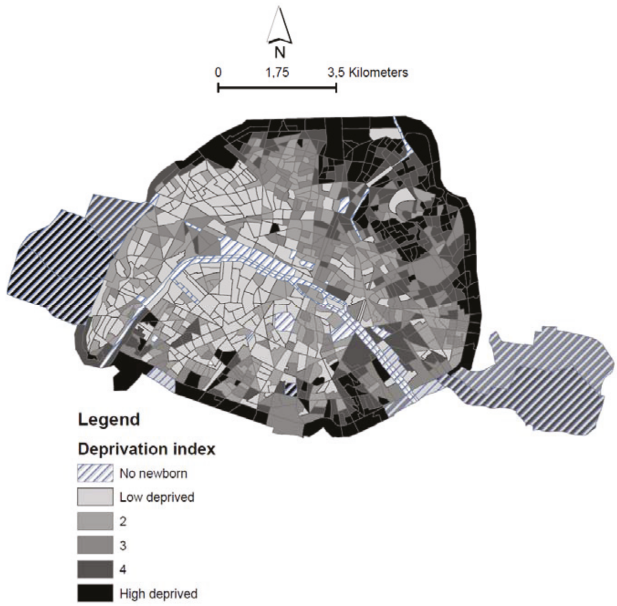
4.1. Description of the Population

A total of 115,112 births were recorded during the study period 2008–2011. After exclusion of all birth with unknown birth weight, gestational age, with birth weights less or 500 g, we counted 110,746 singleton births (about 3.8% of the total births were excluded). When, we excluded also newborn without address (about 4.9%); the total singleton births include in the study is 105,346. Among them, 4871 births occurred before 37weeks of pregnancy. The rate of PTB in Paris was 4.6% overall during the study period (2008–2011). Figure 1A shows that low PTB rates, below 3%, are concentrated in the west-central part of Paris while in the north eastern areas, the PTB rate reaches 6 −15%. The geographical pattern of socioeconomic deprivation index is readily observed: the wealthiest census blocks are located in the western part of Paris while the most socioeconomically deprived neighborhoods are located in the northeast and along the perimeter (much trafficked highway) of Paris (Figure 1B). During study period, the PTB rate among women living in deprived census blocks is 4.9% (n = 241) compared to 3.18% (n = 155) among women living in less deprived census blocks (Significant Kruskal-Wallis test: p-value < 0.0001).

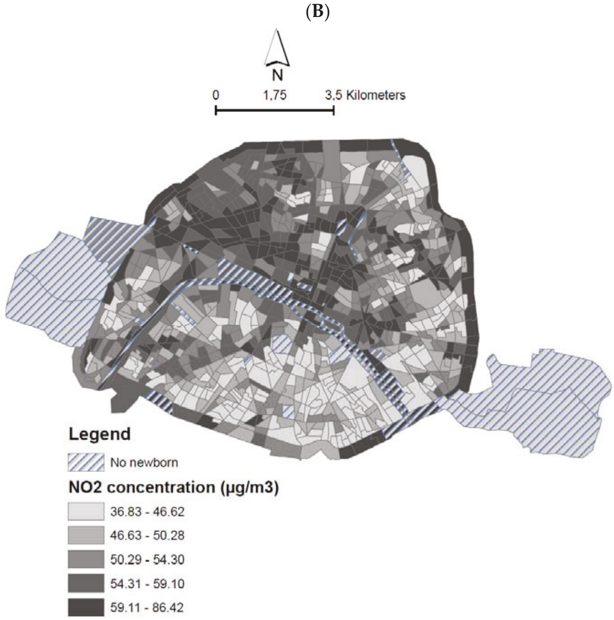


(**A**)

Figure 1. *Cont.*



(**B**)



(**C**)

Figure 1. (**A**): Spatial distribution of crude preterm birth rate in census block areas within Paris; (**B**): Spatial distribution of socio-economic deprivation index in census block areas within Paris; (**C**): Spatial distribution of NO_2 average concentrations from 2008 to 2011 in census block areas within Paris.

All the census blocks have an annual average concentration of NO_2 over the study period 2008–2011 higher than the European limit fixed to 40 µg/m^3. The spatial distribution of the NO_2 concentrations reveals a clear gradient from the north-western part of the city (the highest concentrations level >55.8 µg/m^3) to the south-east part (the lowest concentrations level <50.6 µg/m^3) (Figure 1C).

4.2. Neighborhood Socio-Economic Deprivation, NO_2 Ambient Air Concentrations and Spatial Distribution of PTB in Paris

Figure 2 highlights the census blocks including in the most likely clusters of high risk of PTB, their location and spatial shift of centroids from unadjusted clusters to covariate-adjusted clusters. Table 1 summarizes the results of the spatial analyses: the most likely clusters, the number of census blocks, radius and relative risks (RR, the ratio of the observed- to-expected number in each census blocks estimated by SaTScan) for each cluster.

Table 1. Summary statistics of the most likely clusters of preterm birth risk and spatial relocation resulting from the adjustment analysis.

Analysis	Control Variables	Cluster Radius	No of Census Blocks/No. of Birth in the Cluster	No of Expected Cases	No. of Observed Cases	RR	LLr	Shift	*p*-Value
Unadjusted [a]									
	No adjustment	2816.01	169/25,503	1179.94	1310	1.15	9.23	-	0.06
Adjusted [b]									
	1 Annual concentration of NO_2	1125.2	17/2814	130.84	181	1.40	8.84	Same zone	0.08
	SES[c] index	673.67	19/2396	104.95	140	1.34	5.42	Yes	0.81
	NO_2 and SES level	673.67	19/2396	106.95	140	1.32	4.76	Yes	0.97

RR: relative risk; LLr: log likelihood ratio; [a] Unadjusted analysis, to identify and localize the most likely cluster(s) of high risk of PTB; [b] Adjusted analysis for (1) NO_2 concentration, (2) socio-economic deprivation index, (3) NO_2 concentration and socio-economic deprivation index; [c] Socio-economic deprivation index.

Unadjusted analysis (Figure 2A) reveals that the most likely cluster is located in the northeast part of Paris. Within the cluster, the risk of PTB is 1.15 times greater than in the rest of the study area (p-value < 0.06; Table 1). A total of 169 census blocks composes this most likely cluster, corresponding to about 25,503 inhabitants. The secondary cluster detected is not statistically significant (p-value = 0.89).

After adjustment for NO_2 concentrations (Figure 2B), the most likely significant cluster is reduced (the radius decreases) and hosts 17 census blocks and 2,814 inhabitants. The risk of PTB increases in comparison with the crude estimate (RR = 1.40, p-value = 0.08). The centroid of the cluster shifts and the likelihood ratio slightly decreases from 9.23 (crude model) to 8.84 (adjusted model on air pollution) (Table 1), which suggests that the spatial distribution of NO_2 concentrations partially explain the excess risk of PTB observed in the unadjusted analysis.

After adjustment for socio-economic deprivation (Figure 2C), the most likely significant cluster shifts in South-Eastern Paris and the radius substantially decreases in size as well as the likelihood ratio (from 9.23 in the unadjusted model to 5.42) (Table 1). The remaining excess risk becomes not significant (RR = 1.34, p-value = 0.8). This indicates that socioeconomic deprivation explains a great part of the excess risk of PTB observed in the unadjusted analysis.

After joint adjustment for socioeconomic deprivation index and NO_2 concentrations: the most likely cluster totally disappears. The likelihood ratio falls from 9.23 to 4.76; we also observed a likelihood ratio decrease when comparing with the model adjusted for socioeconomic deprivation index only (Table 1). The most likely cluster is not significant and located in the same zone in South-Eastern Paris (RR = 1.32; p-value = 0.97). This result indicates that the excess risk of PTB detected from the unadjusted analysis is entirely explained, by the spatial distribution of NO_2 concentrations and socioeconomic deprivation.

In our study, the major finding is that while adjustment for socioeconomic deprivation level was the essential variable explaining the most likely cluster (as shown in Table 1), further adjustment for NO_2 concentrations reduces the LLR to a larger degree than that obtained in the model with socioeconomic deprivation level alone or with the NO_2 concentrations alone.

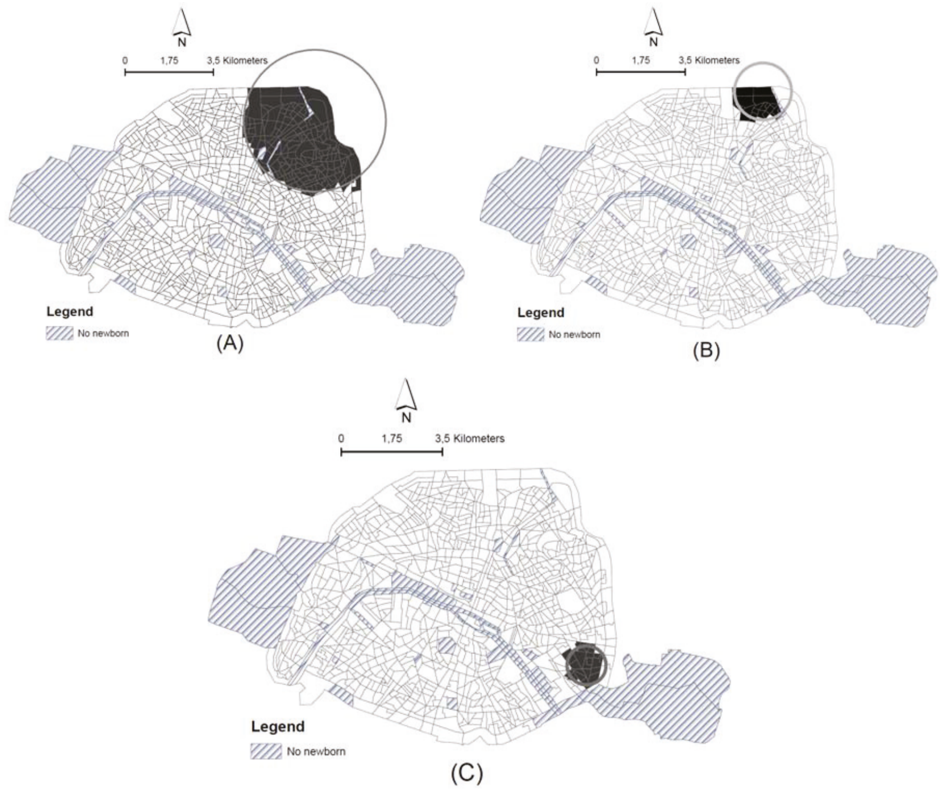


Figure 2. Spatial relocation of the most likely cluster of unadjusted PTB risk (**A**); after adjustment for NO_2 exposure (**B**); after adjustment for NO_2 exposure and socio-economic level (**C**). Legend: the dark area represents the census blocks included in the most likely cluster of high risk of PTB.

5. Discussion

To our knowledge, such a work, exploring the spatial association of neighborhood characteristics on geographical variations of PTB at such a small-scale level had never been performed. For this reason it is difficult to compare our findings with others. Our study shows that neighborhood socioeconomic deprivation and average NO_2 concentrations over years need to be considered in the interpretation of the spatial disparities in PTB in the city of Paris.

First, not surprisingly, NO_2 concentrations only explained a very small part of the spatial variations of PTB across different census blocks. While several studies [12,18] have suggested that maternal exposure to ambient air pollutants (PM_{10}, $PM_{2.5}$, NO_2) are associated with various birth outcome, the evidence regarding preterm birth is mixed and not conclusive. Some studies reported significant associations between exposure during pregnancy to NO_2 and PTB [14,60–63], while others did not [20,64–68]. Recently, Estarlich et al. reported a suggestive association between residential exposure to NO_2 during pregnancy and PTB among pregnant women who spent more time at home [69]. They found that exposure during the second trimester and during the whole pregnancy was associated with a higher risk of PTB. Johnson et al. in 2016, did not confirm the association between NO_2 exposure and PTB in New York City. Using the proximity to traffic as a proxy for air pollution exposure, several studies show that the risk of preterm birth infants is significantly higher among mothers who live near freeways or roadways or to major roads [70–75].

Several biological pathways emerge from the literature to explain the potential impact of exposure to NO_2 on PTB. Potential etiologic factors for PTB include inflammation, oxidative stress and cardiovascular alterations [76,77]. Some studies suggest that maternal exposures to NO_2 can increase the risk of preterm delivery, via oxidative stress [78]. More recently, a second pathway through which NO_2 could alter pregnancy outcomes was proposed. Some studies [61,73] suggest that traffic-related air pollution can related to some cause of PTB such as Preterm premature rupture of membranes (PROM).

Secondly, our findings revealed that the spatial distribution of neighborhood socioeconomic deprivation index explained a great part of spatial repartition of the excess risk of PTB observed in the crude analysis. This finding is coherent with previous works documenting the existence of a social gradient of adverse pregnancy outcome including PTB. Majority revealed an inverse association between PTB and various socioeconomic measures such as income [79–83], unemployment [84], composite socio-economic score including Towsend, carstairs or other socioeconomic deprivation index [85–89].

Recent literature review and meta-analysis concluded that living in a deprived neighborhood is associated with risk of preterm birth [40,41]. Vo et al. in 2014 found that odds ratios for preterm delivery significantly increased in the most deprived neighborhood quintile compared with the least deprived quintile (odds ratio 1.23, (95% CI:1.18–1.28)) [41]. Ncube et al. in 2016 estimated an excess risk of PTB equal to 27% (95%CI: 16%, 39%) among the most disadvantaged neighborhoods compared with least disadvantaged [40].

Many hypotheses have been formulated explaining the pathways through which socioeconomic status could be a potential risk factor of adverse pregnancy outcome including PTB:

(i) Psychological factors such as stressful life events or lack of social support, cohesive social networks and reciprocal exchanges between residents [90–92].
(ii) Unhealthy lifestyle featuring factors such as smoking or poor maternal nutrition and excess alcohol consumption especially around the time of conception [91,93–97].
(iii) Barriers and facilitating factors in access to healthcare such as availability of care, the ability to get to and pay for available care, or to seek and utilize available care [30,98].

The accumulation of these risk factors which is more common in deprived neighborhood [99], can contribute to maternal stress in turn can lead to higher levels of corticotropin-releasing hormone and cortisol which could trigger contractions and/or the premature rupture of the membrane resulting in PTB [100].

Finally, interestingly, our findings showed that the combination of socioeconomic deprivation level and NO_2 concentrations, tacking account the interaction, explain a larger part of the excess risk of PTB estimated in the north-eastern Paris in comparison with analysis considering only socioeconomic deprivation level or NO_2 concentrations (even if the contribution of air pollution is marginal compared to the one of socioeconomic deprivation index). These findings are coherent with previous epidemiological studies. For instance, in the U.S. State of Georgia, Hao et al. in 2015 [101] found that the strength of association between NO_2 and PTB is higher for low education pregnant women. In California, Padula et al. 2014 confirmed a stronger association among pregnant women living in low socioeconomic status neighborhoods [60].

Two main hypotheses are more likely to explain the spatial implication of both NO_2 exposure and socio-economic deprivation in geographical distribution of PTB.

(i) The first mechanisms—*vulnerability differential*—could explain the greater susceptibility to NO_2 exposure of women living in the most deprived neighborhoods. Several studies demonstrated that people with a lower socio-economic status may be more vulnerable to the health effects of proximity to road, air pollution and noise exposure because they experience poorer health due to their economic and psychosocial conditions [15]. Living in communities with lower household income and education levels would also tend to increase vulnerability level to air pollution [102].

(ii) The second mechanism—*combined vulnerability differential with exposure differential*—may explain the greater susceptibility to NO_2 exposure of women living in the deprived neighborhood. Although a majority of studies have found that people living in the most deprived neighborhoods may be more vulnerable to environmental nuisances, some authors have hypothesized that those living in middle deprived neighborhoods may have also a particular vulnerability. In this context, high NO_2 exposure may act on this particular sensitive subpopulation, as an exacerbating factor, which, in combination with unfavorable living conditions, could generate greater health effects than in the rest of the population. The assumption of a synergy of differential exposure and vulnerability to explain our findings therefore seems highly probable.

Some research suggest that socioeconomic deprivation is spatially correlated with air pollution [103,104], and thus may have synergistic health effects through common biological pathways (e.g., chronic stress-induced inflammation, or dysregulation of immune and endocrine systems [105]). Clougherty et al. observed that a heightened susceptibility to pollution, associated with violence exposures or with fear thereof, may lead to synergistic health effects of social and physical environmental conditions. Bandoli et al. provide evidence of synergistic effects of air pollution and psychosocial stressors [106].

Strengths and Limitations

One strength of this work is the use of small area-level analyses allowing a correct understanding of the geographic patterns of PTB. Moreover, this type of analysis is crucial for revealing local-level health inequalities that are often masked when analysis is produced at large spatial scales.

Unlike geographical information system approach used to map and to visualize the spatial trends of PTB risk, in our study, we use a spatial clustering approach allowing us to identify areas of significantly elevated risk of PTB and to investigate spatial implications of adjustment for neighborhood characteristics. Another strength of our study is the databases used in our analysis to investigate PTB in France:

- *Health data:* the advantage of the data used in our study is the rate of completeness of the data which reach 93% on average and the large population size, resulting in a small variability of our estimates [107]. To our knowledge, this is the first French study investigate at fine spatial scale the birth certificates which list all birth in Paris during our study period.
- *Modeled air pollution data*: the air pollutant modeling procedure used provides unbiased estimates of exposure to ambient air pollution at census block level. This type of model was validated by Jerrett et al. who demonstrated its effectiveness and reliability [108].

However, the interpretation of our findings must also consider some weaknesses. Our approach, which uses ecological data, has several limitations. One is the absence of individual data such as maternal age, marital status and number of previous births. In addition, race/ethnic differences were not recorded in the first birth certificate because the French legislation prohibits the collection of any data based on race and ethnicity. Therefore, all statistical unit are considered equal. Data about the race/ethnicity were not available and were thus not included in our analysis. However, our study rests on a fine geographical resolution scale—census block—which has been designed by the Census bureau to be as homogeneous as possible in terms of population size and socio-economic and demographic characteristics. The level of homogeneity of the census blocks ensures the minimization of ecological bias, and the findings from this spatial analysis tend to be close to what could be observed at individual level [109,110]. Nonetheless, some degree of misclassification inevitably exists in individual characteristics and environmental exposures, and these could results in associations being biased towards the null. Another limitation is the absence of certain parental characteristics like lifestyle behaviors [92,95,111] including maternal nutritional deficits or status toxicants such as nicotine, cocaine or alcohol and access to healthcare [91,93–97,112]. Also, while socio-economic characteristics do not change rapidly over time, exposure to air pollution is highly variable and the present study considerer average NO_2 concentrations over several years, the same value being assigned to all births that occurred in the same census block, irrespective of seasonality. From our data, the crude estimate of the PTB rate by season in Paris over the study period increase during the winter (rate of preterm equal 5.12%) while in summer the rate is 4.36%. A recent meta-analysis study [113] revealed that the pooled relative risks of preterm births increase during the winter months (maximum observed in January) and the beginning of summer (maximum observed in June). The air pollution concentrations follow similar temporal trend with the highest and lowest level in winter and in summer respectively. Indeed, the main emission sources of nitrogen oxides are road traffic (56%) and residential sector (18%) [114]. During summer, NO_2 concentrations are lower, due to the slowdown of activities in the city and in particular the decrease of road traffic associated with the holiday period, and also in link with the chemistry of ozone formation. Due to the lack of data (we have not the daily concentrations of NO_2 per census block), it was not possible to explore the air pollution PTB effect by season. This limitation is a common feature of ecological studies as the one we conducted. Epidemiological approaches that allow estimation of personal exposure information provide a complementary viewpoint, with their own limitations.

Finally others characteristic such as green space could be associated with pregnancy outcome. In previous work we describe conceptual framework with 3 hypothetical pathways by which green spaces may have a beneficial effect on adverse pregnancy outcomes [55]. In addition, in recent study conducted in the same study area—city of Paris, we assessed the spatial variability of heat-wave-related mortality risk among elderly at the census block level and the most likely cluster for increased mortality risk is located in the same zone of cluster of high risk of PTB, in the East of Paris. In this study, we found that green space density had a protective effect [115]. In the future, It would be therefore interesting to collect all environmental exposures data from various sources, with negative health impacts (air, water and soil contamination, noise, etc.) or with positive effects (e.g., green space) and assess the effect of cumulative exposure on PTB risk using composite exposure index which performed in French [116].

6. Conclusions

In a public health perspective, regarding maternal and child practice, individual-level interventions predominate. However, adverse birth outcomes result from a complex combination of individual determinants and behavior of the parents (more particularly the mother during pregnancy) and the characteristics of the place where people live requiring appropriate ecological approaches.

Today, spatial approach constitutes a powerful tool to use in the context of the life course perspective of health, and more specifically in reproductive health [117]. A healthy pregnant woman is more likely to have a healthy newborn. In addition, neonates born in healthier place of residence in term of environmental exposure and, living and social conditions will tend to have better health trajectories throughout their life. The theoretical model suggests that adverse exposures (including characteristics of the place where people live) accumulate over time since the birth and will increase adverse outcomes during the adulthood period. In this context, this study is an attempt to fill the gap regarding a need for spatial approaches to support priority setting and guide policy makers in their choice of health interventions in general and on birth outcomes in particular. Our findings underscore the area with increased risk for preterm birth where local authorities should focus their resources and efforts to reduce health inequalities regarding birth outcomes. They highlight significant spatial implication of neighborhood characteristics including socioeconomic deprivation level and maternal exposure to ambient air NO_2, and their combination, which could guide policymakers in choosing and developing the most appropriate and specific community-oriented interventions. We hope that in the future this kind of approach will be more often used in public health studies especially in life courses perspectives.

Supplementary Materials: The following are available online at http://www.mdpi.com/1660-4601/15/9/1895/s1, Figure S1: Scatterplot matrix of NO_2, PM_{10} and $PM_{2.5}$, Table S1: Correlation Matrix between NO_2, PM_{10} and $PM_{2.5}$.

Author Contributions: Conceptualization, S.D. and W.K.; Methodology, S.D., W.K., N.A. and M.G.; Validation, S.D., W.K., M.C., A.D. and D.Z.N.; Formal Analysis, S.D., W.K., N.A.; Investigation, S.D. and W.K.; Writing-Original Draft Preparation, S.D., W.K., N.A., M.C., A.D. and D.Z.N.; Writing-Review & Editing, S.D. and W.K.; Supervision, S.D. and W.K.

Funding: This study was supported by Foundation de France: 201300040943.

Acknowledgments: The authors thank all scientific researchers of the Equit' Area project and the personnel of the local association in air quality of Paris Metropolitan Area.

Conflicts of Interest: The authors declare they have no actual or potential competing financial interests.

References

1. Terzidou, V.; Bennett, P.R. Preterm labour. *Curr. Opin. Obstet. Gynecol.* **2002**, *14*, 105–113. [CrossRef] [PubMed]
2. Hack, M.; Klein, N.K.; Taylor, H.G. Long-term developmental outcomes of low birth weight infants. *Future Child* **1995**, *5*, 176–196. [CrossRef] [PubMed]
3. Lucas, J.S.; Inskip, H.M.; Godfrey, K.M.; Foreman, C.T.; Warner, J.O.; Gregson, R.K.; Clough, J.B. Small size at birth and greater postnatal weight gain: relationships to diminished infant lung function. *Am. J. Respir. Crit. Care Med.* **2004**, *170*, 534–540. [CrossRef] [PubMed]
4. Barker, D.J.; Gluckman, P.D.; Godfrey, K.M.; Harding, J.E.; Owens, J.A.; Robinson, J.S. Fetal nutrition and cardiovascular disease in adult life. *Lancet* **1993**, *341*, 938–941. [CrossRef]
5. Mathews, T.J.; MacDorman, M.F. Infant mortality statistics from the 2005 period linked birth/infant death data set. *Natl. Vital Stat. Rep.* **2008**, *57*, 1–32. [PubMed]
6. Kramer, M.S.; Demissie, K.; Yang, H.; Platt, R.W.; Sauvé, R.; Liston, R. The contribution of mild and moderate preterm birth to infant mortality. Fetal and Infant Health Study Group of the Canadian Perinatal Surveillance System. *JAMA* **2000**, *284*, 843–849. [CrossRef] [PubMed]
7. Leem, J.-H.; Kaplan, B.M.; Shim, Y.K.; Pohl, H.R.; Gotway, C.A.; Bullard, S.M.; Rogers, J.F.; Smith, M.M.; Tylenda, C.A. Exposures to Air Pollutants during Pregnancy and Preterm Delivery. *Environ. Health Perspect.* **2006**, *114*, 905–910. [CrossRef] [PubMed]
8. Mathews, T.J.; MacDorman, M.F. Infant mortality statistics from the 2004 period linked birth/infant death data set. *Natl. Vital Stat. Rep.* **2007**, *55*, 1–32. [PubMed]
9. Hutchinson, E.A.; De Luca, C.R.; Doyle, L.W.; Roberts, G.; Anderson, P.J. Victorian Infant Collaborative Study Group School-age outcomes of extremely preterm or extremely low birth weight children. *Pediatrics* **2013**, *131*, e1053–e1061. [CrossRef] [PubMed]
10. Proietti, E.; Röösli, M.; Frey, U.; Latzin, P. Air pollution during pregnancy and neonatal outcome: a review. *J. Aerosol. Med. Pulm. Drug Deliv.* **2013**, *26*, 9–23. [CrossRef] [PubMed]
11. Shah, P.S.; Balkhair, T.; Knowledge Synthesis Group on Determinants of Preterm/LBW births. Air pollution and birth outcomes: A systematic review. *Environ. Int.* **2011**, *37*, 498–516. [CrossRef] [PubMed]
12. Stieb, D.M.; Chen, L.; Eshoul, M.; Judek, S. Ambient air pollution, birth weight and preterm birth: A systematic review and meta-analysis. *Environ. Res.* **2012**, *117*, 100–111. [CrossRef] [PubMed]
13. Bonzini, M.; Carugno, M.; Grillo, P.; Mensi, C.; Bertazzi, P.A.; Pesatori, A.C. Impact of ambient air pollution on birth outcomes: systematic review of the current evidences. *Med. Lav.* **2010**, *101*, 341–363. [PubMed]
14. Wu, J.; Wilhelm, M.; Chung, J.; Ritz, B. Comparing exposure assessment methods for traffic-related air pollution in an adverse pregnancy outcome study. *Environ. Res.* **2011**, *111*, 685–692. [CrossRef] [PubMed]
15. Miranda, M.L.; Edwards, S.E.; Chang, H.H.; Auten, R.L. Proximity to roadways and pregnancy outcomes. *J. Expo. Sci. Environ. Epidemiol.* **2013**, *23*, 32–38. [CrossRef] [PubMed]

16. Johnson, S.; Bobb, J.F.; Ito, K.; Savitz, D.A.; Elston, B.; Shmool, J.L.C.; Dominici, F.; Ross, Z.; Clougherty, J.E.; Matte, T. Ambient Fine Particulate Matter, Nitrogen Dioxide, and Preterm Birth in New York City. *Environ. Health Perspect.* **2016**, *124*, 1283–1290. [CrossRef] [PubMed]
17. Nieuwenhuijsen, M.J.; Dadvand, P.; Grellier, J.; Martinez, D.; Vrijheid, M. Environmental risk factors of pregnancy outcomes: A summary of recent meta-analyses of epidemiological studies. *Environ. Health* **2013**, *12*, 6. [CrossRef] [PubMed]
18. Bosetti, C.; Nieuwenhuijsen, M.J.; Gallus, S.; Cipriani, S.; La Vecchia, C.; Parazzini, F. Ambient particulate matter and preterm birth or birth weight: A review of the literature. *Arch. Toxicol.* **2010**, *84*, 447–460. [CrossRef] [PubMed]
19. Chang, H.H.; Reich, B.J.; Miranda, M.L. Time-to-event analysis of fine particle air pollution and preterm birth: Results from North Carolina, 2001-2005. *Am. J. Epidemiol.* **2012**, *175*, 91–98. [CrossRef] [PubMed]
20. Gehring, U.; Wijga, A.H.; Fischer, P.; de Jongste, J.C.; Kerkhof, M.; Koppelman, G.H.; Smit, H.A.; Brunekreef, B. Traffic-related air pollution, preterm birth and term birth weight in the PIAMA birth cohort study. *Environ. Res.* **2011**, *111*, 125–135. [CrossRef] [PubMed]
21. Hansen, C.; Neller, A.; Williams, G.; Simpson, R. Maternal exposure to low levels of ambient air pollution and preterm birth in Brisbane, Australia. *BJOG* **2006**, *113*, 935–941. [CrossRef] [PubMed]
22. Ghio, A.J.; Carraway, M.S.; Madden, M.C. Composition of air pollution particles and oxidative stress in cells, tissues, and living systems. *J. Toxicol. Environ. Health B Crit. Rev.* **2012**, *15*, 1–21. [CrossRef] [PubMed]
23. Dominici, F.; Sheppard, L.; Clyde, M. Health Effects of Air Pollution: A Statistical Review. *Int. Stat. Rev.* **2003**, *71*, 243–276. [CrossRef]
24. Šrám, R.J.; Binková, B.; Dejmek, J.; Bobak, M. Ambient Air Pollution and Pregnancy Outcomes: A Review of the Literature. *Environ. Health Perspect.* **2005**, *113*, 375–382. [CrossRef] [PubMed]
25. Strickland, M.J.; Klein, M.; Darrow, L.A.; Flanders, W.D.; Correa, A.; Marcus, M.; Tolbert, P.E. The Issue of Confounding in Epidemiological Studies of Ambient Air Pollution and Pregnancy Outcomes. *J. Epidemiol. Community Health* **2009**, *63*, 500–504. [CrossRef] [PubMed]
26. Sexton, K.; Gong, H.; Bailar, J.C.; Ford, J.G.; Gold, D.R.; Lambert, W.E.; Utell, M.J. Air pollution health risks: Do class and race matter? *Toxicol. Ind. Health* **1993**, *9*, 843–878. [CrossRef] [PubMed]
27. Yi, O.; Kim, H.; Ha, E. Does area level socioeconomic status modify the effects of PM_{10} on preterm delivery? *Environ. Res.* **2010**, *110*, 55–61. [CrossRef] [PubMed]
28. Salihu, H.M.; Ghaji, N.; Mbah, A.K.; Alio, A.P.; August, E.M.; Boubakari, I. Particulate pollutants and racial/ethnic disparity in feto-infant morbidity outcomes. *Matern. Child Health J.* **2012**, *16*, 1679–1687. [CrossRef] [PubMed]
29. Woodruff, T.J.; Parker, J.D.; Kyle, A.D.; Schoendorf, K.C. Disparities in exposure to air pollution during pregnancy. *Environ. Health Perspect.* **2003**, *111*, 942–946. [CrossRef] [PubMed]
30. Ponce, N.A.; Hoggatt, K.J.; Wilhelm, M.; Ritz, B. Preterm Birth: The Interaction of Traffic-related Air Pollution with Economic Hardship in Los Angeles Neighborhoods. *Am. J. Epidemiol.* **2005**, *162*, 140–148. [CrossRef] [PubMed]
31. Carbajal-Arroyo, L.; Miranda-Soberanis, V.; Medina-Ramón, M.; Rojas-Bracho, L.; Tzintzun, G.; Solís-Gutiérrez, P.; Méndez-Ramírez, I.; Hurtado-Díaz, M.; Schwartz, J.; Romieu, I. Effect of PM(10) and O(3) on infant mortality among residents in the Mexico City Metropolitan Area: A case-crossover analysis, 1997–2005. *J. Epidemiol. Community Health* **2011**, *65*, 715–721. [CrossRef] [PubMed]
32. OECD ENV/EPOC/WPNEP. *Rapport Environment and Distributional Issues: Analysis, Evidence and Policy Implications*; OECD Publishing: Paris, France, 2004.
33. Bloch, J.R. Using Geographical Information Systems to Explore Disparities in Preterm Birth Rates among Foreign-born and U.S.-born Black Mothers. *J. Obstet. Gynecol. Neonatal. Nurs.* **2011**, *40*, 544–554. [CrossRef] [PubMed]
34. Ogneva-Himmelberger, Y.; Dahlberg, T.; Kelly, K.; Simas, T.A.M. Using Geographic Information Science to Explore Associations between Air Pollution, Environmental Amenities, and Preterm Births. *AIMS Public Health* **2015**, *2*, 469–486. [CrossRef] [PubMed]
35. Cockings, S.; Martin, D. Zone design for environment and health studies using pre-aggregated data. *Soc. Sci. Med.* **2005**, *60*, 2729–2742. [CrossRef] [PubMed]
36. Haynes, R.; Daras, K.; Reading, R.; Jones, A. Modifiable neighbourhood units, zone design and residents' perceptions. *Health Place* **2007**, *13*, 812–825. [CrossRef] [PubMed]

37. Insaf, T.Z.; Talbot, T. Identifying areas at risk of low birth weight using spatial epidemiology: A small area surveillance study. *Prev. Med.* **2016**, *88*, 108–114. [CrossRef] [PubMed]
38. Sabel, C.E.; Wilson, J.G.; Kingham, S.; Tisch, C.; Epton, M. Spatial implications of covariate adjustment on patterns of risk: Respiratory hospital admissions in Christchurch, New Zealand. *Soc. Sci. Med.* **2007**, *65*, 43–59. [CrossRef] [PubMed]
39. Kihal-Talantikite, W.; Deguen, S.; Padilla, C.; Siebert, M.; Couchoud, C.; Vigneau, C.; Bayat, S. Spatial distribution of end-stage renal disease (ESRD) and social inequalities in mixed urban and rural areas: A study in the Bretagne administrative region of France. *Clin. Kidney J.* **2015**, *8*, 7–13. [CrossRef] [PubMed]
40. Ncube, C.N.; Enquobahrie, D.A.; Albert, S.M.; Herrick, A.L.; Burke, J.G. Association of neighborhood context with offspring risk of preterm birth and low birthweight: A systematic review and meta-analysis of population-based studies. *Soc. Sci. Med.* **2016**, *153*, 156–164. [CrossRef] [PubMed]
41. Vos, A.A.; Posthumus, A.G.; Bonsel, G.J.; Steegers, E.A.P.; Denktaş, S. Deprived neighborhoods and adverse perinatal outcome: A systematic review and meta-analysis. *Acta Obstet. Gynecol. Scand.* **2014**, *93*, 727–740. [CrossRef] [PubMed]
42. WHO. Preterm Birth. Available online: http://www.who.int/news-room/fact-sheets/detail/preterm-birth (accessed on 7 August 2018).
43. Quinn, J.-A.; Munoz, F.M.; Gonik, B.; Frau, L.; Cutland, C.; Mallett-Moore, T.; Kissou, A.; Wittke, F.; Das, M.; Nunes, T.; et al. Preterm birth: Case definition & guidelines for data collection, analysis, and presentation of immunisation safety data. *Vaccine* **2016**, *34*, 6047–6056. [CrossRef] [PubMed]
44. Vardoulakis, S.; Fisher, B.E.A.; Pericleous, K.; Gonzalez-Flesca, N. Modelling air quality in street canyons: A review. *Atmos. Environ.* **2003**, *37*, 155–182. [CrossRef]
45. Wahida, K.-T.; Padilla, C.M.; Denis, Z.-N.; Olivier, B.; Géraldine, L.N.; Philippe, Q.; Séverine, D. A Conceptual Framework for the Assessment of Cumulative Exposure to Air Pollution at a Fine Spatial Scale. *Int. J. Environ. Res. Public Health* **2016**, *13*, 319. [CrossRef] [PubMed]
46. Deguen, S.; Petit, C.; Delbarre, A.; Kihal, W.; Padilla, C.; Benmarhnia, T.; Lapostolle, A.; Chauvin, P.; Zmirou-Navier, D. Neighbourhood Characteristics and Long-Term Air Pollution Levels Modify the Association between the Short-Term Nitrogen Dioxide Concentrations and All-Cause Mortality in Paris. *PLoS ONE* **2015**, *10*, e0131463. [CrossRef] [PubMed]
47. Padilla, C.M.; Kihal-Talantikit, W.; Vieira, V.M.; Deguen, S. City-Specific Spatiotemporal Infant and Neonatal Mortality Clusters: Links with Socioeconomic and Air Pollution Spatial Patterns in France. *Int. J. Environ. Res. Public Health* **2016**, *13*, 624. [CrossRef] [PubMed]
48. Levy, I.; Mihele, C.; Lu, G.; Narayan, J.; Brook, J.R. Evaluating Multipollutant Exposure and Urban Air Quality: Pollutant Interrelationships, Neighborhood Variability, and Nitrogen Dioxide as a Proxy Pollutant. *Environ. Health Perspect.* **2014**, *122*, 65–72. [CrossRef] [PubMed]
49. Ritz, B.; Wilhelm, M. Ambient Air Pollution and Adverse Birth Outcomes: Methodologic Issues in an Emerging Field. *Basic Clin. Pharmacol. Toxicol.* **2008**, *102*, 182–190. [CrossRef] [PubMed]
50. Brauer, M.; Lencar, C.; Tamburic, L.; Koehoorn, M.; Demers, P.; Karr, C. A Cohort Study of Traffic-Related Air Pollution Impacts on Birth Outcomes. *Environ. Health Perspect.* **2008**, *116*, 680–686. [CrossRef] [PubMed]
51. Darrow, L.A.; Klein, M.; Flanders, W.D.; Waller, L.A.; Correa, A.; Marcus, M.; Mulholland, J.A.; Russell, A.G.; Tolbert, P.E. Ambient air pollution and preterm birth: A time-series analysis. *Epidemiology* **2009**, *20*, 689–698. [CrossRef] [PubMed]
52. Lalloué, B.; Monnez, J.-M.; Padilla, C.; Kihal, W.; Le Meur, N.; Zmirou-Navier, D.; Deguen, S. A statistical procedure to create a neighborhood socioeconomic index for health inequalities analysis. *Int. J. Equity Health* **2013**, *12*, 21. [CrossRef] [PubMed]
53. Kihal-Talantikite, W.; Padilla, C.M.; Lalloué, B.; Gelormini, M.; Zmirou-Navier, D.; Deguen, S. Green space, social inequalities and neonatal mortality in France. *BMC Pregnancy Childbirth* **2013**, *13*, 191. [CrossRef] [PubMed]
54. Kihal-Talantikite, W.; Padilla, C.M.; Lalloue, B.; Rougier, C.; Defrance, J.; Zmirou-Navier, D.; Deguen, S. An exploratory spatial analysis to assess the relationship between deprivation, noise and infant mortality: An ecological study. *Environ. Health* **2013**, *12*, 109. [CrossRef] [PubMed]
55. Padilla, C.M.; Deguen, S.; Lalloue, B.; Blanchard, O.; Beaugard, C.; Troude, F.; Navier, D.Z.; Vieira, V.M. Cluster analysis of social and environment inequalities of infant mortality. A spatial study in small areas revealed by local disease mapping in France. *Sci. Total Environ.* **2013**, *454–455*, 433–441. [CrossRef] [PubMed]

56. Kulldorff, M. *SaTScan: Software for the Spatial, Temporal, and Space-Time Scan Statistics*, version 6.0; Information Management Services, Inc.: Rockville, MD, USA, 2005.
57. Kulldorff, M.; Feuer, E.J.; Miller, B.A.; Freedman, L.S. Breast cancer clusters in the northeast United States: A geographic analysis. *Am. J. Epidemiol.* **1997**, *146*, 161–170. [CrossRef] [PubMed]
58. Kulldorff, M. A spatial scan statistic. *Commun. Stat. Theory Methods* **1997**, *26*, 1481–1496. [CrossRef]
59. Dwass, M. Modified Randomization Tests for Nonparametric Hypotheses. *Ann. Math. Stat.* **1957**, *28*, 181–187. [CrossRef]
60. Padula, A.M.; Mortimer, K.M.; Tager, I.B.; Hammond, S.K.; Lurmann, F.W.; Yang, W.; Stevenson, D.K.; Shaw, G.M. Traffic-related air pollution and risk of preterm birth in the San Joaquin Valley of California. *Ann. Epidemiol.* **2014**, *24*, 888–895.e4. [CrossRef] [PubMed]
61. Dadvand, P.; Basagaña, X.; Figueras, F.; Martinez, D.; Beelen, R.; Cirach, M.; de Nazelle, A.; Hoek, G.; Ostro, B.; Nieuwenhuijsen, M.J. Air pollution and preterm premature rupture of membranes: A spatiotemporal analysis. *Am. J. Epidemiol.* **2014**, *179*, 200–207. [CrossRef] [PubMed]
62. Wilhelm, M.; Ghosh, J.K.; Su, J.; Cockburn, M.; Jerrett, M.; Ritz, B. Traffic-related air toxics and preterm birth: A population-based case-control study in Los Angeles County, California. *Environ. Health* **2011**, *10*, 89. [CrossRef] [PubMed]
63. Zhao, Q.; Liang, Z.; Tao, S.; Zhu, J.; Du, Y. Effects of air pollution on neonatal prematurity in Guangzhou of China: A time-series study. *Environ. Health* **2011**, *10*, 2. [CrossRef] [PubMed]
64. Hannam, K.; McNamee, R.; Baker, P.; Sibley, C.; Agius, R. Air pollution exposure and adverse pregnancy outcomes in a large UK birth cohort: Use of a novel spatio-temporal modelling technique. *Scand. J. Work Environ. Health* **2014**, *40*, 518–530. [CrossRef] [PubMed]
65. Gehring, U.; Tamburic, L.; Sbihi, H.; Davies, H.W.; Brauer, M. Impact of noise and air pollution on pregnancy outcomes. *Epidemiology* **2014**, *25*, 351–358. [CrossRef] [PubMed]
66. Schifano, P.; Lallo, A.; Asta, F.; De Sario, M.; Davoli, M.; Michelozzi, P. Effect of ambient temperature and air pollutants on the risk of preterm birth, Rome 2001–2010. *Environ. Int.* **2013**, *61*, 77–87. [CrossRef] [PubMed]
67. Le, H.Q.; Batterman, S.A.; Wirth, J.J.; Wahl, R.L.; Hoggatt, K.J.; Sadeghnejad, A.; Hultin, M.L.; Depa, M. Air pollutant exposure and preterm and term small-for-gestational-age births in Detroit, Michigan: Long-term trends and associations. *Environ. Int.* **2012**, *44*, 7–17. [CrossRef] [PubMed]
68. Van den Hooven, E.H.; Pierik, F.H.; de Kluizenaar, Y.; Willemsen, S.P.; Hofman, A.; van Ratingen, S.W.; Zandveld, P.Y.J.; Mackenbach, J.P.; Steegers, E.A.P.; Miedema, H.M.E.; et al. Air pollution exposure during pregnancy, ultrasound measures of fetal growth, and adverse birth outcomes: A prospective cohort study. *Environ. Health Perspect.* **2012**, *120*, 150–156. [CrossRef] [PubMed]
69. Estarlich, M.; Ballester, F.; Davdand, P.; Llop, S.; Esplugues, A.; Fernández-Somoano, A.; Lertxundi, A.; Guxens, M.; Basterrechea, M.; Tardón, A.; et al. Exposure to ambient air pollution during pregnancy and preterm birth: A Spanish multicenter birth cohort study. *Environ. Res.* **2016**, *147*, 50–58. [CrossRef] [PubMed]
70. Généreux, M.; Auger, N.; Goneau, M.; Daniel, M. Neighbourhood socioeconomic status, maternal education and adverse birth outcomes among mothers living near highways. *J. Epidemiol. Community Health* **2008**, *62*, 695–700. [CrossRef] [PubMed]
71. Yorifuji, T.; Naruse, H.; Kashima, S.; Ohki, S.; Murakoshi, T.; Takao, S.; Tsuda, T.; Doi, H. Residential proximity to major roads and preterm births. *Epidemiology* **2011**, *22*, 74–80. [CrossRef] [PubMed]
72. Yorifuji, T.; Naruse, H.; Kashima, S.; Takao, S.; Murakoshi, T.; Doi, H.; Kawachi, I. Residential proximity to major roads and adverse birth outcomes: A hospital-based study. *Environ. Health* **2013**, *12*, 34. [CrossRef] [PubMed]
73. Yorifuji, T.; Naruse, H.; Kashima, S.; Murakoshi, T.; Doi, H. Residential proximity to major roads and obstetrical complications. *Sci. Total Environ.* **2015**, *508*, 188–192. [CrossRef] [PubMed]
74. Yang, C.-Y.; Chang, C.-C.; Chuang, H.-Y.; Ho, C.-K.; Wu, T.-N.; Tsai, S.-S. Evidence for increased risks of preterm delivery in a population residing near a freeway in Taiwan. *Arch. Environ. Health* **2003**, *58*, 649–654. [CrossRef] [PubMed]
75. Wilhelm, M.; Ritz, B. Residential proximity to traffic and adverse birth outcomes in Los Angeles county, California, 1994–1996. *Environ. Health Perspect.* **2003**, *111*, 207–216. [CrossRef] [PubMed]
76. Vadillo-Ortega, F.; Osornio-Vargas, A.; Buxton, M.A.; Sánchez, B.N.; Rojas-Bracho, L.; Viveros-Alcaráz, M.; Castillo-Castrejón, M.; Beltrán-Montoya, J.; Brown, D.G.; O'Neill, M.S. Air pollution, inflammation and preterm birth: A potential mechanistic link. *Med. Hypotheses* **2014**, *82*, 219–224. [CrossRef] [PubMed]

77. Sapkota, A.; Chelikowsky, A.P.; Nachman, K.E.; Cohen, A.J.; Ritz, B. Exposure to particulate matter and adverse birth outcomes: A comprehensive review and meta-analysis. *Air Qual. Atmos. Health* **2012**, *5*, 369–381. [CrossRef]
78. Longini, M.; Perrone, S.; Vezzosi, P.; Marzocchi, B.; Kenanidis, A.; Centini, G.; Rosignoli, L.; Buonocore, G. Association between oxidative stress in pregnancy and preterm premature rupture of membranes. *Clin. Biochem.* **2007**, *40*, 793–797. [CrossRef] [PubMed]
79. Luo, Z.-C.; Wilkins, R.; Kramer, M.S.; Fetal and Infant Health Study Group of the Canadian Perinatal Surveillance System. Effect of neighbourhood income and maternal education on birth outcomes: A population-based study. *CMAJ* **2006**, *174*, 1415–1420. [CrossRef] [PubMed]
80. Luo, Z.-C.; Kierans, W.J.; Wilkins, R.; Liston, R.M.; Mohamed, J.; Kramer, M.S.; British Columbia Vital Statistics Agency. Disparities in birth outcomes by neighborhood income: Temporal trends in rural and urban areas, British Columbia. *Epidemiology* **2004**, *15*, 679–686. [CrossRef] [PubMed]
81. Agyemang, C.; Vrijkotte, T.G.M.; Droomers, M.; van der Wal, M.F.; Bonsel, G.J.; Stronks, K. The effect of neighbourhood income and deprivation on pregnancy outcomes in Amsterdam, The Netherlands. *J. Epidemiol. Community Health* **2009**, *63*, 755–760. [CrossRef] [PubMed]
82. Liu, N.; Wen, S.W.; Katherine, W.; Bottomley, J.; Yang, Q.; Walker, M.C. Neighbourhood family income and adverse birth outcomes among singleton deliveries. *J. Obstet. Gynaecol. Can.* **2010**, *32*, 1042–1048. [CrossRef]
83. Urquia, M.L.; Frank, J.W.; Glazier, R.H.; Moineddin, R. Birth outcomes by neighbourhood income and recent immigration in Toronto. *Health Rep.* **2007**, *18*, 21–30. [PubMed]
84. Garcia-Subirats, I.; Pérez, G.; Rodríguez-Sanz, M.; Ruiz-Muñoz, D.; Muñoz, D.R.; Salvador, J. Neighborhood inequalities in adverse pregnancy outcomes in an urban setting in Spain: A multilevel approach. *J. Urban Health* **2012**, *89*, 447–463. [CrossRef] [PubMed]
85. Janghorbani, M.; Stenhouse, E.; Millward, A.; Jones, R.B. Neighborhood deprivation and preterm birth in Plymouth, UK. *J. Matern. Fetal. Neonatal. Med.* **2006**, *19*, 85–91. [CrossRef] [PubMed]
86. Gray, R.; Bonellie, S.R.; Chalmers, J.; Greer, I.; Jarvis, S.; Williams, C. Social inequalities in preterm birth in Scotland 1980–2003: Findings from an area-based measure of deprivation. *BJOG* **2008**, *115*, 82–90. [CrossRef] [PubMed]
87. Smith, L.K.; Draper, E.S.; Manktelow, B.N.; Field, D.J. Socioeconomic inequalities in survival and provision of neonatal care: Population based study of very preterm infants. *BMJ* **2009**, *339*, b4702. [CrossRef] [PubMed]
88. Janevic, T.; Stein, C.R.; Savitz, D.A.; Kaufman, J.S.; Mason, S.M.; Herring, A.H. Neighborhood deprivation and adverse birth outcomes among diverse ethnic groups. *Ann. Epidemiol.* **2010**, *20*, 445–451. [CrossRef] [PubMed]
89. Taylor-Robinson, D.; Agarwal, U.; Diggle, P.J.; Platt, M.J.; Yoxall, B.; Alfirevic, Z. Quantifying the impact of deprivation on preterm births: A retrospective cohort study. *PLoS ONE* **2011**, *6*, e23163. [CrossRef] [PubMed]
90. Buka, S.L.; Brennan, R.T.; Rich-Edwards, J.W.; Raudenbush, S.W.; Earls, F. Neighborhood support and the birth weight of urban infants. *Am. J. Epidemiol.* **2003**, *157*, 1–8. [CrossRef] [PubMed]
91. Schempf, A.; Strobino, D.; O'Campo, P. Neighborhood Effects on Birthweight: An Exploration of Psychosocial and Behavioral Pathways in Baltimore, 1995–1996. *Soc. Sci. Med.* **2009**, *68*, 100–110. [CrossRef] [PubMed]
92. Nkansah-Amankra, S.; Luchok, K.J.; Hussey, J.R.; Watkins, K.; Liu, X. Effects of maternal stress on low birth weight and preterm birth outcomes across neighborhoods of South Carolina, 2000–2003. *Matern. Child Health J.* **2010**, *14*, 215–226. [CrossRef] [PubMed]
93. Mitchell, E.A.; Robinson, E.; Clark, P.M.; Becroft, D.M.O.; Glavish, N.; Pattison, N.S.; Pryor, J.E.; Thompson, J.M.D.; Wild, C.J. Maternal nutritional risk factors for small for gestational age babies in a developed country: A case-control study. *Arch. Dis. Child. Fetal Neonatal Ed.* **2004**, *89*, F431–F435. [CrossRef] [PubMed]
94. Kozuki, N.; Lee, A.C.C.; Black, R.E.; Katz, J. Nutritional and Reproductive Risk Factors for Small for Gestational Age and Preterm Births. *Nestle Nutr. Inst. Works. Ser.* **2015**, *81*, 17–28. [CrossRef]
95. Patra, J.; Bakker, R.; Irving, H.; Jaddoe, V.W.V.; Malini, S.; Rehm, J. Dose-response relationship between alcohol consumption before and during pregnancy and the risks of low birthweight, preterm birth and small for gestational age (SGA)—A systematic review and meta-analyses. *BJOG* **2011**, *118*, 1411–1421. [CrossRef] [PubMed]
96. Northstone, K.; Emmett, P.; Rogers, I. Dietary patterns in pregnancy and associations with socio-demographic and lifestyle factors. *Eur. J. Clin. Nutr.* **2008**, *62*, 471–479. [CrossRef] [PubMed]

97. O'Leary, C.M.; Nassar, N.; Kurinczuk, J.J.; Bower, C. The effect of maternal alcohol consumption on fetal growth and preterm birth. *BJOG* **2009**, *116*, 390–400. [CrossRef] [PubMed]
98. Hsieh, V.C.-R.; Shieh, S.-H.; Chen, C.-Y.; Liou, S.-H.; Hsiao, Y.-C.; Wu, T.-N. Does Social Health Insurance Close the Gap: The Case of Socioeconomic Status and Preterm Low-Birth-Weight Survival. *Asia Pac. J. Public Health* **2015**, *27*, 497–508. [CrossRef] [PubMed]
99. Timmermans, S.; Bonsel, G.J.; Steegers-Theunissen, R.P.M.; Mackenbach, J.P.; Steyerberg, E.W.; Raat, H.; Verbrugh, H.A.; Tiemeier, H.W.; Hofman, A.; Birnie, E.; et al. Individual accumulation of heterogeneous risks explains perinatal inequalities within deprived neighbourhoods. *Eur. J. Epidemiol.* **2011**, *26*, 165–180. [CrossRef] [PubMed]
100. Hodgson, E.J.; Lockwood, C.J. Preterm Birth: A Complex Disease. In *Preterm Birth*; Wiley-Blackwell: Oxford, UK, 2010; pp. 8–16.
101. Hao, H.; Chang, H.H.; Holmes, H.A.; Mulholland, J.A.; Klein, M.; Darrow, L.A.; Strickland, M.J. Air Pollution and Preterm Birth in the U.S. State of Georgia (2002–2006): Associations with Concentrations of 11 Ambient Air Pollutants Estimated by Combining Community Multiscale Air Quality Model (CMAQ) Simulations with Stationary Monitor Measurements. *Environ. Health Perspect.* **2016**, *124*, 875–880. [CrossRef] [PubMed]
102. Cakmak, S.; Hebbern, C.; Cakmak, J.D.; Vanos, J. The modifying effect of socioeconomic status on the relationship between traffic, air pollution and respiratory health in elementary schoolchildren. *J. Environ. Manag.* **2016**, *177*, 1–8. [CrossRef] [PubMed]
103. Clark, L.P.; Millet, D.B.; Marshall, J.D. National Patterns in Environmental Injustice and Inequality: Outdoor NO_2 Air Pollution in the United States. *PLoS ONE* **2014**, *9*, e94431. [CrossRef] [PubMed]
104. Tian, N.; Xue, J.; Barzyk, T.M. Evaluating socioeconomic and racial differences in traffic-related metrics in the United States using a GIS approach. *J. Expo. Sci. Environ. Epidemiol.* **2013**, *23*, 215–222. [CrossRef] [PubMed]
105. Clougherty, J.E.; Kubzansky, L.D. A Framework for Examining Social Stress and Susceptibility to Air Pollution in Respiratory Health. *Environ. Health Perspect.* **2009**, *117*, 1351–1358. [CrossRef] [PubMed]
106. Bandoli, G.; von Ehrenstein, O.; Ghosh, J.K.; Ritz, B. Synergistic effects of air pollution and psychosocial stressors on adolescent lung function. *J. Allergy Clin. Immunol.* **2016**, *138*, 918–920.e4. [CrossRef] [PubMed]
107. Premiers Certificats de Santé des Enfants nés en 2010 et Domiciliés en Île-de-France—Observatoire Régional de Santé Île-de-France—15 rue Falguière 75015 Paris—Tel.: 01 77 49 78 60—ors-idf@ors-idf.org. Available online: http://www.erpurs.org/index.php/fr/publications/56-groupes-de-population/meres-et-enfants/87-premiers-certificats-de-sante-des-enfants-nes-en-2010-et-domicilies-en-ile-de-france (accessed on 17 July 2018).
108. Jerrett, M.; Gale, S.; Kontgis, C. Spatial modeling in environmental and public health research. *Int. J. Environ. Res. Public Health* **2010**, *7*, 1302–1329. [CrossRef] [PubMed]
109. Greenland, S. Ecologic versus individual-level sources of bias in ecologic estimates of contextual health effects. *Int. J. Epidemiol.* **2001**, *30*, 1343–1350. [CrossRef] [PubMed]
110. Krieger, N.; Williams, D.R.; Moss, N.E. Measuring social class in US public health research: Concepts, methodologies, and guidelines. *Annu. Rev. Public Health* **1997**, *18*, 341–378. [CrossRef] [PubMed]
111. Miranda, M.L.; Maxson, P.; Edwards, S. Environmental contributions to disparities in pregnancy outcomes. *Epidemiol. Rev.* **2009**, *31*, 67–83. [CrossRef] [PubMed]
112. De Graaf, J.P.; Steegers, E.A.P.; Bonsel, G.J. Inequalities in perinatal and maternal health. *Curr. Opin. Obstet. Gynecol.* **2013**, *25*, 98–108. [CrossRef] [PubMed]
113. Beltran, A.J.; Wu, J.; Laurent, O. Associations of meteorology with adverse pregnancy outcomes: A systematic review of preeclampsia, preterm birth and birth weight. *Int. J. Environ. Res. Public Health* **2013**, *11*, 91–172. [CrossRef] [PubMed]
114. Airparif. Inventaire Régional des Emissions en Ile-de-France—Année de Référence 2012. Available online: https://www.airparif.asso.fr/_pdf/publications/inventaire-emissions-idf-2012-150121.pdf,%22 (accessed on 9 August 2018).
115. Benmarhnia, T.; Kihal-Talantikite, W.; Ragettli, M.S.; Deguen, S. Small-area spatiotemporal analysis of heatwave impacts on elderly mortality in Paris: A cluster analysis approach. *Sci. Total Environ.* **2017**, *592*, 288–294. [CrossRef] [PubMed]

116. Lalloué, B.; Monnez, J.-M.; Padilla, C.; Kihal, W.; Zmirou-Navier, D.; Deguen, S. Data analysis techniques: A tool for cumulative exposure assessment. *J. Expo. Sci. Environ. Epidemiol.* **2015**, *25*, 222–230. [CrossRef] [PubMed]
117. Burton-Jeangros, C.; Cullati, S.; Sacker, A.; Blane, D. (Eds.) *A Life Course Perspective on Health Trajectories and Transitions*; Life Course Research and Social Policies; Springer International Publishing: Cham, Switzherland, 2015; ISBN 978-3-319-20483-3.

International Journal of
Environmental Research and Public Health

MDPI

Article

Developing an Awareness Campaign to Reduce Second Hand Smoke Among Disadvantaged Families—A Participatory M-Health Approach

Tobias Weinmann [1,2,3,*,†], Katja Radon [1,2,3,†], Freya Sukalla [2,4,5], Jessica Gerlich [1,2,3], Swaantje Barth [1,2,3], Dennis Nowak [1,2,3] and Veronika Karnowski [2,4]

1 Institute and Clinic for Occupational, Social and Environmental Medicine, University Hospital, LMU Munich, Ziemssenstr. 1, 80336 Munich, Germany; katja.radon@med.lmu.de (K.R.); jessica.gerlich@med.lmu.de (J.G.); swaantje.barth@gmail.com (S.B.); dennis.nowak@med.lmu.de (D.N.)
2 Munich Center of Health Sciences (MC-Health), LMU Munich, Ludwigstr. 28, 80539 Munich, Germany; freya.sukalla@uni-leipzig.de (F.S.); veronika.karnowski@ifkw.lmu.de (V.K.)
3 Comprehensive Pneumology Center Munich (CPC-M), Deutsches Zentrum für Lungenforschung (DZL), Max-Lebsche-Platz 31, 81377 Munich, Germany
4 Department of Communication Studies and Media Research, LMU Munich, Oettingenstr. 67, 80538 Munich, Germany
5 Institute of Communication and Media Studies, University of Leipzig, Burgstr. 21, 04109 Leipzig, Germany
* Correspondence: tobias.weinmann@med.lmu.de
† These authors contributed equally to this work.

Received: 14 August 2018; Accepted: 5 September 2018; Published: 6 September 2018

Abstract: Children from disadvantaged families are particularly exposed to second hand smoke in their home environment. Using a mixed methods participatory approach, we aimed at identifying suitable media channels and appropriate content for a campaign increasing caregivers' knowledge about the risks of second hand smoke (SHS) exposure for their children and appropriate measures for exposure reduction. In the first phase of the mixed method design, we evaluated knowledge and norms about children's SHS and perceived barriers for avoiding it. To this end, we conducted 26 one-to-one interviews with smoking caregivers of children below the age of six years. Subsequently, we developed and illustrated core messages and identified appropriate communication channels. These were evaluated in focus group discussions by 20 of the 26 interview participants. After a final revision, 121 caregivers evaluated the campaign via an online questionnaire. Online social networks were identified as the most suitable media channel. For these, we developed animated photos with voiceovers addressing the potential consequences of SHS for children. The overall rating of the campaign messages was promising. Participants confirmed that it was important to address the issue in social media. However, sharing the pictures was considered unlikely due to the sensitive topic of the campaign, while the importance of doctors or scientists being recognizable as a source was highlighted. Employing a participatory approach, we developed an m-health campaign, which can now be disseminated in social networks to reach the target population. The effectiveness of the campaign should be evaluated.

Keywords: participatory research; second hand smoking; communication media; vulnerable populations; migrants

1. Introduction

According to the estimates of the World Health Organization (WHO), each year 150,000 children and adolescents die as a consequence of exposure to second hand smoke (SHS) [1]. Scientifically proven is the relationship between SHS exposure and premature births, lower birth weight and irreversible organ damage. The latter include respiratory diseases such as asthma, sudden infant death syndrome and otitis media [2].

In Germany, about 20% of children are exposed to SHS inside the home environment [3] compared to 11% in the US [4] and 79% in Indonesia as examples for countries with relatively small/high exposure levels [5]. While the introduction of smoke-free laws in Germany between 2007 and 2008 led to a considerable reduction of SHS exposure in public places [6], there is no legislation with respect to smoking at home or in cars [7]. The highest levels of SHS exposure are generally found among children of disadvantaged and migrant families [3,8–12]. One explanation for this could be that individuals with lower socio-economic status (SES) are less aware of the detrimental health effects of SHS and have less knowledge regarding strategies to reduce SHS exposure [2,3,13]. Therefore, it has been recommended that interventions should be tailored to the needs of families in difficult social or economic situations and should take into account cultural barriers [3,9,14–17].

The optimal intervention would be to convince parents to stop smoking. However, this approach is frequently not successful [18,19]. Unfortunately, the second best approach—promoting a reduction of SHS exposure in the home environment—has also resulted in only small effects in most studies [19,20]. Practical, social, financial, cultural, and personal issues make it difficult for disadvantaged parents to protect their children from SHS [15,16,21]. To our knowledge, the intervention studies carried out so far followed a top-down rather than a participatory approach [18,19]. However, the latter approach is recommended in order to reach disadvantaged parts of the population [22,23]. For this, a mixed methods iterative design assessing the target populations' perspectives via qualitative studies followed by the development of intervention strategies, which are then evaluated using quantitative methods, is desirable [24].

Furthermore, most of the intervention studies targeted at the reduction of SHS exposure in children were carried out in health settings, while the use of media as an access path was rarely considered [11]. A relatively new option of employing media for interventions is the distribution of health-related information via electronic media (so called "e-health"), especially mobile communication devices and social online networks ("m-health"). One of their main advantages is that they potentially reach a large number of users at low costs as compared to interpersonal counselling. More importantly, they have the potential to reach disadvantaged groups which are hard to approach with conventional methods [25]. A Canadian study, for example, successfully employed a smartphone application for coaching diabetes patients with a rather low socio-economic status [26]. For the effective planning of such m-health interventions, the involvement of the target group from the first development step is crucial [27].

Using a largely participatory approach, we therefore aimed at identifying media channels and appropriate content for an m-health campaign increasing knowledge about the risks of second hand smoke exposure and appropriate measures for exposure reduction among disadvantaged caregivers in Germany.

2. Materials and Methods

The study was conducted using a mixed method iterative design involving the target group in various stages of the project. By doing this, the results of the first two research phases were enriched in a third step by a quantitative online survey. The study was approved by the ethical commission of the faculty for social sciences of the Ludwig-Maximilian University of Munich (GZ 15-01). Each participant provided informed consent prior to participation. Participants were offered shopping vouchers to remunerate them for their time (phase 1 and 2). In phase 3, to increase motivation, respondents could participate in a lottery containing five 100-Euro vouchers.

2.1. Phase 1: Semi-Structured Interviews

In the first qualitative phase of the study, we conducted 26 semi-structured one-to-one interviews lasting between 30 and 60 min exploring existing knowledge and norms of smoking around children as well as perceived barriers to avoiding it [28]. Furthermore, participants were asked to suggest potentially successful key messages and access paths to reach the target group. Given the sensitivity of the topic, one-to-one interviews seemed to be the most appropriate method; providing an empathetic and supportive environment [10]. Male and female adult caregivers of children below the age of six years, who were either smokers themselves or lived in a smokers' household, were eligible. They were also either German, Turkish, Russian or Spanish speakers being unemployed or holding a job with a low job prestige [29]. Recruitment was done at five social institutions offering vocational training, language, integrative or health promotion courses for individuals with low SES, one paediatric practice, the pulmonary service of a children's hospital, and at shopping malls and playgrounds in socially disadvantaged parts of Munich, Germany. Recruitment and interview staff were fluent in German and had either Turkish, Russian or Spanish as their mother tongue so that all interviews could be performed in the participants' preferred language.

2.2. Phase 2: Focus Group Discussions

Based on the results of phase 1, we developed eight key messages, which could be distributed via online social networks. In all messages, a child addressed one health or social consequence of SHS and then offered one potential measure to address the problem. These messages were then visualised by both a comic-style illustrator and a 3D artist. Thereafter, 20 of the 26 participants of the first project phase took part in focus group discussions, of four to six persons, which lasted about 60 min. During the discussions, participants had the chance to give feedback on the design and content of the messages. Based on the focus group discussions, the most promising messages were selected and revised including, among others, the addition of a voiceover.

The interviews, as well as the focus group discussions, were recorded using two audio recorders and transcribed literally, with Turkish, Russian or Spanish transcripts being translated into German. After transcription and translation, we inductively analysed the materials by a stepwise formation of categories [30]. At the same time, analyses were guided by deductive categories that were based on established health behaviour models including evaluation of the perceived threat of SHS for children [31], social norms related to SHS [32] and perceived barriers to avoid SHS [33].

2.3. Phase 3: Quantitative Assessment

Using an online survey (SocSciSurvey; https://www.soscisurvey.de/), key messages and communication channels were evaluated with respect to the fit to the target group, acceptance and general impression. The campaign was offered without audio to ensure that all participants evaluated it in the same way of presentation.

The study population was recruited via snowball sampling. For this, participants in the first two project phases were asked to invite five friends via online social networks to answer the online survey. The number of participants was increased via the social network of the study team and face-to-face recruitment at local playgrounds and in a paediatric practice. Participants could thus answer the questionnaire at home using the link they were given or, in the case of face-to-face recruitment, by direct data entry in a laptop provided by the study team.

The online survey included items on socio-demographics (age, sex, country of birth and educational status) and smoking behaviour. Using five-point Likert scales, participants were asked to assess each of the illustrations and to give an overall evaluation of the campaign. The following aspects were evaluated:

- Attitude towards SHS (4 items)
- Evaluation of each of the illustrations:
    - First impression (1 item)
    - Appeal (1 item)
    - Quality (5 items)

    - ○ Intention to share in social media (4 items)
- Overall evaluation of the campaign:
    - ○ Suitability for social media (3 items)
    - ○ Content (3 items)
    - ○ Effectiveness (4 items)

The exact wording of all items of the online survey is provided in the online Supplement S1.

2.4. Statistical Analyses

For each of the items evaluated on a Likert scale, the relative frequency of each of the five response categories as well as the mean and standard deviation (SD) were calculated. For the scale "intention to share in social media", a total score of the mean of the single item was calculated as Cronbach's alpha and indicated the internal consistency of this scale (alpha > 0.7). In order to evaluate the independence of the evaluation of the individual campaign illustrations, a non-parametric analysis of variance (ANOVA) was carried out. A Dunn-Bonferroni post hoc test was used if pANOVA was <0.05. Data were analyzed using SPSS (IBM SPSS 24.0, SPSS Inc., Chicago, IL, USA). Of the 121 participants, 35 subjects did not complete all items. We did not see a systematic tendency that specific items were not answered, the only observable tendency was that there were more missing data in the items at the end of the questionnaire such as socio-demographic information. Therefore, these item-non-responders were only excluded from analyses for their respective missing items.

3. Results

3.1. Phase 1: Semi-Structured Interviews

A total of 26 people were interviewed, nine men and 17 women, between the ages of 26 and 49. Twenty-two interviewees identified themselves as smokers, while four were non-smokers or ex-smokers but lived together with a smoking partner. Eighteen of them were born in Germany, the others had lived in Germany for an average of 15 years.

In the mixed inductive-deductive analysis of the interviews [34], a lack of knowledge was identified with respect to the definition of SHS exposure and its potential health consequences, especially less severe conditions such as otitis media or tooth decay. For a potential m-health campaign, participants asked for concrete information about health consequences and on prevention measures to effectively reduce SHS exposure, especially in difficult contexts (e.g., the absence of a balcony or garden so that smoking outside would mean leaving the children alone at home). The participants also suggested that messages should be created from the perspective of the children as victims, not from the perspective of the parents.

At the formal level, messages should be (audio)visual, simple, concise, and with a positive appeal. The participants highlighted that the campaign should not only target women but also men as they are less frequently addressed by typical information providers such as gynaecologists or paediatricians. Online social networks, especially Facebook, were classified as the most suitable access paths to the target population and thus the best way to distribute the information.

3.2. Phase 2: Focus Group Discussions

Twenty of the 26 individuals from phase 1 participated in the focus groups. Regarding the design of the illustrations, the participants were in favour of the photo campaign developed by the 3D artist, especially because of the adoption of the children's perspective. However, they suggested to present the key messages more concisely and to have a uniform structure for all illustrations. In addition, illustrations were considered to be even more interesting if they were animated and included a voiceover by a child.

With respect to the content of the illustrations, the focus group participants selected four illustrations with key messages as the most suitable ones. Based on these suggestions, texts were optimised, pictures animated and a voiceover for the first part of the text added (Figure 1). The final illustrations addressed two health effects of SHS exposure in children (otitis media, asthma) and two social effects (children of smokers start smoking at a younger age, smell of SHS in clothes). Each of them were combined with up to two of the following simple and effective measures to reduce SHS exposure: Not smoking inside the house/apartment, not smoking in front of children and not smoking in cars. Finally, participants considered it important that the campaign would be distributed in online social networks by a trustworthy source, e.g., physicians.

Illustration 1: Asthma

MOM, WHEN DOES MY ASTHMA GO AWAY?

Children of smokers suffer twice as often from asthma!

ONLY SMOKE OUTDOORS AND NEVER IN THE CAR!

Illustration 3: Otitis media

DAD, WHY DO I GET MIDDLE EAR INFECTION SO OFTEN?

Children of smokers have a 40% increased risk of middle ear infections!

ONLY SMOKE OUTSIDE!

Illustration 2: Cool

I WANT TO BECOME AS COOL AS MY MOM!

Children of smokers start smoking earlier!

ONLY SMOKE OUTDOORS AND NEVER IN FRONT OF THE CHILDREN!

Illustration 4: Bad smell

EVERYONE IN SCHOOL IS SO MEAN TO ME BECAUSE MY CLOTHES SMELL SO BAD.

Cigarette smoke settles in clothing and remains long after the cigarette has gone out.

ONLY SMOKE OUTDOORS!

Figure 1. Pictures with key messages of the final campaign (Original in German with English translation provided below). MP4 files with the voiceover can be found at http://www.klinikum.uni-muenchen.de/Institut-und-Poliklinik-fuer-Arbeits-Sozial-und-Umweltmedizin/de/forschung/arbeitsgruppen/Prof__Radon/aktuelles/Passivrauchkampagne.

3.3. Phase 3: Quantitative Assessment

Of the 121 participants in the online survey, 45 were women, 41 were men and the remaining 35 did not indicate their sex. The mean age was 35 years (range 20–56 years). All participants were smokers with at least one child below the age of six in the household. About half of the study population did not complete high school and 27% were migrants (either themselves or their parents were not born in Germany).

With respect to attitude towards SHS, participants were mostly aware that SHS has adverse health effects. However, many did not support a restriction on smoking in cars. More than half of the participants agreed that a campaign about SHS and its effects on children would be useful (Figure 2).

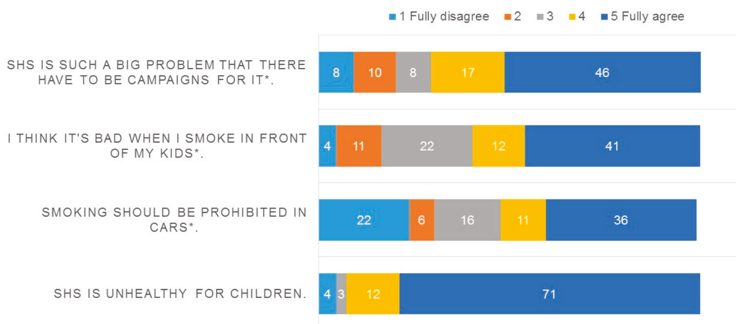


Figure 2. Attitude towards second hand smoke of the 121 participants in the quantitative evaluation of the campaign (* Original item formulated negatively; SHS = Second Hand Smoke).

3.4. Evaluation of the Individual Campaign Illustrations

Concerning the first impression of the illustrations, they were evaluated as very poor by one-third of the participants, while one-third found them (very) appealing without statistically significant differences between the single illustrations (Table 1). The overall quality of the illustrations and their key messages were positively evaluated with small but sometimes statistically significant differences between them (Table 2). The intention to share the pictures was moderate (mean value between 3.14 and 3.21 for the individual illustrations on a Likert scale from 1 = very likely to 5 = very unlikely).

Table 1. Evaluation of the first impression and the appeal of the campaign illustrations by the 121 participants in the quantitative evaluation.

	Illustration 1: Asthma (%)	**Illustration 2: Cool Like My Mom (%)**	**Illustration 3: Otitis Media (%)**	**Illustration 4: Bad Smell (%)**	
		First impression			pANOVA
1 very poor	44.6	31.4	33.9	26.4	
2	3.3	9.9	1.7	8.3	
3	8.3	11.6	11.6	10.7	
4	9.1	7.4	9.1	12.4	
5 very good	16.5	17.4	18.2	15.7	
Mean (SD)	2.38 (1.67)	2.61 (1.62)	2.68 (1.70)	2.76 (1.60)	0.28
		Do you feel personally addressed by the message?			
1 no, not at all	29.8	18.2	24.8	19.0	
2	5.8	5.8	6.6	9.9	
3	14.9	16.5	16.5	13.2	
4	10.7	10.7	12.4	12.4	
5 yes, very much	21.5	26.4	14.9	18.2	
Mean (SD)	2.86 (1.64)	3.28 (1.57)	2.81 (1.53)	3.01 (1.54)	0.08

ANOVA = Analysis of Variance.

Table 2. Evaluation of the quality of the illustrations and their key messages by the 121 participants in the quantitative evaluation.

	Illustration 1: Asthma	Illustration 2: Cool Like My Mom	Illustration 3: Otitis Media	Illustration 4: Bad Smell	
	Mean (Standard deviation)				pANOVA
Incomprehensible (1)/comprehensible (5)	4.64 (0.84)	4.49 (0.90)	4.30 (1.16)	4.40 (1.05)	0.01 *
Not interesting (1)/interesting (5)	4.20 (0.94)	3.96 (1.20)	4.12 (1.05)	4.06 (1.16)	0.21
Implausible (1)/plausible (5)	4.18 (1.13)	4.11 (1.12)	3.93 (1.22)	4.07 (1.30)	0.02 *
Unimportant (1)/important (5)	4.58 (0.92)	4.41 (0.98)	4.69 (0.65)	4.33 (0.97)	<0.01 **
Inappropriate (1)/appropriate (5)	4.21 (1.12)	4.10 (1.07)	3.97 (1.36)	3.90 (1.36)	0.06

ANOVA = Analysis of Variance; * Dunn-Bonferroni post hoc Test $p < 0.05$ between illustration 1 and 3; ** Dunn-Bonferroni post hoc Test $p < 0.05$ between illustration 1 and 4.

3.5. Overall Evaluation of the Campaign

Participants agreed that it would be good to include information about SHS in social media (mean 4.36; SD 0.94 on a scale from 1 'I fully disagree' to 5 'I fully agree'). However, on average they rated the topic as too sensitive to be shared in social media (mean 3.98; SD 1.29 on the same Likert scale).

The content of the key messages was evaluated very positively. The recommended measures were rated as easy to implement (mean 4.32; SD 1.06 on a scale from 1 'I do not at all agree' to 5 'I fully agree') and reasonable (mean 4.56; SD 0.80). The presented consequences of SHS were rated as realistic (mean 4.39; SD 0.92).

4. Discussion

With the chosen iterative participatory mixed methods approach, we identified social media as the best access path to the target population and developed an m-health campaign to increase knowledge about the consequences of SHS in children and about simple measures to reduce exposure. The adequacy of the campaign for the target group was confirmed in a quantitative online survey. Translating the material, the campaign might also be used in other settings and locations.

The main challenge of the project was the recruitment of participants from the target population. This is in agreement with almost all studies targeting sensitive topics like SHS exposure in children among disadvantaged families [11,35]. Due to these difficulties in recruitment, with respect to socio-economic status, the relevant criterion for inclusion in the study was the job status of the parent that took part in the study. Thus, we cannot rule out that their partners might have had a better employment status. However, it is very likely that this applied to no more than a very small proportion of the participants as in Germany spouses largely have the same socio-economic position. In order to improve participation, we worked with community partners and recruited/interviewed participants using staff whose native languages were Turkish, Spanish or Russian. Nevertheless, the target population felt uneasy to discuss this sensitive topic with members of a university hospital. The reason for this was partly that most of the invited members of the target group were aware that SHS exposure has negative health effects. They were thus concerned that researchers would stigmatise them even more. This was confirmed in the quantitative part of the study in which most participants were aware of the harms of SHS. However, as in other studies [10,15], knowledge was incomplete.

One limitation of the quantitative part of the study was the lack of a list of the target population of which a random sample could have been chosen. In addition, based on our own experience and the experience of other researchers aiming at asking disadvantaged families about a sensitive topic, we only expected a response of 10–20% [35,36]. For recruitment, we followed recommendations to contact the target population in their native languages by interculturally trained staff and to distribute the study information through multiple channels (counselling centres, clubs, doctor's practices, Internet) [36]. Furthermore, we offered financial compensation for participation. We assume that by using these measures we were able to recruit a diverse, although more motivated than average, sample. At the same time, the validity of our conclusions was ensured by our participatory mixed methods approach [37]. The consistency of the results of the qualitative and quantitative phases supports this assumption.

The subjects' concern to be confronted with their own smoking, or the smoking of their partner, as well as the possible effects on their children's health, was also evident in the quantitative evaluation of the campaign messages; the first impression of the animated illustrations was rather negative for many of the participants. On the other hand, the participants indicated that they would take a closer look at the illustrations in social media—but would rather not like or share them with others because of the unpleasant content. This converges with the results of the focus groups regarding possible obstacles to dissemination via social media. Hamill and colleagues reported similar experiences regarding the spread of an anti-smoking campaign via Facebook [38]. In their project, they followed the general recommendations for 'anti-smoking campaigns' using deterrent photos of the consequences of smoking. By this, they reached only 10% of the invited users. In accordance with our respondents, their participants stated that the photos were offensive and could provoke others if they were shared [38]. In our focus group discussions, participants suggested that the campaign should be spread by trustworthy persons, such as doctors, rather than sharing them themselves. This, in turn, coincides with the results of an anti-smoking campaign in Egypt, where the campaign was posted and advertised on Facebook and other media [38]. Such social media campaigns are relatively inexpensive and can be specifically tailored to the target audience [38,39].

In summary, our project developed an m-health campaign in close collaboration with the target group. The campaign was finally assessed positively in an independent evaluation, also carried out by members of the target group. The quality of the images and key messages was found to be satisfactory, the effectiveness and credibility rated as high, and it was confirmed that it was good to address this issue in social media. However, the topic has caused some degree of consternation among the participating smoking parents. Hence, it cannot be assumed that the pictures will be shared in social media actively by smoking parents or their partners. As suggested by the participants, dissemination should therefore be done by doctors, scientists or authorities. This could be achieved through the dissemination by paediatricians via social media, as well as through paid advertisements within social media [38,39]. These access routes need to be assessed in a follow-up study. In addition, it would be important to study whether the campaign has a lasting effect on the behaviour of smoking parents; i.e., the extent to which the proposed simple measures to reduce children's exposure to passive smoking are actually implemented. Further light on the efficacy of such campaigns may be shed by the results of a similar project in two other major German cities [40]. A recent representative population-based survey showed considerable support for tobacco control measures in Germany independent of socio-economic status (although, not surprisingly, different among smoker and non-smokers), however, indicated that such campaigns may fall on fertile ground [7].

5. Conclusions

Employing a participatory approach, we developed an m-health campaign to improve knowledge about second hand smoke in socially disadvantaged families. The campaign is ready to be disseminated in social networks, ideally by trustworthy persons such as doctors. Moreover, a follow-up study should evaluate the effectiveness of the campaign.

Supplementary Materials: The following is available online at http://www.mdpi.com/1660-4601/15/9/1945/s1, Supplement S1: List of all items used in the online survey.

Author Contributions: K.R., T.W., V.K. and D.N. planned and designed the study. F.S., J.G., S.B. and T.W. coordinated the field phase. F.S. and V.K. conducted the qualitative data analysis while S.B., K.R. and T.W. performed the quantitative analyses. The manuscript was drafted by K.R. and T.W. with contributions from all authors. All authors read and approved the final manuscript.

Funding: This research was funded by the German Federal Ministry of Health under grant number "Kapitel 15 02 Titel 684 69".

Acknowledgments: We cordially thank all participants, the collaborators and the field workers for their support.

Conflicts of Interest: DN is a delegate in the Aktionsbündnis Nichtrauchen (ABNR, Action for non-smoking) and is a member of a pharmaceutical (Pfizer) Advisory Board on Vareniclin. The other authors declare no conflicts of interest.

References

1. WHO Global Report: Mortality Attributable to Tobacco. Available online: http://www.who.int/tobacco/publications/surveillance/rep_mortality_attributable/en/ (accessed on 5 September 2018).
2. Rockville. *The Health Consequences of Smoking—50 Years of Progress. A Report of the Surgeon General*; U.S. Department of Health and Human Servives: Atlanta, GA, USA, 2014.
3. Tabakrauchbelastung von Kindern in Bayern: Ansatzpunkte für Gesundheitsförderungsstrategien auf Gemeindeebene. Available online: https://www.lgl.bayern.de/gesundheit/arbeitsplatz_umwelt/projekte_a_z/doc/gme_tabakrauchbelastung_kinder_abschlussbericht.pdf (accessed on 14 August 2018).
4. King, B.A.; Patel, R.; Babb, S.D.; Hartman, A.M.; Freeman, A. National and state prevalence of smoke-free rules in homes with and without children and smokers: Two decades of progress. *Prev. Med.* **2016**, *82*, 51–58. [CrossRef] [PubMed]
5. Mbulo, L.; Palipudi, K.M.; Andes, L.; Morton, J.; Bashir, R.; Fouad, H.; Ramanandraibe, N.; Caixeta, R.; Dias, R.C.; Wijnhoven, T.M.; et al. Secondhand smoke exposure at home among one billion children in 21 countries: Findings from the global adult tobacco survey (gats). *Tob. Control* **2016**, *25*, e95–e100. [CrossRef] [PubMed]
6. Müller, S.; Kraus, L.; Piontek, D.; Pabst, A. Changes in exposure to secondhand smoke and smoking behavior. *SUCHT* **2010**, *56*, 373–384. [CrossRef]
7. Boeckmann, M.; Kotz, D.; Shahab, L.; Brown, J.; Kastaun, S. German public support for tobacco control policy measures: Results from the German study on tobacco use (DEBRA), a representative national survey. *Int. J. Environ. Res. Public Health* **2018**, *15*, 696. [CrossRef] [PubMed]
8. Kuntz, B.; Lampert, T. Smoking and passive smoke exposure among adolescents in Germany. *Dtsch Ärztebl Int.* **2016**, *113*, 23–30. [CrossRef] [PubMed]
9. Klingshirn, H.; Hendrowarsito, L.; Fromme, H.; Bolte, G.; Studiengruppe, G.M.E. Die Bedeutung des Migrationshintergrundes fur die Tabakrauchbelastung von Kindern. Eine Querschnittstudie im Rahmen der Gesundheits-Monitoring-Einheiten (GME) in Bayern. *Gesundheitswesen* **2014**, *76*, e14–e22. (In German) [PubMed]
10. Jones, L.L.; Atkinson, O.; Longman, J.; Coleman, T.; McNeill, A.; Lewis, S.A. The motivators and barriers to a smoke-free home among disadvantaged caregivers: Identifying the positive levers for change. *Nicotine Tob. Res.* **2011**, *13*, 479–486. [CrossRef] [PubMed]
11. Rowa-Dewar, N.; Amos, A. Disadvantaged parents' engagement with a national secondhand smoke in the home mass media campaign: A qualitative study. *Int. J. Environ. Res. Public Health* **2016**, *13*, 901. [CrossRef] [PubMed]
12. Mai, Y.; Leonardo, S.; Soulakova, J.N. Smoke-free homes among single-parent families: Differences associated with parental race/ethnicity and smoking behaviors. *Prev. Med. Rep.* **2018**, *9*, 18–23. [CrossRef] [PubMed]
13. Cheah, Y.K.; Teh, C.H.; Lim, H.K. Factors associated with second-hand smoke awareness in Malaysia. *Proc. Singap. Healthc.* **2017**, *27*, 148–156. [CrossRef]
14. Schaeffer, D.; Berens, E.M.; Vogt, D. Health literacy in the German population. *Dtsch. Ärztebl. Int.* **2017**, *114*, 53–60. [PubMed]

15. Passey, M.E.; Longman, J.M.; Robinson, J.; Wiggers, J.; Jones, L.L. Smoke-free homes: What are the barriers, motivators and enablers? A qualitative systematic review and thematic synthesis. *BMJ Open* **2016**, *6*, e010260. [CrossRef] [PubMed]
16. Savas, L.S.; Mullen, P.D.; Hovell, M.F.; Escoffrey, C.; Fernandez, M.E.; Jones, J.A.; Cavazos, J.; Gutierrez Monroy, J.A.A.; Kegler, M.C. A qualitative study among Mexican Americans to understand factors influencing the adoption and enforcement of home smoking bans. *Nicotine Tob. Res.* **2017**, *19*, 1465–1472. [CrossRef] [PubMed]
17. Kuntz, B.; Lampert, T. Social disparities in parental smoking and young children's exposure to secondhand smoke at home: A time-trend analysis of repeated cross-sectional data from the German KiGGS study between 2003–2006 and 2009–2012. *BMC Public Health* **2016**, *16*, 485. [CrossRef] [PubMed]
18. Rosen, L.J.; Myers, V.; Winickoff, J.P.; Kott, J. Effectiveness of interventions to reduce tobacco smoke pollution in homes: A systematic review and meta-analysis. *Int. J. Environ. Res. Public Health* **2015**, *12*, 16043–16059. [CrossRef] [PubMed]
19. Rosen, L.J.; Myers, V.; Hovell, M.; Zucker, D.; Ben Noach, M. Meta-analysis of parental protection of children from tobacco smoke exposure. *Pediatrics* **2014**, *133*, 698–714. [CrossRef] [PubMed]
20. Family and Carer Smoking Control Programmes for Reducing Children's Exposure to Environmental Tobacco Smoke. Available online: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD001746.pub3/abstract (accessed on 1 March 2014).
21. Tobacco Advisory Group of the Royal College of Physicians. *Passive Smoking in Children: A Report*; Royal College of Physicians: London, UK, 2010.
22. Israel, B.A.; Schulz, A.J.; Parker, E.A.; Becker, A.B. Review of community-based research: Assessing partnership approaches to improve public health. *Annu. Rev. Public Health* **1998**, *19*, 173–202. [CrossRef] [PubMed]
23. Weiner, J.; McDonald, J.A. Three models of community-based participatory research. *LDI Issue Briefs* **2013**, *18*, 1–8. [PubMed]
24. Nastasi, B.K.; Hitchcock, J.; Sarkar, S.; Burkholder, G.; Varjas, K.; Jayasena, A. Mixed methods in intervention research: Theory to adaptation. *J. Mixed Methods Res.* **2007**, *1*, 164–182. [CrossRef]
25. Neuhauser, L.; Kreps, G. The advent of e-health. How interactive media are transforming health communication. *M&K* **2003**, *51*, 541–556.
26. Wayne, N.; Ritvo, P. Smartphone-enabled health coach intervention for people with diabetes from a modest socioeconomic strata community: Single-arm longitudinal feasibility study. *J. Med. Int. Res.* **2014**, *16*, e149. [CrossRef] [PubMed]
27. Bock, B.C.; Rosen, R.K.; Barnett, N.P.; Thind, H.; Walaska, K.; Foster, R.; Deutsch, C.; Traficante, R. Translating behavioral interventions onto mhealth platforms: Developing text message interventions for smoking and alcohol. *JMIR Mhealth Uhealth* **2015**, *3*, e22. [CrossRef] [PubMed]
28. Sukalla, F.; Karnowski, V.; Radon, K.; Barth, S.; Gerlich, J.; Weinmann, T. Transdisziplinarität und partizipativität beI der entwicklung einer kampagne zur reduktion der passivrauchbelastung beI kindern. In *Gesundheitskommunikation als Transdisziplinäres Forschungsfeld*, 1st ed.; Lampert, C., Grimm, M.M.A., Eds.; Nomos Verlagsgesellschaft mbH & Co. KG: Baden-Baden, Germany, 2017; pp. 187–198.
29. Thomas, S.; Heinrich, S.; Kuhnlein, A.; Radon, K. The association between socioeconomic status and exposure to mobile telecommunication networks in children and adolescents. *Bioelectromagnetics* **2010**, *31*, 20–27. [PubMed]
30. Mayring, P. *Qualitative Inhaltsanalyse. Grundlagen und Techniken*; Beltz: Weinheim, Deutschland, 2003.
31. Witte, K. Putting the fear back into fear appeals: The extended parallel process model. *Commun. Monogr.* **1992**, *59*, 329–349. [CrossRef]
32. Ajzen, I.; Albarracín, D. Predicting and changing behavior: A reasoned action approach. In *Prediction and Change of Health Behavior: Applying the Reasoned Action Approach*; Lawrence Erlbaum Associates Publishers: Mahwah, NJ, USA, 2007; pp. 3–21.
33. Rosenstock, I.M. Why people use health services. *Milbank Q.* **2005**, *8*. [CrossRef]
34. Mayring, P. *Qualitative Inhaltsanalyse*; Beltz: Weinheim, Deutschland, 2010.
35. Berens, E.M.; Riedel, J.; Reder, M.; Razum, O.; Kolip, P.; Spallek, J. Postalische Befragung von Frauen mit türkischem Migrationshintergrund-Identifizierung, Stichprobenbereinigung und Response im Rahmen der Inema-Studie. *Gesundheitswesen* **2017**, *79*, 1000–1003. (In German) [CrossRef] [PubMed]

36. Reiss, K.; Makarova, N.; Spallek, J.; Zeeb, H.; Razum, O. Identifizierung und Rekrutierung von Menschen mit Migrationshintergrund fur epidemiologische Studien in Deutschland. *Gesundheitswesen* **2013**, *75*, e49–e58. (In German) [CrossRef] [PubMed]
37. Leviton, L.C. Generalizing about public health interventions: A mixed-methods approach to external validity. *Annu. Rev. Public Health* **2017**, *38*, 371–391. [CrossRef] [PubMed]
38. Hamill, S.; Turk, T.; Murukutla, N.; Ghamrawy, M.; Mullin, S. I 'like' mpower: Using facebook, online ads and new media to mobilise tobacco control communities in low-income and middle-income countries. *Tob. Control* **2015**, *24*, 306–312. [CrossRef] [PubMed]
39. Derrick, J.L.; Eliseo-Arras, R.K.; Hanny, C.; Britton, M.; Haddad, S. Comparison of internet and mailing methods to recruit couples into research on unaided smoking cessation. *Addict. Behav.* **2017**, *75*, 12–16. [CrossRef] [PubMed]
40. Zentrum für Interdisziplinäre Suchtforschung der Universität Hamburg. MIRAS-Miteinander Rauchbelastung senken. Available online: http://www.zis-hamburg.de/projekte/projektdetails/MIRAS-Miteinander-Rauchbelastung-senken (accessed on 6 September 2018).

International Journal of
Environmental Research and Public Health

MDPI

Article

Impact of the Regional Network for AMI in the Management of STEMI on Care Processes, Outcomes and Health Inequities in the Veneto Region, Italy

Mario Saia [1], Domenico Mantoan [2], Marco Fonzo [3], Chiara Bertoncello [4,*], Marta Soattin [3], Milena Sperotto [4], Tatjana Baldovin [4], Patrizia Furlan [4], Maria Luisa Scapellato [5], Guido Viel [6], Vincenzo Baldo [4], Silvia Cocchio [4] and Alessandra Buja [4]

1. Medical Directorate, ULSS 6 Euganea, Veneto Region, Via Enrico degli Scrovegni 14, 35131 Padua, Italy; mario.saia@sanita.padova.it
2. Health and Social Services, Veneto Region, Dorsoduro, 3493—Rio Nuovo, 30123 Venice, Italy; domenico.mantoan@regione.veneto.it
3. School of Specialization in Hygiene, Preventive Medicine and Public Health, University of Padua, Via Loredan 18, 35128 Padua, Italy; marco.fonzo@studenti.unipd.it (M.F.); marta.soattin@studenti.unipd.it (M.S.)
4. Department of Cardiological, Thoracic and Vascular Sciences, Hygiene and Public Health Unit, University of Padua, Via Loredan 18, 35128 Padua, Italy; milenasperotto1978@gmail.com (M.S.); tatjana.baldovin@unipd.it (T.B.); patrizia.furlan.73@gmail.com (P.F.); vincenzo.baldo@unipd.it (V.B.); silvia.cocchio@unipd.it (S.C.); alessandra.buja@unipd.it (A.B.)
5. Department of Cardiological, Thoracic and Vascular Sciences, Occupational Medicine Unit, University of Padua, Via Loredan 18, 35128 Padua, Italy; marialuisa.scapellato@unipd.it
6. Department of Cardiological, Thoracic and Vascular Sciences, Legal Medicine Unit, University of Padua, Via Loredan 18, 35128 Padua, Italy; guido.viel@unipd.it

* Correspondence: chiara.bertoncello@unipd.it; Tel.: +39-049-8275406

Received: 13 August 2018; Accepted: 9 September 2018; Published: 11 September 2018

Abstract: Cardiovascular diseases are a leading cause of death in Europe. Outcomes in terms of mortality and health equity in the management of patients with ST-Elevation Myocardial Infarction (STEMI) are influenced by health care service organization. The main aim of the present study was to examine the impact of the new organizational model of the Veneto Region's network for Acute Myocardial Infarction (AMI) to facilitate primary percutaneous coronary intervention (PCI) on STEMI, and its efficacy in reducing health inequities. A retrospective cohort study was conducted on HDRs in the Veneto Region for the period 2007–2016, analyzing 65,261 hospitalizations for AMI. The proportion of patients with STEMI treated with PCI within 24 h increased significantly for men and women, and was statistically much higher for patients over 75 years of age (APC, 75–84: 9.8; >85: 12.5) than for younger patients (APC, <45: 3.3; 45–64: 4.9), with no difference relating to citizenship. The reduction in in-hospital, STEMI-related mortality was only statistically significant for patients aged 75–84 (APC: −3.0 [−4.5;−1.6]), and for Italians (APC: −1.9 [−3.2;−0.6]). Multivariate analyses confirmed a reduction in the disparities between socio-demographic categories. Although the new network improved the care process and reduced health care disparities in all subgroups, these efforts did not result in the expected survival benefit in all patient subgroups.

Keywords: health care research; quality assurance; hospital management; health inequities

1. Introduction

Cardiovascular diseases are a primary cause of death in Europe, responsible for more than half of all deaths across the region, with heart disease or stroke the leading cause of death in all 52 countries [1].

Ischemic heart disease now accounts for almost 1.8 million deaths a year, or 20% of all deaths in Europe, albeit with large variations between countries [2]. There has been an overall declining trend in the mortality due to ischemic heart disease over the past three decades, however [3].

Outcomes in patients with ST-Elevation Myocardial Infarction (STEMI) are influenced by many factors, some of which could relate to how health care services are organized, such as the availability of emergency medical services based on STEMI networks, which can have an impact on the delay before a patient is treated. In patients clinically suspected of having myocardial ischemia and ST-segment elevation, reperfusion therapy needs to be initiated as soon as possible [4,5]. In the literature, some studies have shown that obtaining an electrocardiogram (ECG) before reaching the hospital significantly reduces the door-to-balloon time, and that subsequently admitting ambulances directly to the cardiac catheterization laboratory (CCL), bypassing the emergency room (ER), substantially reduces the time to primary percutaneous coronary intervention (PCI) [6,7], resulting in a significant time saving for patients [4,5]. Changes in health care service organization such as prehospital activation have led to the patients involved having the highest ejection fractions, and the shortest hospital stays [8], and other studies have reported a reduction in mortality too [9].

According to the literature, myocardial infarction management may differ in relation to patients' characteristics, such as age, race and gender [10–12], and whether disparities in access to cardiac procedures translates into a different mortality risk is not known [13]. In recent years, there have been national and global efforts to rectify such health inequities, but few studies have investigated how successful they have been. One framework for action defined six strategies imperative for eliminating disparities in cardiovascular health care [14], and one of these strategies involves collecting health care data by ethnicity and gender to orient program development, implementation and assessment [15].

A study conducted at one American clinic showed that a systems-based approach to STEMI care reduced gender disparities and improved STEMI care and outcomes in women [16]. In another larger American study it emerged that a state-wide STEMI regionalization program was associated with improvements in treatment times for female, black, and elderly patients comparable with those for middle-aged, white, male patients [17]. Finally, in a recent European population-based study, socioeconomic inequity of access to revascularization was no longer apparent following the redesign of revascularization services in the South Wales Cardiac Network [18]. Such studies have been rather scarce in the European context, however. Hence the present study to assess the impact of the network implemented for AMI in the Veneto Region over the past 10 years (based on a 'hub and spoke' organizational model) in terms of the quality and outcome of the care process, and its efficacy in reducing health inequities.

2. Materials and Methods

2.1. Geographical and Socio-Demographic Context

The Veneto Region is located in the north-eastern part of Italy and is the eighth largest region in Italy with a total area of 18,398.9 km^2. In terms of population, on 1 January 2017, the Veneto Region was the fifth largest Italian region, with just over 4.9 million inhabitants and approximately 485,000 immigrants residing in the region, accounting for 9.9% of the total population [19]. In common with the rest of Italy, the main demographic feature of the Veneto Region is its aging population: some 20% of the Veneto population is over 65 years old and 10% are over 75. It is one of the most affluent regions in terms of per-capita income, with only 6% of families living in relative poverty [20]. A quarter of the population has a chronic health problem, including 65% of those over 65, and more than 25% of those aged 75–84; and more than 57% of the population over 75 are disabled [21].

2.2. Health Care System

The Italian health care system is essentially region-based [22]. In the Veneto Region, the regional authorities take full responsibility for organizing and administering public-financed health care through their regional health departments in accordance with a national health plan designed to assure an equitable provision of comprehensive care throughout the country. Responsibility for planning health care is shared by the central government and the regional authorities.

The regional authorities coordinate and control local health units (LHU) (Regional Law No. 19/2016), each of which is a separate unit within the National Health System (NHS) that plans and delivers health care services to its local community, based on a regional health plan, and determines the regional health care reorganization, specific for each region.

In the Veneto Region, the hospital network includes seven large "hubs" with regional or provincial catchment areas, and 20 local "spokes" each serving around 200,000 residents. There are also 40 smaller "node-in-the-net" hospitals that provide integrative or specific health services (such as mental health services or rehab clinics).

The LHUs are responsible for assessing needs and providing comprehensive care to their local population, either using their own staff and facilities or contracting the services out to public hospital enterprises and for-profit and non-profit independent hospitals and specialist outpatient service providers. Private providers must be accredited and have a contract with the LHU.

Tax contributions allocated to the National Health Fund are redistributed horizontally between the regions using a weighted capitation mechanism. LHU services are financed by the regional governments. Each LHU is managed and governed by a general manager appointed by the regional department of health, based on his/her professional qualifications and technical skills. This general manager appoints a financial manager and a medical director.

Services are structured according to a typical division-based model. Each division has financial autonomy over, and technical responsibility for one of three different health care system areas (Legislative decrees No. 502/1992 and No. 229/1999): directly-managed acute care and rehabilitation hospitals, health districts, and health promotion.

2.3. The Veneto Region's AMI Network

The management of Acute Myocardial Infarction (AMI) has been tackled from an organizational standpoint by adopting a 'hub and spoke' organizational model since 2008 [23]. The approach is based on the consensus document approved by the Italian Association of Hospital Cardiologists (Associazione Nazionale Medici Cardiologi Ospedalieri, ANMCO) and the Italian Society for Telemedicine (Società Italiana per la salute digitale e la Telemedicina, SIT) [24].

A characteristic of the Veneto Region's network lies in allowing a selective referral from satellite cardiology units (spokes) to cardiac surgical centers (hubs) with a CCL that operates 24 h a day and takes patients for primary PCI immediately.

A pre-hospital 12-lead ECG is remotely transmitted from every ambulance and every ER to the nearest cardiology unit, where the ECG-based diagnosis can be confirmed by skilled operators with direct access to the CCL at the hub hospital, possibly with fast tracks to bypass the ER, or Coronary Care Unit, or Cardiac Intensive Care Unit. A particular feature of this organizational model is that it goes beyond the borders of the single Local Health Districts: patients are admitted to the nearest CCL anywhere in the Veneto Region, considering only the geographical distance and journey time.

This model is based on the assumption that: (i) primary PCI performed as fast as possible and in good time is the preferred treatment strategy for ST-Elevation Myocardial Infarction (STEMI), as recommended by the European Society of Cardiology (ESC); (ii) the delay due to health care system organizational issues can be reduced by implementing a network enabling spoke hospitals to transfer patients promptly to a hub center for PCI, and by ensuring an efficient ambulance transfer service if the Cardiac Catheterization Laboratory (CCL) can be activated by the ambulance en route to hospital; (iii) high-level, expensive skills are needed for certain situations and complex diseases, and cannot be made available everywhere, but should be concentrated instead at highly-specialized regional centers (hubs) to which patients from local hospitals (spokes) can be promptly referred.

2.4. Materials

A retrospective cohort study was conducted using hospital discharge records (HDRs) from both public and private hospitals. We included residents or non-residents admitted to hospital for AMI who were discharged from any hospital operating under the National Health System (NHS) in the Veneto Region (north-east Italy) between the 1 January 2007 and the 31 December 2016.

The anonymized data obtainable from HDRs include, among others, a patient's gender, date of birth, and citizenship, and details of the type of admission, such as discharge date, length of stay, dates of interventions, discharge status, and codes for diagnoses and procedures according to the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM), with the 24th version of the DRG Grouper. We classified citizenship as Italian, citizenship of Highly-Developed Countries (HDC), or citizenship of High Migratory Pressure Countries (HMPC). The HMPC included new Member States of the European Union, countries in Central-Eastern Africa, Asia (except for Israel, the People's Democratic Republic of Korea, and Japan), and Central and South America; by extension, stateless individuals were also included in this group. The HDC included the other European countries, North America, Oceania, Israel and Japan.

Inclusion criteria were adopted to examine the performance of the Veneto regional network specifically in the management of STEMI. We therefore considered the ICD-9-CM diagnostic codes 410.x1 as the principal diagnosis to define AMI (initial episode of care): STEMI was identified using codes 410.01, 410.11, 410.21, 410.31, 410.41, 410.51, 410.61, 410.81 and 410.91, and non-STEMI AMI (NSTEMI) using code 410.71. Patients who underwent Coronary Artery Bypass Graft (CABG) surgery (ICD-9-CM code 36.1x) were excluded.

The study complied with the Helsinki Declaration and Italian privacy legislation (n. 196/2003). Resolution n. 85/2012 of the Guarantor for the protection of personal data has recently confirmed that anonymized personal data may be processed for medical, biomedical and epidemiological research purposes, and data concerning health status may be used in aggregate form in scientific studies.

2.5. Performance Indicators

Two primary indicators describing the performance of the dedicated network were considered, one for care processes and one for outcomes, using as dependent variables:

- The proportion of patients with STEMI treated with PCI within 24 h, since this was considered the treatment of choice in cases of STEMI; and
- The proportion of in-hospital deaths among patients treated for STEMI.

We used the following ICD-9-CM procedure codes to define PCI: 00.66, 36.01, 36.02, 36.05, 36.06 and 36.07. The Regional Hospital Information System reports the date, but not the exact time when procedures are performed so, for the purposes of this study, it was assumed that PCI was performed within 24 h if it was done on the day of admission to hospital.

2.6. Statistical Analysis

AMI hospitalization rates were calculated based on the size of the population in every year of the study period, and are reported as the rates per 100,000 population. Direct age and gender adjustments of in-hospital admissions for AMI (per 100,000 population) by type of AMI were performed using the data for the population in 2012.

Descriptive statistics were obtained for the variable investigated. Patients' characteristics are presented as means and standard deviations (SD), for continuous variables, and using absolute and relative frequencies for categorical data.

We examined the two performance indicators by subgroup for gender (men and women), age (<45, 45–64, 65–74, 75–84, >84), and citizenship, for each year between 2007 and 2016. The annual percent changes (APC) and relative confidence intervals (CIs) were calculated to identify trends for all the variables investigated.

Two predictive models were built to estimate the relationships between the performance indicators and socio-demographic factors and, to take the hierarchical structure of the data into account, we performed multilevel logistic regression with random intercepts for hospitals, controlling for potential intra-hospital correlations. Next, to investigate disparities in outcomes by gender, age, and citizenship, we pooled the calendar years into two periods, before (2007–2008) and after (2009–2016) the introduction of the network. All statistical analyses were performed using STATA software, version 12.1 (Stata Corp LP, College Station, TX, USA). All p-values reported are two-sided, with a significance threshold of $p < 0.05$.

3. Results

Overall, there were 65,261 hospitalizations with a principal diagnosis at discharge of AMI (initial episode of care) in the decade investigated, among all hospitals operating under the NHS in the Veneto Region (north-east Italy), 55.2% of them (36,035) were STEMI, and 44.8% (29,226) were NSTEMI. The overall mean age of this patient population was 71.1 ± 13.7 years (data not shown).

Figure 1 shows that the incidence of NSTEMI exceeded that of STEMI in the last year of the decade (74.8 vs. 72.7 cases per 100,000 population). This phenomenon was possibly due to a simultaneous decrease in cases of STEMI (−34%; APC:C-4.8 [−5.3;−4.4], $p < 0.001$), and increase in NSTEMI (+7%; APC: 1.1 [0.4;1.7], $p < 0.001$). On the whole, the absolute number of hospital admissions for AMI dropped by 18% (APC: −2.3 [−2.6;−1.9], $p < 0.001$) over the decade.

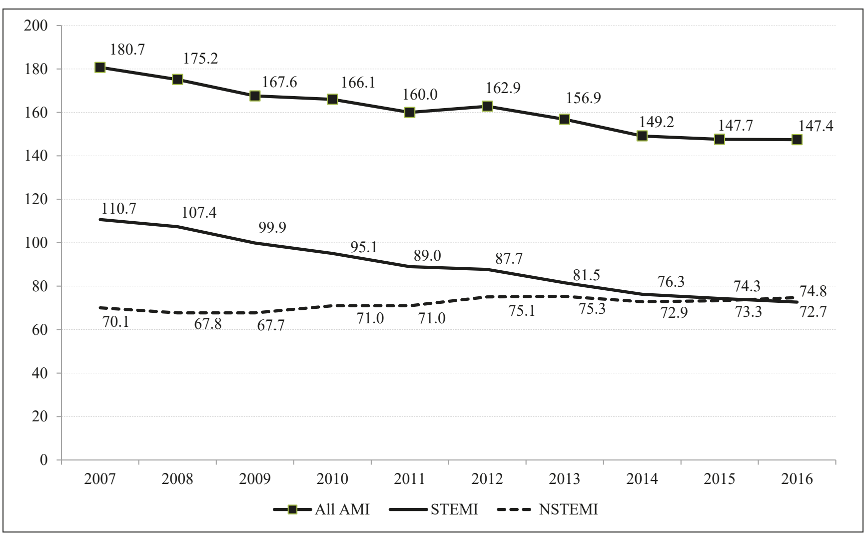


Figure 1. Veneto Region 2007–2016. Temporal trends in hospital admissions for AMI * (per 100,000 population), initial episode of care, by type of AMI, after direct standardization (reference population: 2012). AMI: acute myocardial infarction. * APC (CI 95%), p-value → All STEMI = −2.3 (−2.6;−1.9), $p < 0.001$; STEMI = −4.8 (−5.3;−4.4), $p < 0.001$; NSTEMI = 1.1 (0.4;1.7), $p < 0.001$.

The mean age of the STEMI population was 70.1 ± 14.3 years, and males accounted for nearly two thirds of the patients hospitalized with STEMI (men: 66.2%; women: 33.8%). The majority of patients were of Italian citizenship (95.5%) (Table 1). In the decade from 2007 to 2016, 44.6% of patients admitted to hospital for STEMI as their principal diagnosis underwent a PCI within 24 h of admission; and the in-hospital mortality rate for STEMI was 12.2% (Table 1).

Table 1. Veneto Region 2007–2016. Characteristics of hospital admissions for STEMI, in-hospital mortality and proportion of PCI on day of admission by gender, age, and citizenship.

	Hospital Admissions		Mean Age	Average Hospital Stay	PCI on Day of Admission		In-Hospital Mortality	
	n	%	(y) ± SD	d ± SD	*n*	%	*n*	%
Total	36,035	-	70.1 ± 14.3	9.24 ± 8.80	16,075	44.6	4385	12.2
Gender								
Male	23,862	66.2	66.2 ± 13.5	8.82 ± 8.75	12,284	51.5	2162	9.1
Female	12,173	33.8	77.7 ± 12.7	10.04 ± 8.83	3791	31.1	2223	18.3
Age								
<45 y	1424	3.9	40.0 ± 4.4	7.34 ± 8.44	844	59.3	18	1.3
45–64 y	11,415	31.7	55.9 ± 5.4	7.61 ± 6.51	7012	61.4	312	2.7
65–74 y	7949	22.1	69.6 ± 2.9	9.19 ± 8.96	4039	50.8	536	6.7
75–84 y	8771	24.3	79.6 ± 2.9	10.90 ± 10.63	3147	35.9	1432	16.3
>84 y	6476	18.0	89.4 ± 3.7	10.33 ± 8.86	1033	16.0	2087	32.2
Citizenship								
Italian	34,414	95.5	70.7 ± 14.0	8.4 ± 7.3	15,130	44.0	4314	12.5
HDC	394	10.9	62.9 ± 11.8	8.2 ± 10.4	262	66.5	29	7.4
HMPC	1127	3.1	52.7 ± 12.2	11.8 ± 10.9	645	57.2	38	3.4

PCI: percutaneous coronary intervention; HDC: highly-developed countries; HMPC: high migratory pressure countries.

Table 2 shows the proportion of patients with STEMI treated with PCI within 24 h by socio-demographic factors and year of the study. Throughout the period investigated, this care process indicator was higher for men than for women (e.g., 38.1% vs. 18.6% in 2007), and it increased significantly over time for both genders, and especially for women, though the difference between genders was not statistically significant (APC: 6.3 [5.4;7.3] for men, and APC: 8.6 [6.2;11.1] for women). When age groups were considered, a statistically significant rising trend emerged over the ten-year period, which was statistically far more consistent for patients over 75 years old (APC, 75–84 y: 9.8; >85 y: 12.5) than for younger patients (APC, <45 y: 3.3; 45–64: 4.9). As regards citizenship, the significantly rising trend in the proportion of patients with STEMI treated with PCI within 24 h was confirmed, with no significant differences between the three citizenship groups identified (APC, Italian: 7.2; HDC: 6.1; HMPC: 5.9).

Table 2. Veneto Region 2007–2016. Proportion of PCI within 24 h of hospital admission for STEMI by gender, age, and citizenship.

	% of PCI Within 24 h										
	2007 (*n* = 4040)	**2008 (*n* = 4083)**	**2009 (*n* = 3908)**	**2010 (*n* = 3793)**	**2011 (*n* = 3644)**	**2012 (*n* = 3536)**	**2013 (*n* = 3362)**	**2014 (*n* = 3252)**	**2015 (*n* = 3228)**	**2016 (*n* = 3189)**	**APC (CI 95%)**
Gender											
Male	38.1	40.8	42.5	48.3	49.9	54.4	59.0	61.0	62.2	64.2	6.3 (5.4;7.3) °
Female	18.6	22.5	28.6	28.5	32.4	32.9	38.2	38.8	40.8	40.6	8.6 (6.2;11.1) °
Age											
<45 y	51.3	47.0	60.9	53.0	58.8	63.4	69.4	69.7	66.7	59.6	3.3 (0.9;5.8) °
45–64 y	47.8	49.9	53.2	61.2	60.9	63.2	68.7	68.9	71.8	71.9	4.9 (3.8;6.0) °
65–74 y	35.2	39.2	43.6	44.5	50.6	54.8	59.2	61.4	63.4	64.5	7.2 (6.0;8.4) °
75–84 y	22.0	25.9	27.5	32.8	35.9	37.7	44.9	47.5	47.8	50.5	9.8 (8.2;11.5) °
>84 y	8.0	9.4	12.6	14.9	14.9	19.3	17.3	21.9	22.1	24.2	12.5 (9.4;15.6) °
Citizenship											
Italian	30.4	33.8	37.2	40.8	43.7	46.9	51.1	53.1	54.7	56.2	7.2 (6.1;8.3) °
HDC	50.0	46.4	68.3	46.0	62.9	61.3	92.5	77.3	74.5	76.3	6.1 (1.8;10.6) °
HMPC	41.1	44.6	45.5	60.0	54.3	52.3	65.3	66.1	63.4	70.9	5.9 (3.8;8.1) °

PCI: percutaneous coronary intervention; APC: annual percent change; HDC: highly-developed countries; HMPC: high migratory pressure countries. ° $p < 0.05$.

Table 3 shows the proportion of in-hospital deaths among patients treated for STEMI during the period considered, for all the subgroups investigated. Women with STEMI always had a higher in-hospital mortality rate than men (e.g., 18.9% vs. 9.9% in 2007). From 2007 to 2016, the overall observed in-hospital mortality rate declined for both genders, with no significant difference between the trends of the two groups (APC: men, −1.9 [−3.7;0.0]; women, −1.2 [−3.1;0.6]). When the different age groups were analyzed, it emerged that age had a fundamental impact on the chances of PCI within 24 h, but there was no significant difference between the declining mortality trends for the various age groups, except for the 75- to 84-year-olds (APC: −3.0 [−4.5;−1.6]). The same declining mortality trend was apparent for the citizenship subgroups too, leading to a difference that was significant for Italian patients (APC: −1.9 [−3.2;−0.6]), and greater—though it failed to reach statistical significance—for patients from HDC countries (APC: −4.9 [−12.6;3.5]).

Table 3. Veneto Region 2007–2016. In-hospital mortality after hospital admission for STEMI by gender, age, and citizenship.

	% in-Hospital Mortality										
	2007 (*n* = 4040)	**2008 (*n* = 4083)**	**2009 (*n* = 3908)**	**2010 (*n* = 3793)**	**2011 (*n* = 3644)**	**2012 (*n* = 3536)**	**2013 (*n* = 3362)**	**2014 (*n* = 3252)**	**2015 (*n* = 3228)**	**2016 (*n* = 3189)**	**APC (CI 95%)**
Gender											
Male	9.9	8.7	9.8	9.0	9.6	9.3	8.5	9.8	8.2	7.5	−1.9 (−3.7;0.0)
Female	18.9	19.4	16.7	18.4	20.6	17.8	19.4	16.6	17.7	16.2	−1.2 (−3.1;0.6)
Age											
<45 y	1.3	1.8	-	0.7	3.3	0.6	0.8	1.6	1.8	1.0	−1.6 (−16.0;15.3) *
45–64 y	3.1	1.9	2.6	2.5	2.7	2.7	3.0	4.1	2.3	2.5	1.4 (−3.9;6.9)
65–74 y	7.7	6.5	6.4	7.2	8.1	7.3	6.0	7.2	5.9	4.8	−3.0 (−6.2;0.3)
75–84 y	18.1	17.6	17.3	16.1	16.8	17.9	15.3	14.1	15.1	13.1	−3.0 (−4.5;−1.6) °
>84 y	31.2	34.0	31.7	31.1	35.3	29.5	35.3	31.6	31.6	31.1	−0.2 (−1.8;1.4)
Citizenship											
Italian	13.5	12.9	12.4	12.7	13.6	12.6	12.6	12.4	11.6	10.4	−1.9 (−3.2;−0.6) °
HDC	15.6	7.1	7.3	4.0	8.6	6.5	7.5	6.8	7.3	5.3	−4.9 (−12.6;3.5)
HMPC	5.3	3.0	4.0	3.6	1.7	3.6	1.6	3.3	3.3	4.7	−1.8 (−11.5;8.9)

APC: annual percent change; HDC: highly-developed countries; HMPC: high migratory pressure countries. ° $p < 0.05$. * excluded year 2009.

Figure 2 show the results of multilevel logistic regression analyses and the significant associations between the two performance indicators and the different socio-demographic factors, before (2007–2008) and after (2009–2016) the introduction of the network.

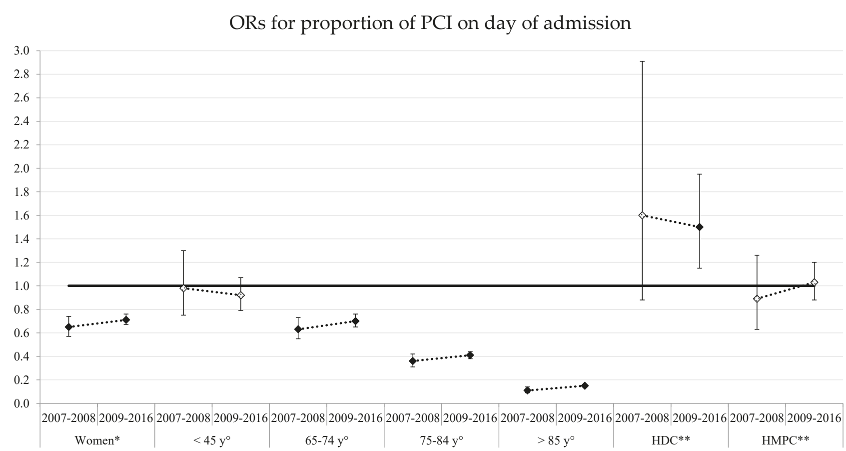


Figure 2. *Cont.*

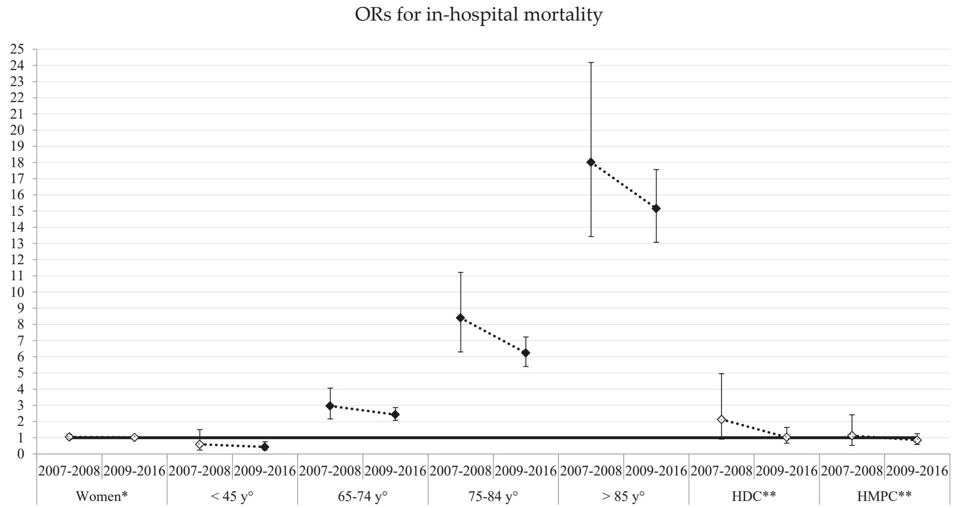


Figure 2. Veneto Region 2007–2016. Multilevel logistic regression analysis of associations between the two indicators investigated and socio-demographic factors; ORs and 95% confidence intervals, p value °°. PCI: percutaneous coronary intervention. * Reference: Men. ° Reference: 45–64 y. ** Reference: Italian. °° full indicator → $p < 0.05$.

Although disparities persisted between genders and age groups, the multivariate analyses confirmed the trend towards a null value in almost all the different categories of socio-demographic factors (where a value of 1 attests to no differences between the factors).

4. Discussion

The present study showed that the proportion of PCI performed within 24 h of admission to hospital with STEMI increased considerably over the decade from 2007 to 2016 for all socio-demographic patient subgroups in the Veneto Region. A greater standardization also reduced disparities in access to PCI by gender and age group. The study also revealed no disparities in access to PCI by citizenship, neither before nor after the implementation of the regional network. During the period studied, there was a significant reduction of the related mortality among Italians, and for people aged 75–84 years.

Overall, we noted a substantial reduction in the number of cases of AMI over time, and this reduction was entirely attributable to declining numbers of STEMI, while the incidence of NSTEMI did not show any noteworthy changes, neither by year, nor over the decade as a whole. This finding is consistent with recent national [25] and international literature [26–30], which attributes a decrease in the incidence of STEMI to improvements in primary prevention measures and a greater use of evidence-based therapeutic changes.

Our study results demonstrate that a "hub and spoke" organizational model could increase the proportion of patients given reperfusion treatment on the day of their admission to hospital, and this applied to all the population subgroups considered. These findings are in line with reports from the United States, where hospitals implementing the key care processes (prehospital CCL activation, a single-call protocol for transfer from facilities without PCI capabilities, bypassing the ER for patients presenting directly to emergency medical services or being transferred) achieved shorter median reperfusion delays than other hospitals [31]. On the other hand, our study only found an impact on the AMI-related mortality rates in some subgroups (Italians citizens, and 75- to 84-year-olds). Though information on the exact door-to-balloon (DTB) time was unavailable for our sample, a possible interpretation for this limited impact may be that the speed of response is actually still outside the critical time-window, and having increased the proportion of PCI performed within 24 h may have no major impact on survival [32,33]. Our data are partly consistent with the findings of a previous study [34], in which the authors documented a marked reduction in median DTB time and a greater compliance with the guidelines, but found in-hospital mortality unchanged. In other words, the successful implementation of organizational efforts to optimize the quality of this care process did not produce the expected survival benefit. Another recent study by Wang et al. on 5881 patients with STEMI undergoing PCI showed that the median DTB time was 101 min from January to December 2005, and dropped to 87 min from July 2006 to June 2007, but this shorter DTB time was not associated with any improvement in mortality [35].

The other aim of our study was to see whether the new network reduces health inequities among socio-demographic subgroups of the population, with a view to ensuring that the structural and procedural features of health service delivery result in an equitable distribution of the services for individuals and population subgroups with comparable needs and wants [36].

Regarding citizenship, we found a statistically significant increase from 2007 to 2016 in the number of PCI performed within 24 h in all categories of citizens, and there were no disparities between them even before the network was implemented. Similarly, a study conducted in the Netherlands [37] found no ethnic inequities in the revascularization rates after STEMI, and a UK study found none in the revascularization rates after AMI [38]. These findings contrast with the abundance of literature from the USA showing considerably lower revascularization rates after coronary events for African Americans compared with White Americans [39–41]. A major difference between Italy, the UK or the Netherlands and the USA lies in the former countries' universal health care provision instead of private health insurance schemes that fail to cover the whole population's health care needs. In fact, Italy and the UK have health systems funded by taxes that guarantee a standard level of care for everyone, and The Netherlands obliges all residents to have health insurance by law. Our study also found citizens from HDC more likely to receive prompt treatment than Italians after the network was reorganized, and this could be because the country of origin's reimbursement of the costs for these patients' treatment is allied to the demonstration of a strict adherence to the guidelines.

After the implementation of the network, there was a statistically significant increase (from 2007 to 2016) in the proportion of PCI performed within 24 h. This was true of both genders, with a greater improvement in women than in men (APC 8.6 vs. 6.3), though the difference between genders was not statistically significant. Multivariate analysis nonetheless showed that gender disparities in STEMI management persisted after the network's implementation. This is consistent with the gender disparities seen in our previous studies [42–44]. This inequity in the administration of the evidence-based best treatment for STEMI might be partly attributable to a greater delay before women access hospital care, which would prevent the adoption of a PCI approach. Whatever its cause, such disparities are a health system concern that can be solved by associating suitable strategies with the network model. In fact, even if high-quality care were accessible to everyone, there is no guaranteeing that all patients would benefit equally, so strategies are needed to ensure that patients are all aware of how to make best use of the health care services available.

We also found that the network significantly improved access to PCI within 24 h for all age groups, and especially for the previously disadvantaged older age groups (>65 years old). We only noted a significant reduction in mortality for the 75- to 84-year-olds, however, during the period examined. Despite these improvements, multivariate analysis showed that some degree of age-related inequity persisted. These results are in line with a similar study conducted in Switzerland [45], where elderly patients (and females) were at a disadvantage in the circadian provision of primary PCI in a cohort of patients with AMI. The Swiss authors surmised that physicians and other hospital staff may slow the process down for some patients more than others and, somewhere along the line, staff members make decisions about patients based on patient group membership rather than guidelines. There is no reason to believe that this reflects a pattern of deliberate behavior, however, since many studies document the presence of unconscious bias among medical personnel, and the effects these have on their decisions and actions [46,47].

More generally speaking, Peterson et al. [48] also analyzed data on approximately 2.5 million patients with STEMI admitted to US hospitals between 1990 and 2006, finding that—despite very clear, and widely-adopted ACCC/AHA STEMI and NSTEMI guidelines—there were still disparities in STEMI treatment for women, blacks, and elderly patients. The results of studies conducted specifically to test adherence to, and effectiveness of standardized guidelines show that, when they are in place and enforced, standardized protocols reduce disparities and improve patient outcomes [49]. The present study underscores the need to monitor the implementation of organizational models and/or standardized protocols by means of up-to-date administrative datasets to see whether a given strategy reduces health care disparities. In fact, hospital administration databases are of considerable practical value for the purpose of broad-based quality assessments, and could be useful for monitoring health equity issues too, identifying problems and targeting areas that might require more in-depth investigations based on more specific data.

The main strength of our study lies in that it was population-based, minimizing selection bias and relying on independently-collected data. Our study also has some limitations, however. First, it was an observational study, which limits our ability to draw causal inferences. In particular, we cannot disentangle whether the disparities arose: because the quality of care provided by the health care services differed according to patients' sociodemographic characteristics; or because patients' sociodemographic factors influenced their attitudes or behavior, or the clinical characteristics predisposing them to different treatment for acute events. That said, studies on disparities can only be observational.

A second limitation lies in that, despite their wide use as a valuable source for health care research on specific disease, hospital discharge data have some drawbacks in terms of their diagnostic coding accuracy (in our case, the use of the ICD-9-CM codes to identify the AMI). These codes have the advantage of being widely available and being less effortful and costly than consulting medical charts, but their effective use hinges on how reliably the codes reflect a given patient's clinical condition.

Finally, cardiovascular risk is related to many economic, social and predisposing factors, apart from age and gender. Unfortunately, our health administration records do not track all the important risk factors for AMI, such as lifestyle indicators, which could have an important influence on survival. That said, such risk factors could only have a confounding effect if they are differently associated with both outcome and exposure.

5. Conclusions

In conclusion, our study confirms the effectiveness of the new hub-and-spoke organizational model of the Veneto Region's network for AMI in facilitating primary PCI on STEMI and reducing in-hospital mortality. The present study also emphasizes that the implementation of organizational models and/or standardized protocols should be monitored with the aid of up-to-date administrative datasets to see whether a given strategy reduces health care disparities.

Based on such indicators, the present study revealed numerous opportunities for improvements and adaptations to the intervention, and for regularly collecting process measures. At the same time, it will be helpful not only to use administrative databases, but also to introduce qualitative assessment methods (e.g., surveys, interviews) to identify problems and target areas that might require more in-depth investigations based on more specific data.

Even though the new network improved the process of care for STEMI and reduced disparities in this process between all population subgroups, this did not generate the expected survival benefit in all patient subgroups.

The present findings could prompt policy-makers to seriously consider the feasibility of designing and implementing programs for health care providers, organizations, and community groups, and policies that focus not only on patients. Finally, sharing feedback with providers, and incentivizing them to reduce disparities could have an important role.

Author Contributions: V.B., M.S. (Mario Saia) obtained funding, coordinated all study phases, and approved the final manuscript as submitted. M.S. (Milena Sperotto) carried out the statistical analyses, and approved the final manuscript as submitted. P.F. designed the data collection instruments, and coordinated and supervised data collection. C.B., M.F., M.S. (Marta Soattin) drafted the manuscript, and approved the final manuscript as submitted. A.B. conceptualized the study, coordinated all study phases, was involved in data collection, and approved the final manuscript as submitted. D.M., M.L.S., G.V., T.B., S.C. critically reviewed and revised the manuscript, and the interpretation of the data.

Funding: The Department of Cardiological, Thoracic and Vascular Sciences funded this research with a competitive call for grants awarded to stimulate collaboration between colleagues within the department.

Conflicts of Interest: The authors have no conflicts of interest to disclose.

References

1. WHO Regional Office for Europe. *Gaining Health: The European Strategy for the Prevention and Control of Noncommunicable Diseases*; WHO Regional Office for Europe: Copenhagen, Denmark, 2006.
2. Townsend, N.; Wilson, L.; Bhatnagar, P.; Wickramasinghe, K.; Rayner, M.; Nichols, M. Cardiovascular disease in Europe: Epidemiological update 2016. *Eur. Heart J.* **2016**, *37*, 3232–3245. [CrossRef] [PubMed]
3. Hartley, A.; Marshall, D.C.; Salciccioli, J.D.; Sikkel, M.B.; Maruthappu, M.; Shalhoub, J. Trends in mortality from ischemic heart disease and cerebrovascular disease in Europe: 1980 to 2009. *Circulation* **2016**, *133*, 1916–1926. [CrossRef] [PubMed]
4. Amit, G.; Cafri, C.; Gilutz, H.; Ilia, R.; Zahger, D. Benefit of direct ambulance to coronary care unit admission of acute myocardial infarction patients undergoing primary percutaneous intervention. *Int. J. Cardiol.* **2007**, *119*, 355–358. [CrossRef] [PubMed]
5. Andersen, H.R.; Nielsen, T.T.; Rasmussen, K.; Thuesen, L.; Kelbaek, H.; Thayssen, P.; Abildgaard, U.; Pedersen, F.; Madsen, J.K.; Grande, P.; et al. DANAMI-2 Investigators. A comparison of coronary angioplasty with fibrinolytic therapy in acute myocardial infarction. *N. Engl. J. Med.* **2003**, *349*, 733–742. [CrossRef] [PubMed]
6. Terkelsen, C.J.; Lassen, J.F.; Nørgaard, B.L.; Gerdes, J.C.; Poulsen, S.H.; Bendix, K.; Ankersen, J.P.; Gøtzsche, L.B.; Rømer, F.K.; Nielsen, T.T.; et al. Reduction of treatment delay in patients with ST-elevation myocardial infarction: Impact of pre-hospital diagnosis and direct referral to primary percutanous coronary intervention. *Eur. Heart J.* **2005**, *26*, 770–777. [CrossRef] [PubMed]
7. Dorsch, M.F.; Greenwood, J.P.; Priestley, C.; Somers, K.; Hague, C.; Blaxill, J.M.; Wheatcroft, S.B.; Mackintosh, A.F.; McLenachan, J.M.; Blackman, D.J. Direct ambulance admission to the cardiac catheterization laboratory significantly reduces door-to-balloon times in primary percutaneous coronary intervention. *Am. Heart J.* **2008**, *155*, 1054–1058. [CrossRef] [PubMed]
8. Brown, J.P.; Mahmud, E.; Dunford, J.V.; Ben-Yehuda, O. Effect of prehospital 12-lead electrocardiogram on activation of the cardiac catheterization laboratory and door-to-balloon time in ST-segment elevation acute myocardial infarction. *Am. J. Cardiol.* **2008**, *101*, 158–161. [CrossRef] [PubMed]
9. Quinn, T.; Johnsen, S.; Gale, C.P.; Snooks, H.; McLean, S.; Woollard, M.; Weston, C.; Myocardial Ischaemia National Audit Project (MINAP) Steering Group. Effects of prehospital 12-lead ECG on processes of care and mortality in acute coronary syndrome: A linked cohort study from the Myocardial Ischaemia National Audit Project. *Heart* **2014**, *100*, 944–950. [CrossRef] [PubMed]
10. Edmund Anstey, D.; Li, S.; Thomas, L.; Wang, T.Y.; Wiviott, S.D. Race and sex differences in management and outcomes of patients after ST-elevation and non-ST-elevation myocardial infarct: Results From the NCDR. *Clin. Cardiol.* **2016**, *39*, 585–595. [CrossRef] [PubMed]
11. Canto, J.G.; Allison, J.J.; Kiefe, C.I.; Fincher, C.; Farmer, R.; Sekar, P.; Person, S.; Weissman, N.W. Relation of race and sex to the use of reperfusion therapy in Medicare beneficiaries with acute myocardial infarction. *N. Engl. J. Med.* **2000**, *342*, 1094–1100. [CrossRef] [PubMed]

12. Vaccarino, V.; Rathore, S.S.; Wenger, N.K.; Frederick, P.D.; Abramson, J.L.; Barron, H.V.; Manhapra, A.; Mallik, S.; Krumholz, H.M.; National Registry of Myocardial Infarction Investigators. Sex and racial differences in the management of acute myocardial infarction, 1994 through 2002. *N. Engl. J. Med.* **2005**, *353*, 671–682. [CrossRef] [PubMed]
13. Hagen, T.P.; Häkkinen, U.; Iversen, T.; Klitkou, S.T.; Moger, T.A.; EuroHOPE Study Group. Socio-economic inequality in the use of procedures and mortality among AMI patients: Quantifying the effects along different paths. *Health Econ.* **2015**, *24*, 102–115. [CrossRef] [PubMed]
14. Mensah, G.A. Eliminating disparities in cardiovascular health: Six strategic imperatives and a framework for action. *Circulation* **2005**, *111*, 1332–1336. [CrossRef] [PubMed]
15. Colais, P.; Agabiti, N.; Fusco, D.; Pinnarelli, L.; Sorge, C.; Perucci, C.A.; Davoli, M. Inequality in 30-day mortality and the wait for surgery after hip fracture: The impact of the regional health care evaluation program in Lazio (Italy). *Int. J. Qual. Health Care* **2013**, *25*, 239–247. [CrossRef] [PubMed]
16. Huded, C.P.; Johnson, M.; Kravitz, K.; Menon, V.; Abdallah, M.; Gullett, T.C.; Hantz, S.; Ellis, S.G.; Podolsky, S.R.; Meldon, S.W. 4-Step Protocol for Disparities in STEMI Care and Outcomes in Women. *J. Am. Coll. Cardiol.* **2018**, *71*, 2122–2132. [CrossRef] [PubMed]
17. Glickman, S.W.; Granger, C.B.; Ou, F.S.; O'Brien, S.; Lytle, B.L.; Cairns, C.B.; Mears, G.; Hoekstra, J.W.; Garvey, J.L.; Peterson, E.D.; et al. Impact of a statewide ST-segment-elevation myocardial infarction regionalization program on treatment times for women, minorities, and the elderly. *Circ. Cardiovasc. Qual. Outcomes* **2010**, 514–521. [CrossRef] [PubMed]
18. Evans, L.W.; van Woerden, H.; Davies, G.R.; Fone, D. Impact of service redesign on the socioeconomic inequity in revascularisation rates for patients with acute myocardial infarction: A natural experiment and electronic record-linked cohort study. *BMJ Open* **2016**, *6*, e011656. [CrossRef] [PubMed]
19. Demografia in Cifre [Demography in Numbers][Website]; Rome: Istituto Nazionale di Statistica. 2018. Available online: http://demo.istat.it/ (accessed on 23 May 2018).
20. Istituto Nazionale di Statistica [National Institute of Statistics][Website]. Rome: National Institute of Statistics. 2017. Available online: http://dati.istat.it (accessed on 25 May 2018).
21. AUR—Anagrafe Unica Assistiti Regionale. *Portale Sanità Regione del Veneto* [*Website*]; Venice: Veneto Region, Italy. 2011. Available online: https://salute.regione.veneto.it/web/bando-ssi/aur-anagrafe-unica-assistiti-regionale (accessed on 14 January 2015).
22. World Health Organization. The Veneto Model—A Regional Approach to Tackling Global and European Health Challenges (2016). Available online: http://www.euro.who.int/en/publications/abstracts/the-veneto-model-a-regional-approach-to-tackling-global-and-european-health-challenges-2016 (accessed on 23 May 2018).
23. Regione del Veneto; Bur n. 15 del 19 Febbraio 2008; Materia: Sanità e Igiene Pubblica; Deliberazione della Giunta Regionale n. 4550 del 28 Dicembre 2007 n.d. Available online: https://bur.regione.veneto.it/BurvServices/pubblica/DettaglioDgr.aspx?id=203192 (accessed on 28 February 2018).
24. Documento di, Consenso. La rete interospedaliera per l'emergenza coronarica. *Ital. Health J.* **2005**, *6*, 56S–62S. Available online: http://www.giornaledicardiologia.it/allegati/00730_2005_12/fulltext/S6-05_02%205-26.pdf (accessed on 9 September 2018).
25. Barchielli, A.; Profili, F.; Balzi, D.; Francesconi, P.; Zuppiroli, A.; Cipriani, F. Trends in occurrence, treatment, and outcomes of acute myocardial infarction in Tuscany Region (Central Italy), 1997–2010. *Epidemiol. Prev.* **2015**, *39*, 167–175. (In Italian) [PubMed]
26. Sugiyama, T.; Hasegawa, K.; Kobayashi, Y.; Takahashi, O.; Fukui, T.; Tsugawa, Y. Differential time trends of outcomes and costs of care for acute myocardial infarction hospitalizations by ST elevation and type of intervention in the United States, 2001–2011. *J. Am. Heart Assoc.* **2015**, *4*, e001445. [CrossRef] [PubMed]
27. McManus, D.D.; Gore, J.; Yarzebski, J.; Spencer, F.; Lessard, D.; Goldberg, R.J. Recent trends in the incidence, treatment, and outcomes of patients with STEMI and NSTEMI. *Am. J. Med.* **2011**, *124*, 40–47. [CrossRef] [PubMed]
28. Yeh, R.W.; Sidney, S.; Chandra, M.; Sorel, M.; Selby, J.V.; Go, A.S. Population trends in the incidence and outcomes of acute myocardial infarction. *N. Engl. J. Med.* **2010**, *362*, 2155–2165. [CrossRef] [PubMed]
29. Chen, J.; Normand, S.L.; Wang, Y.; Drye, E.E.; Schreiner, G.C.; Krumholz, H.M. Recent declines in hospitalizations for acute myocardial infarction for Medicare fee-for-service beneficiaries: Progress and continuing challenges. *Circulation* **2010**, *121*, 1322–1328. [CrossRef] [PubMed]

30. Luepker, R.V.; Berger, A.K. Is acute myocardial infarction disappearing? *Circulation* **2010**, *121*, 1280–1282. [CrossRef] [PubMed]
31. Fordyce, C.B.; Al-Khalidi, H.R.; Jollis, J.G.; Roettig, M.L.; Gu, J.; Bagai, A.; Berger, P.B.; Corbett, C.C.; Dauerman, H.L.; Fox, K.; et al. STEMI Systems Accelerator Project. Association of rapid care process implementation on reperfusion times across multiple ST-segment-elevation myocardial infarction networks. *Circ. Cardiovasc. Interv.* **2017**, *10*, e004061. [CrossRef] [PubMed]
32. Gersh, B.J.; Stone, G.W.; White, H.D.; Holmes, D.R. Pharmacological facilitation of primary percutaneous coronary intervention for acute myocardial infarction: Is the slope of the curve the shape of the future? *JAMA* **2005**, *293*, 979–986. [CrossRef] [PubMed]
33. Reimer, K.A.; Lowe, J.E.; Rasmussen, M.M.; Jennings, R.B. The wavefront phenomenon of ischemic cell death. 1. Myocardial infarct size vs duration of coronary occlusion in dogs. *Circulation* **1977**, *56*, 786–794. [CrossRef] [PubMed]
34. Flynn, A.; Moscucci, M.; Share, D.; Smith, D.; LaLonde, T.; Changezi, H.; Riba, A.; Gurm, H.S. Trends in door-to-balloon time and mortality in patients with ST-elevation myocardial infarction undergoing primary percutaneous coronary intervention. *Arch. Intern. Med.* **2010**, *170*, 1842–1849. [CrossRef] [PubMed]
35. Wang, T.Y.; Fonarow, G.C.; Hernandez, A.F.; Liang, L.; Ellrodt, G.; Nallamothu, B.K.; Shah, B.R.; Cannon, C.P.; Peterson, E.D. The dissociation between door-to-balloon time improvement and improvements in other acute myocardial infarction care processes and patient outcomes. *Arch. Intern. Med.* **2009**, *169*, 1411–1419. [CrossRef] [PubMed]
36. Aday, L.A.; Begley, C.E.; Lairson, D.R.; Slater, C.H.; Richard, A.J.; Montoya, I.D. A framework for assessing the effectiveness, efficiency, and equity of behavioral healthcare. *Am. J. Manag. Care* **1999**, *5*, SP25–SP44. [PubMed]
37. Van Oeffelen, A.A.; Rittersma, S.; Vaartjes, I.; Stronks, K.; Bots, M.L.; Agyemang, C. Are there ethnic inequalities in revascularisation procedure rate after an ST-elevation myocardial infarction? *PLoS ONE* **2015**, *10*, e0136415. [CrossRef] [PubMed]
38. Bansal, N.; Fischbacher, C.M.; Bhopal, R.S.; Brown, H.; Steiner, M.F.; Capewell, S.; Scottish Health and Ethnicity Linkage Study. Myocardial infarction incidence and survival by ethnic group: Scottish Health and Ethnicity Linkage retrospective cohort study. *BMJ Open* **2013**, *3*, e003415. [CrossRef] [PubMed]
39. Freund, K.M.; Jacobs, A.K.; Pechacek, J.A.; White, H.F.; Ash, A.S. Disparities by race, ethnicity, and sex in treating acute coronary syndromes. *J. Womens Health* **2012**, *21*, 126–132. [CrossRef] [PubMed]
40. Iribarren, C.; Tolstykh, I.; Somkin, C.P.; Ackerson, L.M.; Brown, T.T.; Scheffler, R.; Syme, L.; Kawachi, I. Sex and racial/ethnic disparities in outcomes after acute myocardial infarction: A cohort study among members of a large integrated health care delivery system in northern California. *Arch. Intern. Med.* **2005**, *165*, 2105–2113. [CrossRef] [PubMed]
41. Thomas, K.L.; Honeycutt, E.; Shaw, L.K.; Peterson, E.D. Racial differences in long-term survival among patients with coronary artery disease. *Am. Heart J.* **2010**, *160*, 744–751. [CrossRef] [PubMed]
42. Buja, A.; Canavese, D.; Furlan, P.; Lago, L.; Saia, M.; Baldo, V. Are hospital process quality indicators influenced by socio-demographic health determinants. *Eur. J. Public Health* **2015**, *25*, 759–765. [CrossRef] [PubMed]
43. Srichaiveth, B.; Ruengsakulrach, P.; Visudharom, K.; Sanguanwong, S.; Tangsubutr, W.; Insamian, P. Impact of gender on treatment and clinical outcomes in acute ST elevation myocardial infarction patients in Thailand. *J. Med. Assoc. Thai.* **2007**, *90*, 65–73. [PubMed]
44. Falcone, M.; Del Santo, S.; Forni, S.; Pepe, P.; Marchi, M.; Rossi, G. Equity of access to percutaneous transluminal coronary angioplasty (PTCA) among patients with acute myocardial infarction in Tuscany Region (Central Italy), 2001–2008. *Epidemiol. Prev.* **2013**, *37*, 386–395. (In Italian) [PubMed]
45. Pilgrim, T.; Heg, D.; Tal, K.; Erne, P.; Radovanovic, D.; Windecker, S.; Jüni, P.; AMIS Plus Investigators. Age- and gender-related disparities in primary percutaneous coronary interventions for acute ST-segment elevation myocardial infarction. *PLoS ONE* **2015**, *10*, e0137047. [CrossRef] [PubMed]
46. Blair, I.V.; Steiner, J.F.; Havranek, E.P. Unconscious (implicit) bias and health disparities: Where do we go from here? *Perm. J.* **2011**, *15*, 71–78. [PubMed]
47. Schulman, K.A.; Berlin, J.A.; Harless, W.; Kerner, J.F.; Sistrunk, S.; Gersh, B.J.; Dubé, R.; Taleghani, C.K.; Burke, J.E.; Williams, S.; et al. The effect of race and sex on physicians' recommendations for cardiac catheterization. *N. Engl. J. Med.* **1999**, *340*, 618–626. [CrossRef] [PubMed]

48. Peterson, E.D.; Shah, B.R.; Parsons, L.; Pollack, CV.; French, W.J.; Canto, J.G.; Gibson, C.M.; Rogers, W.J. Trends in quality of care for patients with acute myocardial infarction in the National Registry of Myocardial Infarction from 1990 to 2006. *Am. Heart J.* **2008**, *156*, 1045–1055. [CrossRef] [PubMed]
49. Wang, T.Y.; Dai, D.; Hernandez, A.F.; Bhatt, D.L.; Heidenreich, P.A.; Fonarow, G.C.; Peterson, E.D. The importance of consistent, high-quality acute myocardial infarction and heart failure care results from the American Heart Association's Get with the Guidelines Program. *J. Am. Coll. Cardiol.* **2011**, *58*, 637–644. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

Diagnosis on Transport Risk Based on a Combined Assessment of Road Accidents and Watershed Vulnerability to Spills of Hazardous Substances

Emerson Ribeiro Machado [1], Renato Farias do Valle Junior [1], Teresa Cristina Tarlé Pissarra [2], Hygor Evangelista Siqueira [2], Luís Filipe Sanches Fernandes [3] and Fernando António Leal Pacheco [4,*]

1 Laboratório de Geoprocessamento, Instituto Federal do Triângulo Mineiro, Campus Uberaba, Uberaba MG 38064-790, Brasil; emersonrimac@oi.com.br (E.R.M.); renato@iftm.edu.br (R.F.d.V.J.)
2 Departamento de Engenharia Rural, Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Jaboticabal SP 14884-900, Brasil; teresap1204@gmail.com (T.C.T.P.); hygorsiqueira@yahoo.com.br (H.E.S.)
3 Centro de Investigação e Tecnologias Agroambientais e Biológicas, Universidade de Trás-os-Montes e Alto Douro, Ap 1013, 5001-801 Vila Real, Portugal; lfilipe@utad.pt
4 Centro de Química de Vila Real, Universidade de Trás-os-Montes e Alto Douro, Ap 1013, 5001-801 Vila Real, Portugal
* Correspondence: fpacheco@utad.pt

Received: 13 August 2018; Accepted: 9 September 2018; Published: 14 September 2018

Abstract: Roads play an important role in the economic development of cities and regions, but the transport of cargo along highways may represent a serious environmental problem because a large portion of transported goods is composed of dangerous products. In this context, the development and validation of risk management tools becomes extremely important to support the decision-making of people and agencies responsible for the management of road enterprises. In the present study, a method for determination of environmental vulnerability to road spills of hazardous substances is coupled with accident occurrence data in a highway, with the purpose to achieve a diagnosis on soil and water contamination risk and propose prevention measures and emergency alerts. The data on accident occurrences involving hazardous and potentially harmful products refer to the highway BR 050, namely the segment between the Brazilian municipalities of Uberaba and Uberlândia. The results show that many accidents occurred where vulnerability is high, especially in the southern sector of the segment, justifying the implementation of prevention and alert systems. The coupling of vulnerability and road accident data in a geographic information system proved efficient in the preparation of quick risk management maps, which are essential for alert systems and immediate environmental protection. Overall, the present study contributes with an example on how the management of risk can be conducted in practice when the transport of dangerous substances along roads is the focus problem.

Keywords: environmental vulnerability; multi-criteria spatial analysis; risk management tool; hazardous substance; road accidents

1. Introduction

According to data from the National Confederation of Transport (http://www.cnt.org.br/), in Brazil about 61% of cargo and 95% of passengers are transported via highways, indicating that this transportation sector is economically relevant at the national scale, handling a substantial amount of financial resources [1]. Despite the recognized importance of highways for the country's economic development, road transport involves a diversity of cargo types, including dangerous products, which renders this activity a preoccupying risk of environmental and social impacts. The risk associated to the transport of a dangerous product depends not only on the substance being transported but also on the road network characteristics, weather conditions, driver skills and population concentration along the selected routes [2]. The potential impacts of accidents related to road transport of hazardous products were discussed in various studies, which triggered social concerns and were helpful as guides to the conception and implementation of adequate corrective and preventive actions [3]. However, before implementation of any specific measure, a correct diagnosis on vulnerability and risk is required.

The Multi Criteria Analysis (MCA) method embedded in a geographic information system (GIS), also known as Spatial MCA, is a computational tool that can assist the preparation of vulnerability and risk maps for decision making on the planning and operation road transport networks [4–6], as well as in many other environmental applications [7–12]. In the study of [5] MCA was used to determine the vulnerability of soil and water resources in rural basins, while in the study of [6] it was used to model the environmental risk of accidents involving the transport of dangerous products based on static and dynamic data. The work group of Van [13] developed a method to evaluate statistical data on road accidents involving dangerous products, from which it was possible to generate a global risk map. Regardless the method on which the vulnerability, the risk or the hazards are assessed, from a safety standpoint road risk management requires the implementation of methods that are capable to provide a quick diagnosis on the potential environmental impact of an accident involving the spill and leaching of dangerous substances from a road. The challenge is therefore to develop and validate robust but expeditious diagnostic tools. The resort to statistical data on road accidents demands a significant time span dedicated to monitoring, while the generation of vulnerability maps cannot stand alone as method to identify the risky areas [2]. According to [14], besides the potential lack of reliable numbers the management of road transport risk based on statistical data requires the capability to evaluate, in a short period of time, the diversity of transported materials, and the possible environmental consequences related to their road spill. This evaluation becomes even more complex given the multiplicity of road accident circumstances. The route to follow is therefore to combine hazard assessments (road accident counts) with vulnerability assessments at site and catchment scales.

The main purpose of this study is to combine a method already used to determine the environmental vulnerability of areas adjacent roads to spills of dangerous products [15] with road accident data, in a manner that becomes possible to analyze the accident scenario immediately after the occurrence and thus to implement an alert system whereby corrective measures can readily be triggered, such as the sending of resources, isolation of a certain area, withdrawal of the population, protection of springs, soil or water decontamination, among others.

2. Area of Study

The area where the risk management method is to be implemented is located between the municipalities of Uberaba and Uberlândia, in the Brazilian State of Minas Gerais, mesoregion of Triângulo Mineiro. The municipality of Uberlândia is the second most populous in Minas Gerais, and the 30th in Brazil, with 676,613 inhabitants according to the Brazilian Institute of Geography and Statistics—IBGE [16]. The municipality of Uberaba has a population of 328,272 inhabitants, as estimated in 2017, being the 8th largest municipality in the state and the 81st in the country [17]. The sector under study comprises a 97-km segment of highway BR 050, from km 77 at the junction of BR 050 to kilometer 174 at the intersection with BR 262. The BR 050 is an important connecting corridor between the central-west and southeast regions of Brazil, receiving the flow of several highways that cross the country from north to south, namely the traffic from the Federal Highways BR 365/452/455/497. The BR 050 highway was built in the 1970s with the main purpose to connect the capital of the country to the Port of Santos SP. In 2010 the segment under study was doubled, increasing the highway capacity.

Along the studied segment the BR 050 highway intersects several watercourses that drain small watersheds (Figure 1). The climate of Triângulo Mineiro region is qualified as Aw according to the Köeppen classification. The Aw climate is a tropical mega thermal climate, with winter droughts and an average temperature for the coldest month around 18 °C. Precipitation is characterized by annual averages of 1200 to 1450 mm, according to the climatic norm of Brazil (1961–1990) published by the National Institute of Meteorology [18]. The dry period runs from May to September and the wet period from October on [19]. As regards geomorphology, the Triângulo Mineiro is located in the so-called "Plateaus and Mesas of the Paraná Sedimentary Basin", which comprise the "Northern Plateau sub-unit" [20]. Geologically, this plateau is characterized by deposits of Uberaba, Marília and Vale do Rio do Peixe formations belonging to the Bauru group; Serra Geral formation belonging to the São Bento group; and undifferentiated dendritic and/or lateritic [21].

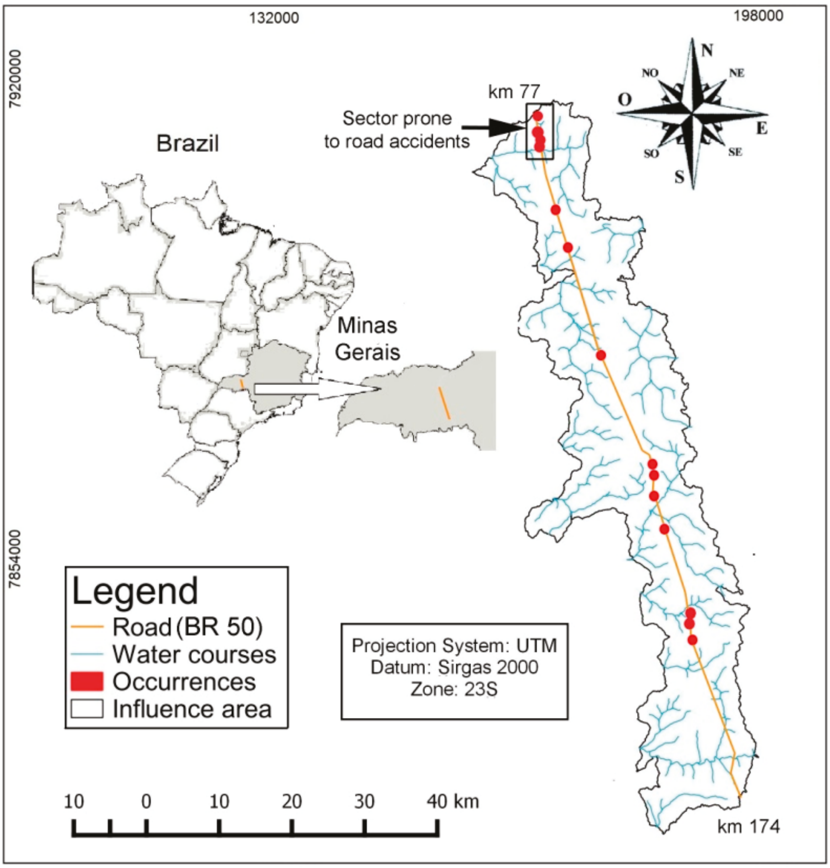


Figure 1. Geographic location of the studied BR 050 highway segment, in Brazil and Minas Gerais State, with representation of intercepted water courses and distribution of road accidents involving spills of hazardous substances. In the northern and central parts of this segment the road was built nearly along a water divide. In these sectors the water channels are likely to be equally vulnerable to contamination at both sides of the road, because the spill of a harmful substance will potentially leach in both directions. For similar reasons, in the southern part the water channels from the west side are potentially more vulnerable than the channels from the east side.

3. Materials and Methods

The method to investigate road accident scenarios based on the assessment of environmental vulnerability and road accident data related to transport of hazardous substances is illustrated in

Figure 2 in the form of a workflow. The Multi Criteria Analysis represented in this diagram and used to assess environmental vulnerability has been developed and presented in the previous work of [15]. The coupling of those results with road accident data is performed in this study. The method developed in the earlier work of Machado and co-authors is briefly described in Section 3.1. The complement related to analysis of road accident data is described in detail in Sections 3.2 and 3.3. Both methods were implemented in Geographic Information System (GIS), frequently used in environmental studies [22–34]. In the present study, the specific GIS was IDRISI Selva software [35] that resorted to various sources of digital information, mostly public institution websites (e.g., http://www.webmapit.com.br/inpe/topodata/ for topographic data or https://earthexplorer.usgs.gov/ for land use/land occupation data). The full inventory of information sources is listed in [15].

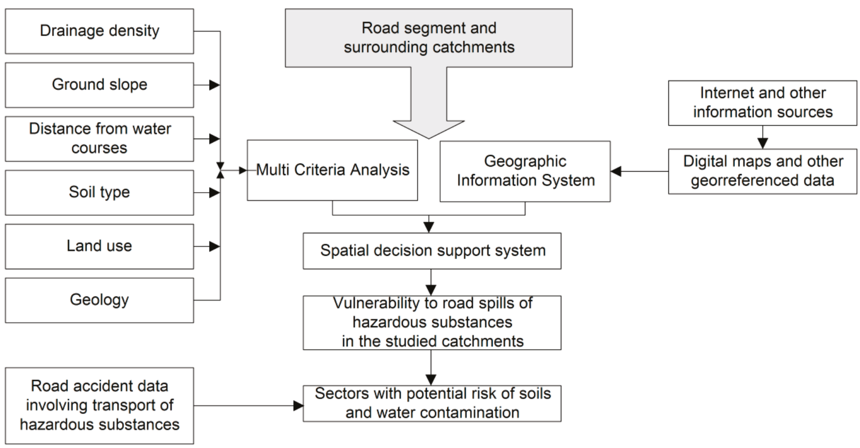


Figure 2. Flowchart illustrating how contamination risk has been assessed in the present study, associated with spills of hazardous substances following a road accident.

3.1. Determination of Vulnerability

As mentioned, an environmental vulnerability evaluation along the BR 050 highway was accomplished by the authors of this study in a previous publication [15]. To complete the task, a Multi Criteria Analysis (MCA) embedded in a geographic information system was applied to the areas under direct influence (within a 210 m buffer from each margin of the highway) and indirect influence (within the limits of the micro basins along the 97 km segment) of spills of dangerous products. The MCA approach is a four-step process, which involves (1) selection of factors to describe vulnerability, with subsequent normalization of factor scales into a common dimensionless range, and elaboration of raster maps that describe the spatial distribution of normalized factors; (2) the allocation of a weight to each factor; (3) the weighted combination of factors to compose a final vulnerability map; and (4) a sensitivity analysis of vulnerability results based on scenarios [36]. The four steps are briefly outlined in the next paragraphs:

Step 1: In the study of [15] the vulnerability maps were based on the following factors: (1) drainage density, (2) distance from water courses, (3) ground slope, (4) soil type, (5) land use/occupation, and (6) geology (Figure 3; Table 1, part (a)). These factors were selected because they are comparable to key variables of drainage models and flow routing algorithms that describe the detachment and transportation of pollutants in catchments [5]. The normalization of factor classes was based on a byte-level interval (0 to 255), which linked a 0 level to the least important class and 255 level to the most important class (Table 1, part (b)). The association of factor classes to levels of importance (ratings) was based on the authors' personal experiences about vulnerability assessments.

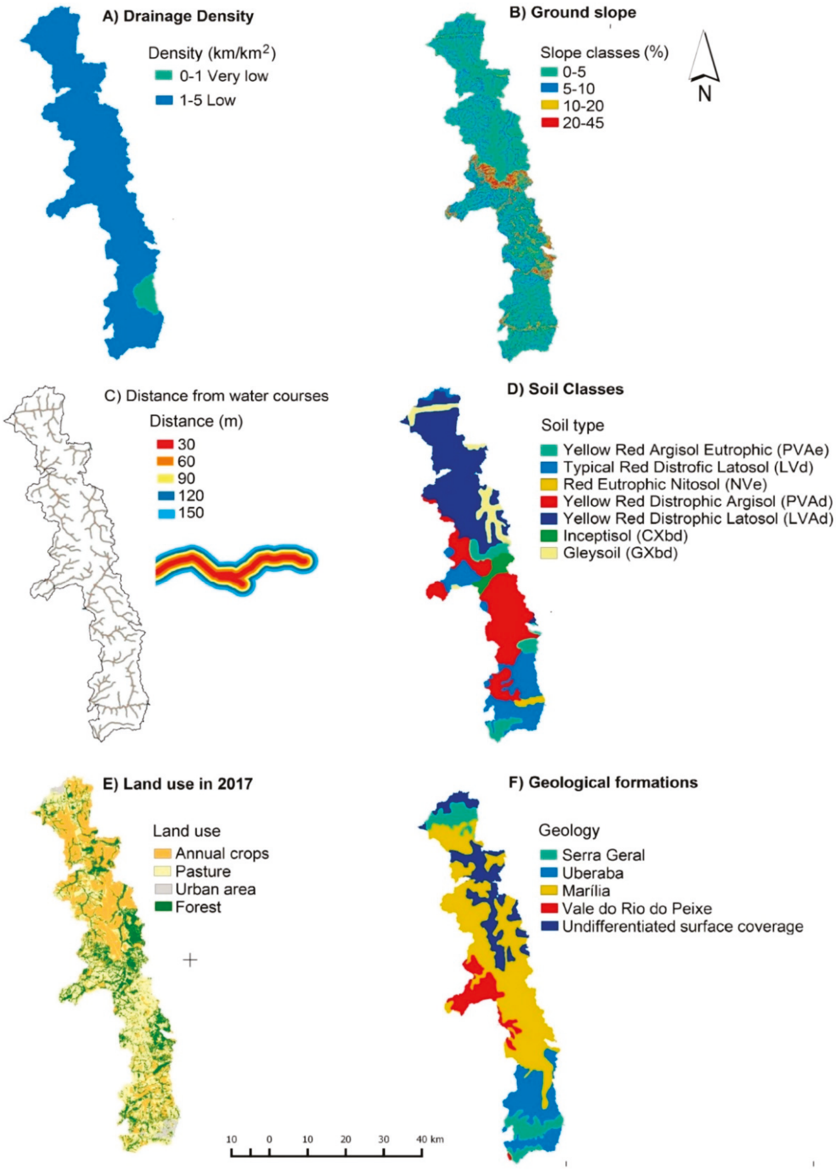


Figure 3. Spatial distribution of vulnerability-relevant factors included in the Multi Criteria Analysis. Adapted from [15].

Table 1. (**a**) Factors used by [15] in the Multi Criteria Analysis of environmental vulnerability related to road accidents along the studied segment of BR 050 highway involving the transport of hazardous substances; (**b**) Normalization of factors within the Multi Criteria Analysis—MCA (step 2) designed to evaluate soil and water vulnerability along roads. The higher the value of a normalized factor the greater its importance for vulnerability. The MCA model was applied to a segment of BR 050 highway where transport of hazardous substances is intense and spills of those products following a road accident can cause severe damage to the surrounding environment. Adapted from [15].

(a)

Environmental Factors	Accident Scenario Implications
Ground slope	Factor that describes important aspects related to the control of erosion, transport of sediments and contaminants.
Drainage density/distance from water courses	It describes factors related to the likelihood of water resources and biotic environment contamination.
Geology	Factors related to likelihood of contamination, socioeconomic impact and the extent of damage in accident scenarios.
Soil Classes/land use or occupation	It exposes factors related to the likelihood of soil and groundwater contamination and contaminant movement in accident scenarios.

(b)

Factors	Values	Normalized Values
Drainage density ($km \cdot km^{-2}$)	Very low	0–1
	Low	1–5
	Medium	5–13
	High	13–15
	Very high	>15
Distance of water course (m)	30	255
	60	175
	90	115
	120	75
	150	50
Ground slope (%)	0 a 5%	25
	5 a10%	75
	10 a 20%	125
	20 a 45%	255
Soil classes	Latosol	100
	Acrisol	150
	Nitisol	180
	Gleysol	200
	Cambisol	250
Land use and occupation	Annual crops	75
	Pasture	125
	Forest	200
	Urban Area	255
	Undifferentiated surface coverage	50
	Serra Geral	100
	Vale do Rio do Peixe	150
	Marília	200
	Uberaba	255

Step 2: The allocation of weights was based on the Analytical Hierarchy Process (AHP; [37]) whereby the user (or a group of experts) assign a relative importance to each factor based on pairwise comparisons with the other factors, and then this hierarchy is processed in the AHP algorithm to obtain a set of optimized levels of importance (weights). Because the attribution of weights can be subjective, in the study of [15] a sensitivity analysis was performed (step 4 below) where various factors were given the largest relative importance and hence maximum weights.

Step 3: The overall vulnerability was calculated by Equation (1), implemented in the GIS software using map algebra tools and factor maps in raster format (Figure 4):

$$S_i = \sum_{i=1}^{p} w_j X_{ij} \tag{1}$$

where S_j represents the vulnerability at pixel i, w_j represents the weight of the factor j, and X_{ij} represents the normalized value of factor j at pixel i. The S_i values are reclassified into five classes using the same byte-level range: Invulnerable (0–50), Weakly vulnerable (50–100), Vulnerable (100–150), Strongly Vulnerable (150–200), and Extremely Vulnerable (200–250).

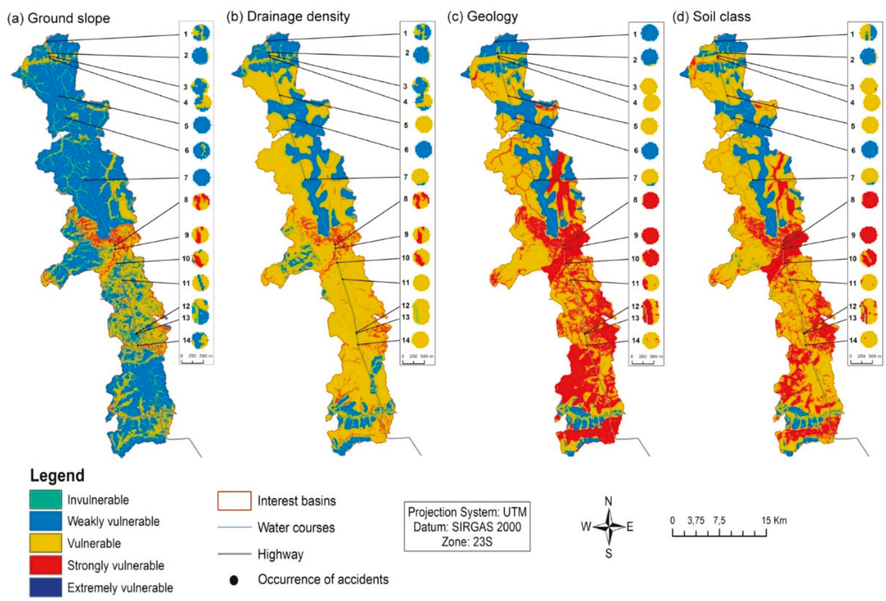


Figure 4. Vulnerability maps of the intercepted water course catchments, highlighting the vulnerability at the road accident sites (also termed hazard; labeled circles). The maps are outcomes of a Multi Criteria Analysis where vulnerability-relevant factors ground slope (map (**a**)), drainage density (**b**), geology (**c**) and soil type (**d**) were given the largest weight [15]. The concomitant effects on vulnerability are reflexes of factor heterogeneity across the studied area. The largest effect occurs when factors geology or soil type are maximized highlighting the importance of these factors.

Step 4: The sensitivity of S to changing factor weights was evaluated through generation of four scenarios where one of these factors has been given the largest relative importance maximizing its weight. The factors that have been given maximum weights were drainage density, ground slope, soil type and geology. The scenarios were created because the aforementioned factors are heterogeneous across the studied region and in that context associated with an ample range of scores. For these reasons there is no easy way to define a universal hierarchy to describe the importance of each factor. For example, the studied segment of BR 050 highway is contrasting as regards ground slope, because the north and south sectors are occupied by plains linked to low vulnerability while the central part is mountainous and linked to high vulnerability. In a scenario that maximizes the importance of ground slope, these contrasting topographic features will be highlighted in the final vulnerability map, while being smoothed otherwise. The same rationale holds for the other factors as well. Vulnerability in the four scenarios is illustrated in Figure 4a–d.

As expected, when factor ground slope is maximized (Figure 4a) the vulnerability map shows a central area with high vulnerability bordered to the north and south by areas with low vulnerability. However, it is evident from analysis of Figure 4c,d that rising the role of geology or soils in the vulnerability assessment results in larger overall vulnerability. It is also worth to note that the BR 050 highway in the studied segment was built nearly along a water divide in the northern and central sectors but away from it in the southern part (Figure 1). For that reason, spills of dangerous substances in the northern-central parts of the segment will potentially affect the water courses in both sides of the highway while in the southern part spill drainage will primarily affect the western channels.

3.2. Occurrence Data Involving Hazardous and Potentially Harmful Products to the Environment

The data on accidents involving hazardous and potentially harmful products in the studied segment was obtained from the Concessionaire who manages approximately 700 km of BR 050 highway. The collected data is summarized in Table 2. It is important to note that some products represented in the table are not classified as dangerous by the United Nations (http://www.unece.org/trans/danger). These products are identified as "not applicable" under the heading "UN Code" (Column 3). However, because the road spill of these products can contribute significantly to soil and water contamination in accident scenarios, they were used in the present study of risk management. The accident data was compiled from operational resources such as Operational Control Center (OCC), Traffic Inspection and Mechanical Rescue Vehicles, Emergency Medical Service Vehicles (rescue and salvage), Closed Circuit TV and Radio Communication System. Through the OCC, all the occurrence data are recorded using the software Kria Operational Control for Highways. Among other issues, the recording involves the generation of a GIS database through the conversion of site details (i.e., the exact kilometer of the occurrence) into geographical coordinates (last columns of Table 2). Besides generation of data the software releases management reports according to the periodicity and type of occurrence desired, allowing analysis, treatment and decision making. The data for the present study spans the period from July 2014 to December 2017, which represents a 2 years and 6 months interval. Figure 1 shows the distribution of occurrences (red circles) involving hazardous products in the studied segment, obtained through the Concessionaire. We recognize that the data record is not long to provide a clear image of the situation, but are confident that enables a preliminary view.

Table 2. Occurrences involving dangerous products in BR 050 during the monitored period. Symbols: UN—United Nations; UTM—Universal Transverse Mercator (coordinate system); X, Y—planimetric coordinates of the accident.

Date	Product	UN Code	Time	Kilometer	UTM—Zone 23 S	
					X	Y
09/29/14	Diesel oil	1202	13:04:00	082 + 180	161,681	7,898,073
10/09/14	Ethanol	1170	15:07:00	149 + 500	181,617	7,834,447
10/30/14	Diesel oil	1202	21:32:00	096 + 500	165,445	7,884,280
01/23/15	Toluene	1294	02:21:00	111 + 500	169,793	7,869,927
08/31/15	Ethanol	1170	15:52:00	152 + 120	181,988	7,831,857
09/12/15	GLP	1075	05:35:00	078 + 340	161,054	7,901,861
10/29/15	Oil S10	1202	11:30:00	081 + 800	161,619	7,898,448
10/14/16	Hydrated alcohol	1170	06:04:00	091 + 200	163,904	7,889,351
02/10/17	Diesel oil	1202	12:37:00	129 + 100	176,754	7,854,079
07/17/17	Hydrochloric acid	1789	06:31:00	149 + 300	181,607	7,834,653
10/24/17	Vegetable oil	not applicable	10:40:00	136 + 600	178,378	7,846,904
10/25/17	Limestone	not applicable	08:45:00	132 + 540	177,162	7,850,754
11/13/17	Cement	not applicable	18:09:00	081 + 100	161,494	7,899,147
11/22/17	Kerozene	1223	23:01:00	128 + 300	176,647	7,854,871

3.3. Environmental Vulnerability at Occurrence Sites (Risk)

The vulnerability around the road accident sites listed in Table 2 was assessed by the IDRISI Selva software [35], taking into account the four predefined scenarios (Figure 4). The IDRISI Selva software embeds a set of tools capable to determine the environmental vulnerability of an area according to the necessary steps. The vulnerability profile of each site was defined through the following steps: (1) a 200 m buffer was drawn around the site. This area covers the environmental resources immediately affected after the occurrence of spills; (2) The buffers were plotted over the four vulnerability maps (Figure 4a–d); (3) For each vulnerability scenario, the area related to a vulnerability level (e.g., strongly vulnerable) was determined using raster map operations.

Steps 1–3 were repeated for all the vulnerability levels and all sites, and aggregated per vulnerability level in each scenario. To distinguish the vulnerability evaluated within the studied segment (catchment scale) from the vulnerability evaluated around the road accident sites (buffer scale) the latter was termed hazard. Having determined the hazard area within the 14 buffers, the risk of soil and water contamination is estimated for every vulnerability level using the formula:

$$R_j = \frac{H_j}{V_j} = \frac{Ab_j}{A_j} \tag{2}$$

where R_j is the risk for level j, H_j is the hazard for level j evaluated within the 14 buffers and represented by the corresponding area (Ab_j, in percentage of total buffer area), and V_j is the vulnerability for level j evaluated within the studied segment and represented by its area (A_j, in percentage of segment area). If the R_j value is >1 then road accident sites are considered risky at that level. If sites are risky for the preoccupying levels (e.g., "strongly vulnerable" or "extremely vulnerable") then the implementation of prevention and alert systems should be mandatory. These systems should also be considered for the "vulnerable level". The analysis of risk can be refined, which means executed site by site. In this case, the Ab_j represents the area of hazard level j within the specific site, in percentage of buffer area. It is worth mentioning that, besides hazard incidence and medium vulnerability the risk of soil and water contamination by dangerous substance, including public health issues, also depends on the extension of contaminant propagation, the amount and chemical properties (toxicity) of the spilled product, and the proximity of human presence [38]. Toxicity and proximity to urban centers will not be addressed in this study, because the vulnerability assessment on which the risk analysis is standing has been focused on the protection of environmental resources, soils and water.

4. Results and Discussion

4.1. Accident Count over the Monitored Period

A total of 14 accidents were reported to the monitoring system within the studied period (Table 2). Among these occurrences, 11 involved the transport of hazardous products and three the transport of products potentially harmful to the environment such as vegetable oil, limestone and cement. The largest number of episodes occurred in the morning (from 06:00 to 12:00, 35.71% of occurrences), being followed by the afternoon (from 12:00 to 18:00, 28.57%), night time (from 18:00 to 00:00, 21.43%), and early morning (from 00:00 to 06:00, 14.29%). As regards seasonality, there were nine occurrences in the rainy season (64.29%) and five in the period of low or no rainfall. The results also show that one quarter (28.57%) of all accidents occurred in the section between km 77 and 83 of BR 050. This section is characterized by steep slopes, but paradoxically is also distinguished by fast-speed traffic. Figure 5, generated in Google Earth, shows the elevation profile of km 77–km 83 section, with an elevation of 932 m at km 77 and 810 m at km 83 (2% slope, on average). Ferreira [39] studied the causes of road accidents with hazardous products in the State of São Paulo, Brazil, and observed a greater predominance of accidents in the afternoon (between 12:00 and 18:00) and that the routes between petrochemical poles also influence the number of accidents, due to the greater flow of vehicles transporting dangerous products along these routes. Overall, the results obtained in this study as well as by other authors [2] demonstrate the heterogeneity of factors influencing the incidence of accidents with hazardous products, which hampers the selection of priority sections for hazard management. The alternative path to follow relies on combining the assessment of hazard distribution with environmental vulnerability assessments, and hence moving from a conventional hazard management to the more integrated approach that is risk management. The assemblage of hazard and vulnerability data into a common framework of risk data, especially if using a geographic information system to accommodate and process the maps and associated attribute tables, has the additional virtue to help finding priority sections for management, because risky areas are fewer and smaller than the sum of hazard and vulnerable areas.

Figure 5. Elevation profile of the BR 050 highway segment involved in a large number of road accidents. This segment is located between km 77 and km 83 of the highway, as illustrated in Figure 1. The elevation profile was generated using the Google Earth software. The accidents are mostly caused by fast-speed traffic in a relatively steep-slope road.

4.2. Vulnerability, Hazard and Risk

The 200 m-buffers around the 14 accidents sum a hazard area of approximately 168.27 hectares. The buffers were defined because large areas around the accident sites can be affected by the road spills, even extending to the entire watershed. Table 3 summarizes the results obtained for vulnerability, hazard and risk in the four predefined scenarios of factor maximization, respectively expressed as vulnerable areas within the BR 050 segment watersheds (V), hazard areas within the 200 m buffers surrounding the accident sites (H) and the percent ratio between H and V ($R = H(\%)/V(\%)$). The V areas were evaluated within the maps of Figure 4a–d, while the H areas were measured within the buffer areas (labeled circles) represented in the same figures. For example, for vulnerability level "strongly vulnerable" (red color in Figure 4) the V area is 4337.79 hectares (3.4% of road segment area; Table 3) in the ground slope scenario (Figure 4a), while raises to 40,073.87 hectares (31.4%) in the geology scenario (Figure 4c). The corresponding H areas are 11.50 hectares (6.83%) and 48.75 hectares (28.97%). According to Equation (2) this gives a risk $R = 2.01$ for the ground slope scenario and $R = 0.92$ in the geology scenario. In case ground slope is adopted as reference scenario for decision making on soil and water protection, then road accident sites located where the environment is strongly vulnerable are considered risky because $R > 1$. In general, as regards vulnerability the areas were mostly classified as weakly vulnerable or vulnerable, for the ground slope and drainage density factors, and as vulnerable or strongly vulnerable for the soil class and geology factors. The coverage by extremely vulnerable areas or invulnerable areas was insignificant. The results obtained for the areas where the accidents have occurred (hazards) follow the general results obtained for vulnerability, because the percentage of area ascribed to the vulnerability classes are similar in both cases. The exceptions occur for the scenarios where ground slope or drainage density factors were maximized, because in some cases the areas where the accidents have occurred are more vulnerable than the general vulnerability areas in those scenarios. As mentioned above, for the scenario that maximized ground slope the areas classified as strongly vulnerable along the highway watersheds represent $V = 3.4\%$ of the total watershed area while the homologous areas around the accident sites represent $H = 6.83\%$. The same holds for the strongly vulnerable areas in the scenario that maximized the drainage density factor, which rise from $V = 4.06\%$ to $H = 6.83\%$. Put another way, the strongly vulnerable areas in these two scenarios can be classified as risky, because $R = H/V > 1$ in both cases, namely 2 and 1.7 (Equation (1)). In that context, these vulnerability levels and corresponding areas of influence would deserve special attention in risk management plans.

Table 3. Vulnerability assessments within the watersheds that surround the studied segment of BR 050 highway (*V*), considering the four scenarios. Vulnerability assessments within the 200 m buffers that surround the 14 road accidents (also termed hazard assessments; *H*), considering the same scenarios. Risk assessments (*R* = *H*/*V*, in percent ratio).

Intercepted Basins [15]			Buffers Around Road Accident Sites		
Scenario 1—Maximize ground slope factor					
Category	**Vulnerability**		**Hazard**		**Risk—*R***
	Area—*V* (hectare)	**%**	**Area—*H* (hectare)**	**%**	
Invulnerable	1425.79	1.12	1.33	0.79	0.71
Weakly vulnerable	79,725.12	62.46	85.90	51.05	0.82
Vulnerable	42,135.85	33.01	69.54	41.32	1.25
Strongly vulnerable	4337.79	3.40	11.50	6.83	2.01
Extremely vulnerable	17.07	0.01	0.00	0.00	0.00
Total	127,641.62	100.00	168.27	100	
Scenario 2—Maximize drainage density factor					
Category	**Vulnerability**		**Hazard**		**Risk—*R***
	Area—*V* (hectare)	**%**	**Area—*H* (hectare)**	**%**	
Invulnerable	1416.06	1.11	1.33	0.79	0.7
Weakly vulnerable	31,358.31	24.57	44.59	26.50	1.1
Vulnerable	89,678.65	70.26	110.85	65.88	0.9
Strongly vulnerable	5188.61	4.06	11.50	6.83	1.7
Extremely vulnerable	0.00	0.00	0.00	0.00	nd
Total	127,641.62	100	168.27	100.00	
Scenario 3—Maximize geology factor					
Category	**Vulnerability**		**Hazard**		**Risk—*R***
	Area—*V* (hectare)	**%**	**Area—*H* (hectare)**	**%**	
Invulnerable	1421.19	1.11	1.33	0.79	0.71
Weakly vulnerable	25,391.51	19.89	36.89	21.92	1.10
Vulnerable	60,624.47	47.50	81.22	48.26	1.02
Strongly vulnerable	40,073.87	31.40	48.75	28.97	0.92
Extremely vulnerable	130.58	0.10	0.09	0.05	0.50
Total	127,641.62	100	168.27	100	
Scenario 4—Maximize soil class factor					
Category	**Vulnerability**		**Hazard**		**Risk—*R***
	Area—*V* (hectare)	**%**	**Area—*H* (hectare)**	**%**	
Invulnerable	1416.06	1.11	1.33	0.79	0.71
Weakly vulnerable	22,612.48	17.72	28.58	16.98	0.96
Vulnerable	77,085.70	60.39	100.15	59.52	0.99
Strongly vulnerable	26,190.75	20.52	37.95	22.56	1.10
Extremely vulnerable	336.63	0.26	0.27	0.16	0.62
Total	127,641.62	100	168.27	100	

Figure 4a–d display the vulnerability maps obtained at the 14 sites where the accidents occurred during the studied period. The maps also represent the surrounding watersheds, because they can also be environmentally affected. In all cases, these figures provide visual information for rapid environmental risk assessment of accident sites, enabling immediate prevention and alerts for the sites classified as vulnerable or strongly vulnerable. The representation of accident sites in a color scale related to environmental vulnerability validates the method of [15] as support for an expeditious risk management tool, and hence represents the achievement of a proposed objective. In that context, it is important to note the large number of accident sites located in strongly vulnerable areas, especially in the southern sector watersheds and when the focus of vulnerability is put on the catchments' soil and geologic characteristics. An environmental alert is due in these cases to ensure the safety of soil and water quality within the involved watersheds.

5. Conclusions

The management of road transport risk involving hazardous substances is a challenging exercise because the factors influencing this variable depend on the vulnerability of the medium as well as on a myriad of road accident causes and circumstances. The challenge also results from the fact that, to be effective, a risk management tool needs to release the relevant information on local vulnerability immediately after the occurrence of an accident that spilled a dangerous product over the road. In this study, a combination of vulnerability and hazard assessments proved efficient to readily identify risky sections in a segment of highway BR 050 located in Brazil, which correspond to areas with a larger incidence of accidents located on strongly vulnerable areas. In these areas, the risk is mostly determined by ground slope and drainage density. The study could also depict accident sites and associated influence buffers as colored circles to readily represent vulnerability at site scale. It became evident at first sight that the southern sector of the highway BR 050 segment requires closer attention as regards environmental risk. In case of accident, soil and water contamination is highly probable because soils and geological formations are barely capable to sustain the propagation of contaminants in that sector. Overall, the study proved efficient in providing a comprehensive diagnosis on contamination risk along the studied road, as well as in providing clues about sectors of the highway requiring particular attention from risk managers.

Author Contributions: E.R.M. contributed to resources, data curation, investigation, conceptualization, methodology, and writing original draft. R.F.d.V.J. and T.C.T.P. contributed to formal analysis, supervision, project administration and funding acquisition. H.E.S. contributed to software and visualization. L.F.S.F. contributed to formal analysis, validation, writing-review, editing, and funding acquisition, F.A.L.P. contributed to funding acquisition, methodology, supervision, and writing-review editing.

Funding: As regards the Portuguese authors, the research was funded by national funds (FCT-Portuguese Foundation for Science and Technology) under the strategic project of the Vila Real Chemistry Research Center (PEst-OE/QUI/UI0616) and the CITAB (Centre for the Research and Technology of Agro-Environmental and Biological Sciences) project UID/AGR/04033.

Acknowledgments: The authors would like to thank the Federal Institute of the Triângulo Mineiro (IFTM) for their support with infrastructure and software for the development of the research, and the road company MGO Rodovias for the availability of data. The authors also wish to thank the University of Trás-Montes and Alto Douro (UTAD), the Chemistry Research Center and the Center for the Research and Technology of Agro-Enviromental and Biological Sciences (CITAB), for technical support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. CNT. *Pesquisa CNT de Rodovias 2015: Relatório Gerencial*; CNT: Brasília DF, Brasíl, 2015; p. 420.
2. Torretta, V.; Rada, E.C.; Schiavon, M.; Viotti, P. Decision support systems for assessing risks involved in transporting hazardous materials: A review. *Saf. Sci.* **2017**, *92*, 1–9. [CrossRef]

3. Tinoco, M.A.C.; Nodari, C.T.; Pereira, K.R.S. *Vulnerabilidade Ambiental, Social e viáRia em Acidentes com Transporte de Produtos Perigosos: Estudo de caso na BR-101 entre Osório e Torres, Rio Grande do Sul, Brasil*; Caderno de Saúde Pública: Rio de Janeiro, Brazil, 2016; Volume 32, p. 9.
4. Siqueira, H.E. Vulnerabilidade Ambiental dos Recursos Hídricos Interceptados por Rodovias na Bacia do rio Uberaba. Master's Thesis, Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, São Paulo, Brazil, 2016.
5. Siqueira, H.E.; Pissarra, T.C.T.; Valle Junior, R.F.; Fernandes, L.F.S.; Pacheco, F.A.L.A. Multi criteria analog model for assessing the vulnerability of rural catchments to road spills of hazardous substances. *Environ. Impact Assess. Rev.* **2017**, *64*, 26–36. [CrossRef]
6. Cordeiro, F.G.; Bezerra, B.S.; Peixoto, A.S.P.; Ramos, R.A.R. Methodological aspects for modeling the environmental risk of transporting hazardous materials by road. *Transp. Res. D* **2016**, *44*, 105–121. [CrossRef]
7. Bellu, A.; Sanches Fernandes, L.F.; Cortes, R.M.V.; Pacheco, F.A.L. A framework model for the dimensioning and allocation of a detention basin system: The case of a flood-prone mountainous watershed. *J. Hydrol.* **2016**, *533*, 567–580. [CrossRef]
8. Pacheco, F.A.L.; Sanches Fernandes, L.F. The multivariate statistical structure of DRASTIC model. *J. Hydrol.* **2013**, *476*, 442–459. [CrossRef]
9. Pacheco, F.A.L.; Pires, L.M.G.R.; Santos, R.M.B.; Sanches Fernandes, L.F. Factor weighting in DRASTIC modeling. *Sci. Total Environ.* **2015**, *505*, 474–486. [CrossRef] [PubMed]
10. Terêncio, D.P.S.; Sanches Fernandes, L.F.; Cortes, R.M.V.; Pacheco, F.A.L. Improved framework model to allocate optimal rainwater harvesting sites in small watersheds for agro-forestry uses. *J. Hydrol.* **2017**, *550*, 318–330. [CrossRef]
11. Terêncio, D.P.S.; Sanches Fernandes, L.F.; Cortes, R.M.V.; Moura, J.P.; Pacheco, F.A.L. Rainwater harvesting in catchments for agro-forestry uses: A study focused on the balance between sustainability values and storage capacity. *Sci. Total Environ.* **2018**, *613–614*, 1079–1092. [CrossRef] [PubMed]
12. Valle Junior, R.F.; Varandas, S.G.P.; Sanches Fernandes, L.F.; Pacheco, F.A.L. Multi Criteria Analysis for the monitoring of aquifer vulnerability: A scientific tool in environmental policy. *Environ. Sci. Policy* **2015**, *48*, 250–264. [CrossRef]
13. Van Raemdonck, K.; Macharis, C.; Mairesse, O. Risk analysis system for the transport of hazardous materials. *J. Saf. Res.* **2013**, *45*, 55–63. [CrossRef] [PubMed]
14. Bubbico, R.; Maschio, G.; Mazzarottaa, B.; Milazzo, M.F.; Parisia, E. Risk management of road and rail transport of hazardous materials in Sicily. *J. Loss Prev. Process Ind.* **2006**, *19*, 32–38. [CrossRef]
15. Machado, E.R.; Valle Júnior, R.F.; Sanches Fernandes, L.F.; Pacheco, F.A.L. The vulnerability of the environment to spills of dangerous substances on highways: A diagnosis based on Multi Criteria Modeling. *Transp. Res. Part D* **2018**, *62*, 748–759. [CrossRef]
16. Instituto Brasileiro de Geografia e Estatística (IBGE). Available online: https://cidades.ibge.gov.br/brasil/mg/uberlandia/panorama (accessed on 8 September 2018).
17. Instituto Brasileiro de Geografia e Estatística (IBGE). Available online: https://cidades.ibge.gov.br/brasil/mg/uberaba/panorama (accessed on 8 September 2018).
18. INMET. Normais Climatológicas Do Brasil 1961–1990: Precipitação Anual (mm). Available online: http://www.inmet.gov.br/portal/index.php?r=clima/normaisclimatologicas (accessed on 9 March 2018).
19. Costa, F.S.; Silva, J.J.D.; Souza, C.M.D.; Mendes, J. Dinâmica populacional de Aedes aegypti (L) em área urbana de alta incidência de dengue. *Rev. Soc. Bras. Med. Trop.* **2008**, *41*, 309–312. [CrossRef] [PubMed]
20. RADAMBRASIL. *Levantamento de Recursos Naturais*; Folha SE: Rio de Janeiro, Brazil, 1983; Volume 31, p. 22.
21. CODEMIG; CPRM. Mapa Geológico do Estado de Minas Gerais. Escala 1:1.000.000. Edição Atualizada. 2014. Available online: http://www.portalgeologia.com.br/index.php/mapa/ (accessed on 5 June 2017).
22. Pacheco, F.A.L. Regional groundwater flow in hard rocks. *Sci. Total Environ.* **2015**, *506–507*, 182–195. [CrossRef] [PubMed]
23. Pacheco, F.A.L.; Landim, P.M.B. Two-way regionalized classification of multivariate data sets and its application to the assessment of hydrodynamic dispersion. *Math. Geol.* **2005**, *37*, 393–417. [CrossRef]
24. Pacheco, F.A.L.; Van der Weijden, C.H. Weathering of plagioclase across variable flow and solute transport regimes. *J. Hydrol.* **2012**, *420–421*, 46–58. [CrossRef]
25. Pacheco, F.A.L.; Van der Weijden, C.H. Integrating Topography, Hydrology and Rock Structure in Weathering Rate Models of Spring Watersheds. *J. Hydrol.* **2012**, *428–429*, 32–50. [CrossRef]

26. Pacheco, F.A.L.; Van der Weijden, C.H. Role of hydraulic diffusivity in the decrease of weathering rates over time. *J. Hydrol.* **2014**, *512*, 87–106. [CrossRef]
27. Pacheco, F.A.L.; Van der Weijden, C.H. Modeling rock weathering in small watersheds. *J. Hydrol.* **2014**, *513*, 13–27. [CrossRef]
28. Pacheco, F.A.L.; Landim, P.M.B.; Szocs, T. Anthropogenic impacts on mineral weathering: A statistical perspective. *Appl. Geochem.* **2013**, *36*, 34–48. [CrossRef]
29. Pacheco, F.A.L.; Sousa Oliveira, A.; Van der Weijden, A.J.; Van der Weijden, C.H. Weathering, biomass production and groundwater chemistry in an area of dominant anthropogenic influence, the Chaves-Vila Pouca de Aguiar region, north of Portugal. *Water Air Soil Pollut.* **1999**, *115*, 481–512. [CrossRef]
30. Pacheco, F.A.L.; Santos, R.M.B.; Sanches Fernandes, L.F.; Pereira, M.G.; Cortes, R.M.V. Controls and forecasts of nitrate fluxes in forested watersheds: A view over mainland Portugal. *Sci. Total Environ.* **2015**, *537*, 421–440. [CrossRef] [PubMed]
31. Sanches Fernandes, L.F.; Santos, C.; Pereira, A.; Moura, J. Model of management and decision support systems in the distribution of water for consumption: Case study in North Portugal. *Eur. J. Environ. Civ. Eng.* **2011**, *15*, 411–426. [CrossRef]
32. Sanches Fernandes, L.F.; Seixas, F.; Oliveira, P.; Leitao, S.; Moura, J. Climate change impacts on nitrogen in a hydrographical basin in the northeast of Portugal. *Fresenius Environ. Bull.* **2012**, *21*, 3643–3650.
33. Sanches Fernandes, L.F.; Marques, M.J.; Oliveira, P.C.; Moura, J.P. Decision support systems in water resources in the demarcated region of Douro—Case study in Pinhão River Basin, Portugal. *Water Environ. J.* **2014**, *28*, 350–357.
34. Sanches Fernandes, L.F.; Terêncio, D.P.S.; Pacheco, F.A.L. Rainwater harvesting systems for low demanding applications. *Sci. Total Environ.* **2015**, *529*, 91–100. [CrossRef] [PubMed]
35. Eastman, J.R. *IDRISI Selva Manual*; Clark Labs, Clark University: Worcester, MA, USA, 2012; 322p, Provided as a PDF with the IDRISI Selva Software Package; Available online: http://clarklabs.org/products/spanish-manual.cfm (accessed on 11 September 2018).
36. Malczewski, J. *GIS and Multicriteria Decision Analysis*; John Wiley and Sons: New York, NY, USA, 1999; p. 392.
37. Saaty, T.L. *The Analytic Hierarchy Process. Planning, Priority Setting, Resource Allocation*; McGraw-Hill: New York, NY, USA, 1980.
38. Ambituuni, A.; Amezaga, J.M.; Werner, D. Risk assessment of petroleum product transportation by road: A framework for regulatory improvement. *Saf. Sci.* **2015**, *79*, 324–335. [CrossRef]
39. Ferreira, C.E.C. Acidentes com motoristas no transporte rodoviário de produtos perigosos. *São Paulo Perspect.* **2003**, *17*, 68–80. [CrossRef]

International Journal of
Environmental Research and Public Health

MDPI

Brief Report

Associations between Bystanders and Perpetrators of Online Hate: The Moderating Role of Toxic Online Disinhibition

Sebastian Wachs [1,*] and Michelle F. Wright [2,3]

1 Department of Educational Studies, University of Potsdam, 14476 Potsdam, Germany
2 Department of Psychology, Pennsylvania State University, PA 16802, USA; mfw5215@psu.edu
3 Faculty of Social Studies, Masaryk University, 60200 Brno, Czech Republic
* Correspondence: wachs@uni-potsdam.de; Tel.: +49-331-977-2702

Received: 29 August 2018; Accepted: 15 September 2018; Published: 17 September 2018

Abstract: Hatred directed at members of groups due to their origin, race, gender, religion, or sexual orientation is not new, but it has taken on a new dimension in the online world. To date, very little is known about online hate among adolescents. It is also unknown how online disinhibition might influence the association between being bystanders and being perpetrators of online hate. Thus, the present study focused on examining the associations among being bystanders of online hate, being perpetrators of online hate, and the moderating role of toxic online disinhibition in the relationship between being bystanders and perpetrators of online hate. In total, 1480 students aged between 12 and 17 years old were included in this study. Results revealed positive associations between being online hate bystanders and perpetrators, regardless of whether adolescents had or had not been victims of online hate themselves. The results also showed an association between toxic online disinhibition and online hate perpetration. Further, toxic online disinhibition moderated the relationship between being bystanders of online hate and being perpetrators of online hate. Implications for prevention programs and future research are discussed.

Keywords: online hate; hate speech; bystander; perpetrator; online disinhibition; online discrimination; cyber aggression

1. Introduction

Hatred directed at members of groups due to their origin, race, gender, religion, or sexual orientation is not new, but it has taken on a new dimension in the online world. Online hate involves actions involving the denigration, harassment, exclusion, and advocacy of violence against specific groups on the basis of assigned or selected characteristics (i.e., origin, race, gender, religion, or sexual orientation) through information and communication technologies (ICTs) [1–3]. Although research regarding online hate is in its infancy, research on offline discrimination has shown that the consequences of exposure and victimization can be severe and it can promote deviant behavior, social disintegration tendencies, and negative health outcomes (i.e., psychosomatic problems, externalizing behavior problems) [4,5]. Therefore, it is important to understand why some adolescents perpetrate online hate. Because not much is known about possible correlates of online hate perpetration, the present study examined the association between being bystanders and perpetrators of online hate, and toxic online disinhibition as a potential moderator of this association. The results might help to deepen the knowledge concerning involvement of adolescents in online hate and how the online environment promotes online hate. The findings might provide information for prevention and intervention efforts to tackle online hate among adolescents, thereby promoting democratic coexistence in a pluralistic society.

There are different roles that adolescents can play in online hate, including being bystanders who observe online hate without being directly involved, being victims directly targeted by online hate material, comments, or posts, and being perpetrators who post, forward, or share harmful or hostile online hate material, comments, or posts [6]. Initial research on prevalence suggests that the most common way to experience online hate is by witnessing these behaviors as bystanders. For example, in a study with Finnish adolescents between 15 and 18 years, 53% had witnessed online hate, 6% had perpetrated online hate, and 23% had been victimized by online hate [1]. A common reaction of bystanders of online hate is taking no action, making a comment disagreeing with the online hate post, reporting it, liking the content, or blocking the post [6]. However, researchers have found that being bystanders and perpetrators of offline and online aggression are related [7–9]. Therefore, it seems assumable that some bystanders of online hate might also perpetrate online hate, which might be explained as follows. First, Social Learning Theory postulates that adolescents who observe deviant behavior and/or perceive that the peer group accepts these behaviors are more likely to engage in similar deviant behavior [10]. In these cases, perpetrating online hate might be explained in light of observational learning, adopting inappropriate coping strategies, and dynamic group processes. The Social Learning Theory has also been expanded to the online context by some researchers that have found that individuals tend to use more aggressive expressions in their online communication and interaction when their peers behave aggressively [11,12]. Second, some adolescents might become desensitized when observing online hate. For instance, initial research has shown that exposure to cyberbullying predicts lower levels of empathic responsiveness [13]. Third, some adolescents might be exposed more often to these behaviors because they are friends with perpetrators and share common values. For example, research on social networks of online and offline bullies has shown evidence for "nests" of cyberbullying perpetrators, assistants, and reinforcers [14,15].

The online environment involves anonymity, invisibility, asynchronicity, textuality, and lack of face-to-face contact, and punishment and repercussions are considered less likely to occur as compared with the offline world [16]. These circumstances can promote rude language, hatred, and threats, also referred to as toxic online disinhibition or the tendency to feel less inhibited [16]. Toxic online disinhibition can also decrease the ability for empathy, self-control, and the ability to recognize social cues [16,17]. When compared to the offline world, there is an increased likelihood that fewer adults are present in the online world of adolescents, which can also increase aggressive behavior and discrimination [18,19]. To the authors' knowledge, no study has investigated the association between toxic online disinhibition and online hate perpetration. However, past research has revealed that higher levels of toxic online disinhibition are positively associated with cyberbullying perpetration, flaming, and trolling [17,20–23]. Therefore, it can be proposed that toxic online disinhibition might also lead to less self-monitoring when expressing beliefs through hateful or degrading writing or speech online, making inappropriate attacks on minorities more likely.

The present study aims to contribute to the existing knowledge about online hate exposure and perpetration among adolescents by focusing on possible moderation effects of toxic online disinhibition in the association between being bystanders and perpetrators of online hate. In contrast to previous research on exposure and perpetration of online aggression, the current study will be the first to investigate these associations among online hate. To guide this purpose, the present study included the following hypotheses:

Hypothesis 1 (H1). *Being bystanders of online hate is related positively to being perpetrators of online hate.*

Hypothesis 2 (H2). *Higher levels of toxic online disinhibition are positively associated with being perpetrators of online hate.*

Hypothesis 3 (H3). *Higher levels of toxic online disinhibition increase the association between being bystanders and being perpetrators of online hate.*

2. Materials and Methods

2.1. Participants

In this study, 1480 students aged between 12 and 17 years old (students from grades seven to ten, respectively) were included (M_{age} = 14.21; SD = 1.22). They were from seven middle schools from the federal states of Bremen, Berlin, and Brandenburg in Germany. In terms of gender, 50.3% (n = 744) were girls. Regarding migration background, 9.7% (n = 144) reported that German was not the main language spoken at home. Around 33.6% (n = 483) of students reported living in families of low affluence, 33% (n = 474) in families of middle affluence, and 33.4 % (n = 479) in families of high affluence.

2.2. Measures

Online Hate Involvement. To measure online hate involvement, three items were adopted from the work of Hawdon et al. [2]. For assessing those who were bystanders of online hate, participants were asked: "How often in the past 12 months have you observed hateful or degrading writing or speech online, which inappropriately attacks certain groups of people or individuals because of their sex, religious affiliation, race, or sexual orientation?". For online hate perpetration, they were asked: "How often in the past 12 months have you posted hateful or degrading writing or speech online, which inappropriately attacks certain groups of people or individuals based on their sex, religious affiliation, race, or sexual orientation?". For online hate victimization, they were also asked: "How often in the past 12 months have you personally been the target of hateful or degrading writing or speech online because of your sex, religious affiliation, race, or sexual orientation?". Participants rated each item on a scale of 0 (never) to 4 (very frequently).

Toxic Online Disinhibition. The four-item Toxic Online Disinhibition Scale assessed the extent to which adolescents believed that they were less inhibited while interacting or engaging in certain behaviors online, with response options ranging from 0 (definitely do not believe) to 4 (definitely do believe) [22]. A confirmatory factor analysis revealed a good fit of the toxic online disinhibition scale: χ^2 = 29.33, df = 8, p < 0.001, comparative fit index (CFI) = 0.99, Tucker–Lewis index (TLI) = 0.99, root mean square error of approximation ($RMSEA$) = 0.04, and standardized root mean square residual ($SRMR$) = 0.01. A mean score was computed by averaging all items. Cronbach's alpha was 0.79.

Control Variables. Participants were asked for their age and sex to determine demographic characteristics. Migration background was assessed by asking which language is mainly spoken at home. Family socioeconomic status was measured with the Family Affluence Scale (FAS) [24]. The FAS was trichotomized into low, medium, and high socioeconomic status.

2.3. Procedures

All materials and procedures were approved by the data protection officer and educational authority of the federal state of Bremen, Germany, as well as University Institutional Review Board. Twenty schools were randomly selected from a list of 167 schools. From these 20 schools nine principals did not reply to the recruitment email, four expressed interest but had existing commitments that prevented them from participating, and seven provided agreement to have their school participate. There were 1788 parental permission slips passed out to the students. Of these, 1480 parents/guardians agreed to allow their child to participate. Reasons for not participating in this study were missing written parental consent, sick note, absence because of projects, internship, refusal to participate, unexcused absence at school, being new to the class and therefore not informed about the survey, or having refugee status (missing German language skills). An online survey was conducted during one regular school hour in the school's computer lab. All participants received instructions and were informed that their participation was optional, that they could choose not to answer questions, and that participation could be stopped at any time without giving a reason and with no consequences. In order to prevent distress and further harm by participating in this study, oral and written information where those who had taken part in the research could get counseling online and offline was given. Less than 3% of the data was incomplete and the missing data were handled with mean imputation [25].

2.4. Data Analyses

Descriptive statistics were used to determine the frequency rates of online hate. Pearson's r correlations were used to investigate the bivariate associations among the main study's variables. The *t*-test was used to investigate sex differences among the online hate variables, Cohen's *d* used to calculate the effect size. Confirmatory Factor Analysis was completed with Mplus 8.1 software (Muthén & Muthén, Los Angeles, CA, USA) [26]. The proposed regression-based moderated model was examined using the Process Macro for SPSS (SPSS Inc., Chicago, IL, USA) [27], applying Model 1 with 5000 bias-corrected bootstrap samples. Being bystanders of online hate was the independent variable, toxic online disinhibition was the moderator, and online hate perpetration was the dependent variable, while controlling for participants' age, sex, migration background, socioeconomic background, and online hate victimization. Cohen's f^2 was used as an effect size. According to Cohen [28] $f^2 \geq 0.10$, $f^2 \geq 0.25$, and $f^2 \geq 0.40$ represent small, medium, and large effect sizes, respectively. Multicollinearity diagnostics were assessed and were within an acceptable range (see Table 1).

Table 1. Means, standard deviations, and correlations between online hate bystanders, online hate perpetrators, online hate victims, and toxic online disinhibition.

Variable	1	2	3	4
1. Online hate bystanders	-	-	-	-
2. Online hate perpetrators	0.28 **	-	-	-
3. Online hate victims	0.40 **	0.31 **	-	-
4. Toxic online disinhibition	0.18 **	0.20 *	0.18 *	-
Mean	1.15	0.19	0.29	0.61
SD	1.32	0.62	0.74	0.73

* $p < 0.05$; ** $p < 0.01$.

3. Results

3.1. Descriptive Statistics

Correlations, means, and standard deviations for online hate bystanders, perpetrators, victims, and toxic online disinhibition are shown in Table 1. All variables were significantly correlated with each other.

Overall 53.7% (*n* = 761) of participants reported that they observed at least one incident of hateful or degrading writing or speech online, inappropriately attacking certain groups of people or individuals because of their sex, religious affiliation, race, or sexual orientation. Regarding the frequencies, 46.3% (*n* = 655) reported they have never had observed online hate, 18.6% (*n* = 263) reported observing online hate very rarely, 16.8% (*n* = 238) occasionally, 10.2% (*n* = 145) frequently, and 8.1% (*n* = 115) very frequently. Concerning online hate perpetration, 11.3% (*n* = 160) of participants reported that they had posted at least item of one hateful or degrading writing or speech online, inappropriately attacking certain groups of people or individuals because of their sex, religious affiliation, race or sexual orientation. Furthermore, 88.7% (*n* = 1256) reported they have never had posted online hate, 7.6% (*n* = 104) reported posting online hate very rarely, 1.8% (*n* = 26) occasionally, 0.9% (*n* = 13) frequently, and 1.2% (*n* = 17) very frequently. Regarding online hate victimization, 16.9% (*n* = 240) of participants reported that they have personally been the target of hateful or degrading writing or speech online because of their sex, religious affiliation, race, or sexual orientation. Additionally, 83.1% (*n* = 1178) reported they had never personally been targeted by online hate, 9.6% (n = 136) very rarely, 4.3% (n = 61) occasionally, 1.6% (n = 23) frequently, and 1.4% (n = 20) very frequently.

There was a positive correlation between age and observing online hate, $r = 0.10$, $p \leq 0.001$, and posting online hate $r = 0.10$, $p \leq 0.001$, but not with victimization through online hate. Girls (M = 1.37, SD = 1.36) reported more often than boys (M = 0.93, SD = 1.24) observing online hate online (t (1403) = 6.35, $p < 0.001$, Cohen's d = 0.35). Boys (M = 0.26, SD = 0.73) reported more often than girls (M = 0.11, SD = 0.47) posting online hate (t (1208) = −4.42, $p < 0.001$, Cohen's d = 0.24). However, no sex differences were found regarding online hate victimization.

3.2. Association between Online Hate Bystanders, Perpetrators, and Toxic Online Disinhibition

The overall model was significant, $F(6, 1357) = 11.87$, $p < 0.001$, $R^2 = 0.19$, indicating a large effect (Cohen's $f^2 = 0.53$). As Table 2 illustrates, there were statistically significant correlates of online hate perpetration. While controlling for participants' age, sex, migration background, and socioeconomic background, increases in being bystanders of online hate were positively related to being perpetrators of online hate (b = 0.08, SE = 0.19, $p < 0.001$). Toxic online disinhibition was positively associated with being perpetrators of online hate (b = 0.11, SE = 0.02, $p < 0.001$). Although age, migration background, and socioeconomic background were not significant predictors, online hate victimization (b = 0.16, SE = 0.04, p = 0.007) and sex (b = 0.16, SE = 0.05, $p < 0.001$) were significant predictors of online hate perpetration.

Table 2. Coefficients of the model predicting online hate perpetration.

Predictor	*b* (*)	*SE*	*t*	*p*
Constant	−0.530 [−0.915, −0.145]	0.196	−2.70	0.007
Toxic online disinhibition	0.116 [0.060, 0.172]	0.028	4.09	0.000
Online hate bystanders	0.086 [0.052, 0.119]	0.017	5.01	0.000
OHB × TOD	0.074 [0.019, 0.129]	0.279	2.67	0.007
Control Variables				
Age	0.021 [−0.002, 0.045]	0.031	5.08	0.080
Sex (male)	0.162 [0.099, 0.225]	0.059	0.967	0.000
Migration background	0.050 [−0.059, 0.175]	0.059	0.967	0.333
SES	0.022 [−0.014, 0.060]	0.019	1.19	0.230
Online hate victimization	0.168 [0.079, 0.256]	0.045	3.73	0.007

Note: OHB = online hate bystanders; TOD = toxic online disinhibition; SES = socioeconomic status; * 95% BCa = bootstrap confidence intervals based on 5000 samples.

As Figure 1 shows, significant moderation effects were found between bystanders of online hate and toxic online disinhibition when predicting online hate perpetration (b = 0.07, SE = 0.27, p = 0.007). Probing the interaction further revealed that bystanders of online hate reported more online hate perpetration when they reported higher levels of online disinhibition (b = 0.14, SE = 0.02, $p < 0.001$ at +1 SD) and less frequent online hate perpetration when they reported lower levels of toxic online disinhibition (b = 0.04, SE = 0.02, p = 0.029 at −1 SD).

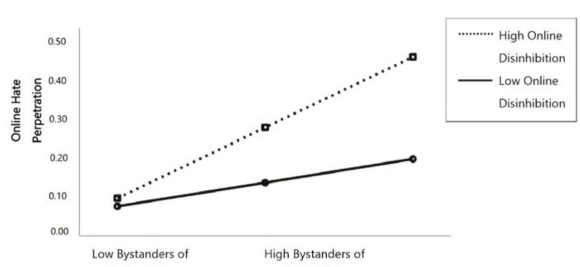


Figure 1. Simple slopes equations of the regression of online hate bystanders on online hate perpetrators at high and low levels of toxic online disinhibition.

4. Discussion

The purpose of this study was to fill a gap in the literature regarding the associations between being bystanders of online hate, toxic online disinhibition, and online hate perpetration. To address this aim, data were gathered from a sample of 1480 German adolescents aged between 12 and 17 years old. Notably, 53.7% had observed at least one online hate incident, 11.3% reported having perpetrated at least one incident of online hate, and 16.9% reported being victimized at least once by online hate. Therefore, online hate appears to be a prevalent issue among adolescents that warrants further investigation in the future. Our finding that the majority of students report observing online hate parallels that of Räsänen et al. [1]. These findings also underscore the need for more research into the experiences of adolescents who are bystanders of online hate.

We found support for our prediction that being bystanders and perpetrators of online hate would correlate (Hypothesis 1), even after controlling for the effects of being victims of online hate. Thus, it seems to be important to limit online hate exposure among adolescents. More broadly, our findings align with those indicating that being bystanders and being perpetrators of online and offline aggression is correlated [7,8]. A possible explanation might be that adolescents who observe online hate or perceive that their peers accept it perceive online hate as normal and unexceptional behavior are therefore more likely to perpetrate online hate. Another explanation might be that some adolescents might become desensitized by observing online hate, making online hate seem like a potentially normative behavior.

The evidence showed that, as expected, higher levels of toxic online disinhibition were positively associated with online hate perpetration (Hypothesis 2). This result extends previous research that revealed positive associations between online disinhibition and cyberbullying perpetration, flaming, and trolling [17,20–23]. Nevertheless, personal features like impulsivity, repressed emotions, personal drives, and one's own experiences with exclusion and discrimination may also be important predictors of online hate perpetration, which need to be investigated in future research thoroughly.

We add to the literature that, consistent with expectations, toxic online disinhibition moderated the associations between being bystanders and being perpetrators of online hate (Hypothesis 3). Thus, the online disinhibition effect might be a key variable in understanding why adolescents who observe online hate also perpetrate online hate. More research is needed whether toxic online disinhibition might also moderate associations between other participating roles, such as between victims and perpetrators of online hate. Although this finding sheds light on possible contextual factors that explain the association between being bystanders and perpetrators of online hate, more research is needed to understand whether intra- and interpersonal factors (i.e., desensitization process, social norms within the peer group, popularity of the perpetrator, the subjective nature of the perceived severity of online hate) might also contribute to the correlation between being bystanders and being perpetrators of online hate. Finally, more research is needed to understand whether benign online disinhibition controversially might buffer the association between being bystanders and perpetrators of online hate.

5. Limitations

There are several limitations of this research requiring some discussion. First, the cross-sectional nature of the study's design limits the ability to draw any causal conclusions and temporal ordering of the main study constructs. Future research would benefit from longitudinal studies. Second, the data were exclusively collected through self-reports. Therefore, the observed relationships might be inflated due to shared method variance. In addition, measuring online hate solely through self-reports may affect adolescents' reports of these experiences. For example, some adolescents who perpetrated online hate might choose not to answer honestly for fear of consequences (i.e., restricted ICT use). Adolescents who think that they have a legitimate opinion or are just joking around might not recognize that they perpetrate online hate and consequently they might also underreport online hate. There is some research suggesting that people who recognize racism are more likely to perceive it as a deviant and negative behavior [6]. More specifically, the items used to measure online hate refer to "inappropriate attacks" which might also lead to underreporting as adolescents might not perceive such behavior as inappropriate. On the other hand, it might be that some adolescents overreport perpetrating online hate to appear tough. Follow-up research should apply a multi-informant approach. Third, we did not control for involvement in other forms of cyber aggressions (i.e., cyberbullying, trolling), ICT access, time spent online, or online activities, all of which may have an impact on online hate perpetration. Future research should include theses control variables. Finally, we relied on single item measurement for the assessment of online hate. Follow-up studies should try to include validated scales to overcome concerns with one-item measurements (i.e., low content validity, sensitivity, and lack a measure of internal consistency reliability).

6. Practical Implications

The findings of the present study signify a need for school staff, policy makers, and providers of social media to be aware of adolescents' exposure to online hate. Schools and their educational mission face a double challenge regarding online hate. On one hand, online hate is not just an online phenomenon, it can also affect peaceful coexistence at school. On the other hand, as a democracy-fostering authority, schools are predestined to counter online hate by teaching appropriate skills (i.e., media literacy, conflict strategies, democratic, and social skills). These skills can be enhanced through prevention programs that help adolescents understand that democratic values and basic human rights also apply to the online world. These programs should also aim to foster empathy with victims, take the perspective of the victim, embrace diversity, and enable adolescents to recognize and cope with online hate. The present study showed that toxic online disinhibition might prevent adolescents from becoming online hate perpetrator. Therefore, it seems important to increase the awareness among adolescents concerning how the online environment influences their own behavior. Increasing self-control, learning techniques for critical self-monitoring, and fostering of the ability to recognize social cues and self-reflection might reduce the effects of online disinhibition. Furthermore, policy makers have to ensure that they recognize the balance between enabling freedom of expression, protection from online hate, and punishing perpetrators. It is important that policy is developed to persuade people to not post hateful content in the first place or convince people to remove the content themselves and apologize. Policy makers should also urge social media platforms to rank such content lower in social media news feeds and/or to implement more efficient procedures for reporting the content. Social media platforms need to intensify their efforts to protect adolescents from online hate exposure by implementing practical systems to reduce exposure to online hate and for removal of online hate material. Social media platforms should also be expected to remove online hate content in a more efficient manner, as sometimes there is a lengthy amount of time between reporting and removal of the content.

7. Conclusions

This study was one of the first to examine the association between being bystanders and being perpetrators of online hate. It was also one of the first to investigate the moderating effect of toxic online disinhibition in this relationship. Findings indicate that increases in being bystanders of online hate and toxic online disinhibition were positively associated with being an online hate perpetrator, while controlling for participants' age, sex, migration background, socioeconomic status, and experience of online hate victimization. Moreover, a significant interaction was found between being bystanders of online hate and toxic online disinhibition when predicting online hate perpetration. The findings of the present study indicate a need for school staff, policy makers, and providers of social media to be aware of the possible impact that witnessing online hate and online disinhibition can have on adolescents' behavior. In addition, the identification of correlates, such as toxic online disinhibition, could help with the development of prevention and intervention programs aimed at changing adolescents' online behavior.

Author Contributions: S.W. and M.F.W. developed the study design. S.W. performed the statistical analyses and drafted the manuscript. M.F.W. provided feedback on drafts of the manuscript. S.W. processed all feedback from the coauthor and reviewers. Both authors read and approved of the final manuscript.

Funding: We acknowledge the support of the Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of University of Potsdam.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Räsänen, P.; Hawdon, J.; Holkeri, E.; Keipi, T.; Näsi, M.; Oksanen, A. Targets of online hate: Examining determinants of victimization among young Finnish Facebook users. *Violence Vict.* **2016**, *31*, 708–726. [CrossRef] [PubMed]
2. Hawdon, J.; Oksanen, A.; Räsänen, P. Exposure to online hate in four nations: A cross-national consideration. *Deviant Behav.* **2017**, *38*, 254–266. [CrossRef]
3. Sponholz, L. Theoretische Grundlagen und empirische Umsetzung. In *Hate Speech in den Massenmedien*; VS Verlag: Wiesbaden, Germany, 2018; ISBN 978-3-658-15077-8.
4. Unnever, J.D.; Cullen, F.T.; Barnes, J. Racial discrimination, weakened school bonds, and problematic behaviors: Testing a theory of African American offending. *J. Res. Crime Delinq.* **2016**, *53*, 139–164. [CrossRef]
5. Kauff, M.; Wölfer, R.; Hewstone, M. Impact of discrimination on health among adolescent immigrant minorities in Europe: The role of perceived discrimination by police and security personnel. *J. Soc. Issues* **2017**, *73*, 831–851. [CrossRef]
6. Jakubowicz, A.; Dunn, K.; Mason, G.; Paradies, Y.; Bliuc, A.-M.; Bahfen, N.; Oboler, A.; Atie, R.; Connelly, K. *Cyber Racism and Community Resilience*; Springer: Cham, Switzerland, 2017; ISBN 978-3-319-64388-5.
7. Rivers, I.; Poteat, V.P.; Noret, N.; Ashurst, N. Observing bullying at school: The mental health implications of witness status. *Sch. Psychol. Q.* **2009**, *24*, 211–223. [CrossRef]
8. Erreygers, S.; Pabian, S.; Vandebosch, H.; Baillien, E. Helping behavior among adolescent bystanders of cyberbullying: The role of impulsivity. *Learn. Individ. Differ.* **2016**, *48*, 61–67. [CrossRef]
9. Pabian, S.; Vandebosch, H.; Poels, K.; Van Cleemput, K.; Bastiaensens, S. Exposure to cyberbullying as a bystander: An investigation of desensitization effects among early adolescents. *Comput. Hum. Behav.* **2016**, *62*, 480–487. [CrossRef]
10. Akers, R.L.; Krohn, M.D.; Lanza-Kaduce, L.; Radosevich, M. Social learning and deviant behavior: A specific test of a general theory. *Am. Sociol. Rev.* **1979**, *44*, 636–655. [CrossRef] [PubMed]
11. Rösner, L.; Krämer, N.C. Verbal venting in the social web: Effects of anonymity and group norms on aggressive language use in online comments. *Soc. Media Soc.* **2016**. [CrossRef]
12. Miller, B.; Morris, R.G. Virtual peer effects in social learning theory. *Crime Delinq.* **2016**, *62*, 1543–1569. [CrossRef]
13. Mrug, S.; Madan, A.; Windle, M. Emotional desensitization to violence contributes to adolescents' violent behavior. *J. Abnorm. Child Psychol.* **2016**, *44*, 75–86. [CrossRef] [PubMed]
14. Festl, R.; Scharkow, M.; Quandt, T. Peer influence, internet use and cyberbullying: A comparison of different context effects among German adolescents. *J. Child. Media* **2013**, *7*, 446–462. [CrossRef]
15. Salmivalli, C. Participant roles in bullying: How can peer bystanders be utilized in interventions? *Theor. Pract.* **2014**, *53*, 286–292. [CrossRef]
16. Suler, J. The online disinhibition effect. *Cyberpsychol. Behav.* **2004**, *7*, 321–326. [CrossRef] [PubMed]
17. Voggeser, B.J.; Singh, R.K.; Göritz, A.S. Self-control in Online Discussions: Disinhibited Online Behavior as a Failure to Recognize Social Cues. *Front. Psychol.* **2018**, *8*, 2372. [CrossRef] [PubMed]
18. Hinduja, S.; Patchin, J.W. Cyberbullying: An exploratory analysis of factors related to offending and victimization. *Deviant Behav.* **2008**, *29*, 129–156. [CrossRef]
19. Tynes, B.; Reynolds, L.; Greenfield, P.M. Adolescence, race, and ethnicity on the Internet: A comparison of discourse in monitored vs. unmonitored chat rooms. *J. Appl. Dev. Psychol.* **2004**, *25*, 667–684. [CrossRef]
20. Wright, M.F. Predictors of anonymous cyber aggression: The role of adolescents' beliefs about anonymity, aggression, and the permanency of digital content. *Cyberpsychol. Behav. Soc. Netw.* **2014**, *17*, 431–438. [CrossRef] [PubMed]
21. Wright, M.F.; Harper, B.D.; Wachs, S. The associations between cyberbullying and callous-unemotional traits among adolescents: The moderating effect of online disinhibition. *Personal. Individ. Differ.* **2018**. [CrossRef]

22. Udris, R. Cyberbullying among high school students in Japan: Development and validation of the Online Disinhibition Scale. *Comput. Hum. Behav.* **2014**, *41*, 253–261. [CrossRef]
23. Görzig, A.; Ólafsson, K. What makes a bully a cyberbully? Unravelling the characteristics of cyberbullies across twenty-five European countries. *J. Child. Media* **2013**, *7*, 9–27. [CrossRef]
24. Boyce, W.; Torsheim, T.; Currie, C.; Zambon, A. The family affluence scale as a measure of national wealth: Validation of an adolescent self-report measure. *Soc. Indic. Res.* **2006**, *78*, 473–487. [CrossRef]
25. Little, R.J.; Rubin, D.B. *Statistical Analysis with Missing Data*; Wiley: New York, NY, USA, 2002; ISBN 978-0-471-18386-0.
26. Muthen, L.K.; Muthen, B.O. *Mplus (Computer Software)*; Muthén & Muthén: Los Angeles, CA, USA, 2017.
27. Hayes, A. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*; The Guilford Press: New York, NY, USA, 2013; ISBN 9781462534654.
28. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Academic Press: New York, NY, USA, 1988; ISBN 9781483276489.

International Journal of
Environmental Research and Public Health

MDPI

Article

Oral Microbiota of the Snake *Bothrops lanceolatus* in Martinique

Dabor Résière [1,*], Claude Olive [2], Hatem Kallel [3], André Cabié [4], Rémi Névière [1], Bruno Mégarbane [5,†], José María Gutiérrez [6,†] and Hossein Mehdaoui [1]

1 Intensive Care Unit, University Hospital of Martinique, Fort-de-France, 97200 Martinique, France; Remi.NEVIERE@chu-martinique.fr (R.N.); Hossein.MEHDAOUI@chu-martinique.fr (H.M.)
2 Department of Microbiology, University Hospital of Martinique, 97200 Martinique, France; Claude.OLIVE@chu-martinique.fr
3 Intensive Care Unit, Rosemond André General Hospital, Cayenne, 97300 French Guiana, France; hatem.kallel@ch-cayenne.fr
4 Department of Infectious Diseases, University Hospital of Martinique, 97200 Martinique, France; Andre.CABIE@chu-martinique.fr
5 Department of Medical and Toxicological Critical Care, Lariboisière Hospital, Paris-Diderot University, INSERM UMRS-1144, 75010 Paris, France; bruno.megarbane@lrb.aphp.fr
6 Instituto Clodomiro Picado, Facultad de Microbiología, Universidad de Costa Rica, 11501 San José, Costa Rica; jose.gutierrez@ucr.ac.cr
* Correspondence: dabor.resiere@chu-martinique.fr; Tel.: +1-(596)-696-203-184
† The authors contribute equally to this work.

Received: 17 August 2018; Accepted: 19 September 2018; Published: 27 September 2018

Abstract: In Martinique, *Bothrops lanceolatus* snakebite, although relatively uncommon (~30 cases/year), may result in serious complications such as systemic thrombosis and local infections. Infections have been hypothesized to be related to bacteria present in the snake's oral cavity. In this investigation, we isolated, identified, and studied the susceptibility to beta-lactams of bacteria sampled from the oral cavity of twenty-six *B. lanceolatus* specimens collected from various areas in Martinique. Microbiota from *B. lanceolatus* oral cavity was polymicrobial. Isolated bacteria belonged to fifteen different taxa; the most frequent being *Aeromonas hydrophyla* (present in 50% of the samples), *Morganella morganii*, *Klebsiella pneumoniae*, *Bacillus* spp., and *Enterococcus* spp. Analysis of antibiotic susceptibility revealed that 66.7% of the isolated bacteria were resistant to amoxicillin/clavulanate. In contrast, the majority of isolated bacteria were susceptible to the third-generation cephalosporins (i.e., 73.3% with cefotaxime and 80.0% with ceftazidime). Microbiota from *B. lanceolatus* oral cavity is polymicrobial with bacteria mostly susceptible to third-generation cephalosporins but rarely to amoxicillin/clavulanate. In conclusion, our findings clearly support that first-line antibiotic therapy in the *B. lanceolatus*-bitten patients, when there is evidence of infection, should include a third-generation cephalosporin rather than amoxicillin/clavulanate.

Keywords: *Bothrops lanceolatus*; envenomation; snakebite; bacteria; infection; antibiotic susceptibility

1. Introduction

Snakebite envenomation is a relatively uncommon but serious medical emergency in Martinique, an overseas region of France with a population of 386,000 inhabitants, located in the Lesser Antilles of the Eastern Caribbean. Approximately thirty cases are declared every year. Envenomation is due to *Bothrops lanceolatus* (Order: Squamata; suborder: Serpentes; family: Viperidae; subfamily: Crotalinae; Figure 1), the only venomous snake in this Caribbean French territory and a snake not found elsewhere in the world [1]. This envenoming is known to be responsible for a high rate of multiple systemic thrombotic events as well as for possible local infectious complications.

Figure 1. *Bothrops lanceolatus* from Martinique, known as trigonocéphale or fer-de-lance is one the most dangerous venomous snakes in the Caribbean. (Photo courtesy of Mr. Michel Tanasi, a member of the snake working group).

As suggested by previous observations at the bedside [2,3], bacterial infection from snakebites in Martinique occurs in about one third of the envenomed cases, mainly in the most severely envenomed patients (grade II or III). Interestingly, local infection was hypothesized to be caused by the oral and fang microbiota of *B. lanceolatus*.

The oral microbiota of snakes comprises a wide range of aerobic and anaerobic microorganisms, including *Enterobacteriaceae* (*Morganella* spp. and *Escherichia coli*), *Streptococcus* spp., *Aeromonas* spp., *Staphylococcus aureus*, and *Clostridium* spp., as reported in several observational studies from snakebites worldwide [4–9]. However, the predominant microorganisms change according to the geographic region and environmental conditions. Additionally, bacteria susceptibility to beta-lactams, the most frequently used antibiotics to treat snake-envenomed patients, varies with possible resistant phenotypes, especially since the emergence of multidrug-resistant bacteria in the environment during the last years.

Therefore, we designed an experimental study aiming to identify the main bacterial microbiota from *B. lanceolatus* oral cavity. Then, based on the nature and susceptibility of the isolated bacteria to beta-lactams, our objective was to propose the most adequate preemptive antibiotic therapy that should be administered in the suspicion of infection in *B. lanceolatus*-bitten patients in Martinique.

2. Materials and Methods

2.1. Experimental Design

This experimental study was conducted at the Microbiology Laboratory of the University Hospital of Martinique. During eighteen months, all *B. lanceolatus* snakes captured by personnel of the National Office of Forests in Martinique were studied. Snakes were grouped according to their geographical origin, i.e., wet (forest and near water areas) versus dry regions (peri-urban areas). The snake's oral cavity was opened and sampled from the vicinity of the fangs using sterile cotton swabs. Samples were subjected to Gram staining and examined for bacterial growth. They were plated on non-selective blood agar and chocolate agar and cultured at 37 °C for 2–7 days and the color and shape of the colonies were observed. Species identification was performed with API-20E and API-20NE systems (BioMérieux, Marcy L'Etoile, France). Antimicrobial susceptibilities of all isolates to beta-lactams were determined by the disk diffusion methods based on the definitions of the Antibiogram Committee of the French Microbiology Society [10]. The inhibition zone diameters of each drug for each isolate were determined after overnight incubation at 35.8 °C in ambient air. The interpretive criteria of the inhibition zone and minimum inhibitory concentrations were in accordance with those of the Antibiogram Committee of the French Society of Microbiology.

2.2. Statistical Analysis

Data are expressed as absolute values and percentages. Comparisons between the subgroups were performed using Chi-2 tests. p-Values less than 0.05 were considered to be statistically significant.

3. Results

During the study period, twenty-six specimens of *B. lanceolatus* were captured and their oral cavity sampled for bacteriological culture. Twenty snakes were captured from the "wet" zones versus six from the "dry" zones. All samples obtained from the snake mouths tested positive for bacterial growth.

In 20 cases (76.9%), the sample was polymicrobial. The most frequently isolated bacteria were *Aeromonas hydrophila*, *Morganella morganii*, and *Klebsiella pneumonia* (Table 1). *A. hydrophila* was isolated in 50% of the samples (13 cases). In four cases (15.4%), pure cultures were obtained or the bacterium was quantitatively predominant. *A. hydrophila* was isolated in 60.0% of cases from snakes captured from a "wet" zone compared to 16.7% from snakes captured from a "dry" zone ($p = 0.06$).

Table 1. Bacteria isolated from the oral cavity of *Bothrops lanceolatus* in Martinique and their susceptibility to beta-lactams.

Microorganism	N	AMX	AMX-Clav	CTX	CAZ	% of Bacteria ($N = 46$)	% of Specimens ($N = 26$)
Aeromonas hydrophila	13	R	R	S	S	28.3	50
Morganella morganii	7	R	R	S	S	15.2	26.9
Klebsiella pneumoniae	5	R	S	S	S	10.9	19.2
Bacillus spp.	4	R	R	R	R	8.7	15.4
Enterococcus spp.	3	S	S	R	R	6.5	11.5
Proteus mirabilis	2	S	S	S	S	4.3	7.7
Serratia marcescens	2	R	R	S	S	4.3	7.7
Shewanella putrefaciens	2	R	R	S	S	4.3	7.7
Clostridium bifermentans	2	R	R	S	S	4.3	7.7
Proteus penneri	1	S	S	S	S	2.2	3.8
Proteus vulgaris	1	S	S	S	S	2.2	3.8
Enterobacter cloacae	1	R	R	S	S	2.2	3.8
Citrobacter freundii	1	R	R	S	S	2.2	3.8
Chryseomonas violaceum	1	R	R	R	R	2.2	3.8
Pseudomonas pickettii	1	R	R	R	S	2.2	3.8
Total	46	-	-	-	-	100	100
% susceptible	-	26.7	33.3	73.3	80.0	-	-

S, susceptible; R, resistant; AMX, amoxicillin; AMX-Clav, amoxicillin/clavulanate; CTX, cefotaxime; CAZ, ceftazidime.

In 66.7% of the cases, the isolated microorganisms were resistant to amoxicillin/clavulanate, while bacteria isolated were more frequently susceptible to the third-generation cephalosporins including cefotaxime (73.3%) and ceftazidime (80.0%).

4. Discussion

In this study, the most frequently isolated microorganisms from *B. lanceolatus* oral cavity in Martinique were *A. hydrophila*, followed by other members of the *Enterobacteriaceae* family. The cultures were polymicrobial and bacteria were susceptible to third-generation cephalosporins in most cases, coinciding with the majority of studies that focused on other snakes from different regions of the Americas [8,10–12].

B. lanceolatus is endemic to Martinique in the French West Indies. Its venom can cause severe local damage, ranging from blistering and tissue necrosis to secondary cellulitis and abscess in severe cases [1]. As previously shown [2–7], the oral microbiota of snakes is known to contain a wide range of microorganisms. Conformingly, *B. lanceolatus* oral flora is polymicrobial, combining Gram-positive, Gram-negative bacteria were isolated, with varying proportions across the Martinique sub-regions.

A. hydrophila was the most commonly isolated bacterial species, especially when the snake was captured from a "wet" zone. This bacterium is known to cause tissue damage and necrotizing fasciitis [13–15], raising the hypothesis that the tissue damage observed after *B. lanceolatus* bite may result from the tissue-damaging properties of the venom combined with the necrosis potency of *A. hydrophila*. It is obvious that tissue damage induced by the venom is responsible for perfusion abnormalities in the zone of the bite. This is due to microvascular damage and thrombosis induced by the venom itself, in addition to microcirculatory compression due to local edema and compartment syndrome in some cases. In these conditions, tissue damage and ischemia, with an associated decrease in the local innate immunity, is an excellent medium for the growth of bacteria ejected from the snake's oral cavity at the moment of the bite, as previously demonstrated for *Staphylococcus aureus* in an experimental model [16].

In the case of a moderate to severe snakebite, most authors and guidelines recommend the use of antibiotics to reduce complications by preventing secondary infection. However, the 2010 World Health Organization statements advised against the use of preemptive antibiotics in snake bites except in certain circumstances [17]. Our study clearly showed that envenomings by *B. lanceolatus* species have a high incidence of bite abscesses and, thus, represent an ideal candidate to use antibiotics, if they are given when there is the evidence of infection soon enough after the incident and when appropriate antibiotics are employed.

The most recommended antibiotic for snakebite treatment is amoxicillin/clavulanate, albeit without strong evidence supporting this recommendation. The Infectious Diseases Society of America (IDSA) guidelines for the diagnosis and management of skin and soft-tissue infections recommended amoxicillin/clavulanate as the preemptive antibiotic of choice to reduce complications by preventing secondary infection from animal bites other than snakes [18]. However, preemptive amoxicillin/clavulanate was reported to be ineffective in preventing secondary infections from *Bothrops* snakebites in the Western Brazilian Amazon region [19]. Similarly, in another study based on a 10-year experience in a northern Taiwan medical center, amoxicillin/clavulanate alone appeared non-convenient for the empirical or definitive treatment of soft tissue infections after snakebite [20]. In this setting, the authors advised to use the combination of amoxicillin/clavulanate with ciprofloxacin or to choose parenteral piperacillin/tazobactam.

Our findings were consistent with these two studies [19,20], since we showed that amoxicillin/clavulanate was not effective against 66.7% of the isolated bacteria. By contrast, we found that the most appropriate first-line antibiotics were the third-generation cephalosporins that should be preferred in *B. lanceolatus*-bitten patients with signs of local infection. However, since the systematic prophylactic antibiotic use in snakebite patients remains a matter of debate [21], the use of antibiotics in *B. lanceolatus* cases should be restricted to cases where there is evidence of infection.

5. Conclusions

Microbiota from *B. lanceolatus* oral cavity is polymicrobial with bacteria mostly susceptible to third-generation cephalosporins and more rarely to amoxicillin/clavulanate. In the absence of definitive international guidelines for the management of snakebite-associated infections, our findings support that the first-line antibiotic therapy should include a third-generation cephalosporin rather than amoxicillin/clavulanate in patients suffering infections as a consequence of *B. lanceolatus* snakebites in Martinique.

Author Contributions: Conceptualization, D.R., C.O., and H.M.; methodology, C.O. and R.N.; software, H.M.; validation, C.O., R.N., and H.M.; formal Analysis, D.R., C.O., H.M.; investigation, D.R., C.O., H.K., and A.C.; resources, C.O. and H.M.; data curation, C.O. and H.M.; writing—original draft preparation, D.R. and H.K.; writing—review and editing, J.M.G. and B.M.; visualization, H.M.; supervision, J.M.G., B.M., and H.M.; project administration, B.M. and H.M.

Funding: The authors received no specific funding for this work.

Acknowledgments: We would like to thank the personnel of the National Office of Forests in Martinique who caught the snakes.

Conflicts of Interest: The authors have declared that no competing interests exist.

References

1. Resiere, D.; Mégarbane, B.; Valentino, R.; Mehdaoui, H.; Thomas, L. *Bothrops lanceolatus* bites: Guidelines for severity assessment and emergent management. *Toxins* **2010**, *2*, 163–173. [CrossRef] [PubMed]
2. Jorge, M.T.; Mendonça, J.S.; Ribeiro, L.A.; Silva, M.L.R.; Kusano, E.J.U.; Cordeiro, C.L.S. Flora bacteriana da cavidade oral, presas e veneno de *Bothrops jararaca*: Possivel fonte de infecção no local da picada. *Rev. Inst. Med. Trop São Paulo* **1990**, *32*, 6–10. [CrossRef] [PubMed]
3. Thomas, L.; Tyburn, B.; Bucher, B.; Pecout, F.; Ketterle, J.; Rieux, D.; Smadja, D.; Garnier, D.; Plumelle, Y. Prevention of thromboses in human patients with *Bothrops lanceolatus* envenoming in Martinique: Failure of anticoagulants and efficacy of a monospecific antivenom. Research Group on Snake Bites in Martinique. *Am. J. Trop. Med. Hyg.* **1995**, *52*, 419–426. [CrossRef] [PubMed]
4. Thomas, L.; Tyburn, B.; Ketterlé, J.; Biao, T.; Mehdaoui, H.; Moravie, V.; Rouvel, C.; Plumelle, Y.; Bucher, B.; Canonge, D.; Marie-Nelly, C.A.; et al. Prognostic significance of clinical grading of patients envenomed by *Bothrops lanceolatus* in Martinique. Members of the Research Group on Snake Bite in Martinique. *Trans. R. Soc. Trop. Med. Hyg.* **1998**, *92*, 542–545. [CrossRef]
5. Goldstein, E.J.C. Bite wounds and infection. *Clin. Infect. Dis.* **1992**, *14*, 633–640. [CrossRef] [PubMed]
6. Garg, A.; Sujatha, S.; Garg, J.; Acharya, N.S.; Parija, S.C. Wound infections secondary to snakebite. *J. Infect. Dev. Ctries* **2009**, *3*, 221–223. [CrossRef] [PubMed]
7. Brazilian Ministry of Health; Fundaçao Nacional de Saúde. *Manual de Diagnóstico e Tratamento de Acidentes por Animais Peçonhentos*; Fundaçao Nacional de SaúÂde: Brasilia, Brazil, 2001.
8. Jorge, M.T.; Mendonça, J.S.; Ribeiro, L.A.; Cardoso, J.L.C.; Da Silva, M.L.R. Bacilos Gram-negativos aeroÂbios em abscessos por acidente botrópico. *Rev. Soc. Bras. Med. Trop.* **1987**, *20*, 55.
9. Bastos, H.M.; Lopes, L.F.L.; Gattamorta, M.A.; Matushima, E.R. Prevalence of enterobacteria in *Bothrops jararaca* in São Paulo State: Microbiological survey and antimicrobial resistance standards. *Acta Sci. Biol. Sci.* **2008**, *30*, 321–326. [CrossRef]
10. Soussy, C.J.; Carret, G.; Cavallo, J.D.; Chardon, H.; Chidiac, C.; Choutet, P.; Courvalin, P.; Dabernat, H.; Drugeon, H.; Dubreuil, L.; et al. Antibiogram Committee of the French Microbiology Society. Report 2000–2001. *Pathol. Biol.* **2000**, *48*, 832–871. [PubMed]
11. Fonseca, M.G.; Moreira, W.M.Q.; Counha, K.C.; Ribeiro, A.C.M.G.; Almeida, M.T.G. Oral microbiota of Brazilian captive snakes. *J. Venom. Anim. Toxins Incl. Trop. Dis.* **2009**, *15*, 54–60. [CrossRef]
12. Jorge, M.T.; Ribeiro, L.A.; Da Silva, M.L.; Kusano, E.J.; Mendonça, J.S. Microbiological studies of abscesses complicating Bothrops snakebite in humans: A prospective study. *Toxicon* **1994**, *32*, 261–270. [CrossRef]
13. Minnaganti, V.R.; Patel, P.J.; Iancu, D.; Schoch, P.E.; Cunha, B.A. Necrotizing fasciitis caused by *Aeromonas hydrophila*. *Heart Lung* **2000**, *29*, 306–308. [CrossRef] [PubMed]
14. Apisarnthanarak, A.; Pheerapiboon, P.; Apisarnthanarak, P.; Kiratisin, P.; Mundy, L.M. Fulminant epiglottitis with evolution to necrotizing soft tissue infections and fasciitis due to *Aeromonas hydrophila*. *Infection* **2008**, *36*, 94–95. [CrossRef] [PubMed]
15. Chern, C.H.; How, C.K.; Huang, L.J. Images in emergency medicine. Necrotizing fasciitis caused by *Aeromonas hydrophila*. *Ann. Emerg. Med.* **2006**, *48*, 216–225. [CrossRef] [PubMed]

16. Saravia-Otten, P.; Gutierrez, J.M.; Arvidson, S.; Thelestam, M.; Flock, J.I. Increased infectivity of *Staphylococcus aureus* in an experimental model of snake venom-induced tissue damage. *J. Infect. Dis.* **2007**, *196*, 748–754. [CrossRef] [PubMed]
17. World Health Organization. Guidelines for the Prevention and Clinical Management of Snakebite in Africa. Available online: http://apps.who.int/medicinedocs/documents/s17810en/s17810en.pdf (accessed on 17 September 2018).
18. Stevens, D.L.; Bisno, A.L.; Chambers, H.F.; Dellinger, E.P.; Goldstein, E.J.C.; Gorbach, S.L.; Hirschmann, J.V.; Kaplan, S.L.; Montoya, J.G.; Wade, J.C. Practice guidelines for the diagnosis and management of skin and soft tissue infections: 2014 update by the infectious diseases society of America. *Clin. Infect. Dis.* **2014**, *59*, 147–159. [CrossRef] [PubMed]
19. Sachett, J.A.G.; da Silva, I.M.; Alves, E.C.; Oliveira, S.S.; Sampaio, V.S.; do Vale, F.F.; Romero, G.A.S.; Dos Santos, M.C.; Marques, H.O.; Colombini, M.; et al. Poor efficacy of preemptive amoxicillin clavulanate for preventing secondary infection from *Bothrops* snakebites in the Brazilian Amazon: A randomized controlled clinical trial. *PLoS Negl. Trop. Dis.* **2017**, *11*, e0005745. [CrossRef] [PubMed]
20. Chen, C.M.; Wu, K.G.; Chen, C.J.; Wang, C.M. Bacterial infection in association with snakebite: A 10-year experience in a northern Taiwan medical center. *J. Microbiol. Immunol. Infect.* **2011**, *44*, 456–460. [CrossRef] [PubMed]
21. Blaylock, R.S. Antibiotic use and infection in snakebite victims. *S. Afr. Med. J.* **1999**, *89*, 874–876. [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

Health-Related Quality of Life among People Applying for Housing Adaptations: Associated Factors

Lovisa Boström [1], Carlos Chiatti [1], Björg Thordardottir [2], Lisa Ekstam [1] and Agneta Malmgren Fänge [1,*]

[1] Department of Health Sciences, Faculty of Medicine, Lund University, P.O. Box 157, 221 00 Lund, Sweden; mev15lbo@student.lu.se (L.B.); carlos.chiatti@med.lu.se (C.C.); lisa.ekstam@med.lu.se (L.E.)
[2] Department of Occupational Therapy, Prosthetics and Orthotics, Faculty of Health Sciences, OsloMet—Oslo Metropolitan University, P.O. Box 4 St. Olavs Plass, 0130 Oslo, Norway; bjorgt@oslomet.no
* Correspondence: agneta.malmgren_fange@med.lu.se; Tel.: +46-046-222-1972

Received: 30 August 2018; Accepted: 25 September 2018; Published: 27 September 2018

Abstract: Housing adaptations (HA) clients are a heterogeneous group of people with disabilities experiencing restricted performance and participation in everyday life. While health-related quality of life is a common and relevant outcome in health care research, associated factors among HA clients are largely unknown. Thus, the aim of this study was to investigate which factors are associated with health-related quality of life among HA clients in Sweden. The study has a cross-sectional design, using baseline data collected among 224 participants in three municipalities in Sweden. The main outcome was health-related quality of life measured by the EQ-5D. Factors investigated as potentially associated were age, sex, living conditions, cognitive impairment, usability of the home, activities of daily living (ADL) dependence, participation, and fear of falling. The associations were explored using multiple linear regression analysis. Younger age and higher dependence in ADL were associated with lower scores on the EQ-5D. The social aspect of usability in the home had a positive association with the EQ VAS. Results suggest that certain groups of HA clients might be at risk for low health-related quality of life. Knowledge of their characteristics can potentially improve development and implementation of tailored interventions aiming at increasing their health-related quality of life.

Keywords: ADL; disability; health; home; usability

---

1. Introduction

Health-related quality of life is widely accepted as important outcomes for evaluating the effect of interventions, both on individual and societal level [1,2]. Apart from having proved to be a strong predictor of mortality [3,4], self-rated health is a valuable alternative indicator of a health condition [2]. Studies have shown that a variety of factors affect how people rate their own health. In general, personal factors such as higher age, being a woman, low income, low education level, living alone, and body functions and structure such as cognitive impairment and having a chronic disease are all related with worse scores on self-rated health assessments [5–7]. Mobility, disability, and high risk of falling are also known factors associated with lower self-rated health [8,9]. Falls are common among older people, and most falls occur in people's own homes [10–12], as consequence of environmental factor hazards such as tripping over thresholds or loose rugs [11,12].

The ageing-in-place principle, aiming to allow people to remain living in their own homes while ageing, has been inspiring policy reform in Sweden and many other countries [13]. Several older adults aspire to remain in their homes for several personal and practical reasons. This desire is also beneficial from a societal point of view, as it relates to lower costs for the society, compared for example to the cost of nursing home placements [14]. In this respect, a factor contributing to successful ageing-in-place strategies is the design of the home environment. Environmental barriers and accessibility problems in the home or in the close neighborhood can indeed have a negative impact on activity and participation both inside and outside the home [15], which in turn may lead to reduced self-rated health and quality of life [16].

In Sweden, people experiencing activity limitations and dependence on other people due to the presence of physical environmental barriers in the home can apply to the municipality for a housing adaptation (HA) grant. The individual can submit a certification of his/her needs and an application for a HA grant, which the municipality can entirely or partly approve or reject. Unlike to what is the case in most countries, HA in Sweden are publicly funded, and regulated by the Housing Adaptations Act [17]. Approximately 73,000 HAs are funded in Sweden each year, to a total cost of 1.039 million Swedish kronor. Common adaptations granted were mounting of grab bars, installation of ramps, and adaptations in hygiene areas [18].

People applying for HA are a heterogeneous group, in terms of e.g., age, independence in activities of daily living, participation in social activities and living conditions [19,20]. Most of them are 70 years or older [18] with declining physical and cognitive capacity due to normal ageing, or diseases such as stroke and dementia. However, younger people with injuries or diseases belong to this population [20,21]. Though different in terms of sociodemographic, clinical, and functional characteristics, they share the common need for having their physical home environment adapted to live independently. While it is known that a home environment unfit for their needs can negatively impact their health [22], other factors related to the person, the activity and the home environment specifically associated with health-related quality of life among this population remains a largely uninvestigated area. Knowledge about such determinants might support the HA process by providing valuable information on those concomitant factors which could be addressed by the professionals working in the municipalities during their interventions to benefit this group of clients. This may influence the overall effectiveness of the HA intervention. To address the current research gaps, the aim of this study was to investigate which factors are associated with health-related quality of life among people applying for HAs in Sweden.

2. Materials and Methods

This study has a cross-sectional design. It draws on baseline data from a larger trial (ClinicalTrials.gov.NCT01960582), implemented among HA applicants in Southern Sweden [19], approved by the Ethical Review Board at Lund University (Dnr 2012/556).

2.1. Sampling and Participants

Three municipalities were included in the study. The selection is based on their number of inhabitants, their geographical dispersion, and their willingness to participate in the project. Individuals living in these municipalities applying for a HA were asked to participate in the study, if they met the following inclusion criteria: aged 20 years or older, living in the community and applying for a HA after contact with the municipality occupational therapist. Excluded were those living in sheltered housing and those not able to communicate in Swedish [19]. In total, 224 participants were included in this study.

2.2. Data Collection

Data were collected by occupational therapists during home visits before the start of the HA. The home visit took approximately 90 min and the data collection included both self-assessments and observations by means of standardized instruments and study specific questions. In two municipalities, the occupational therapists employed by the municipality gathered the data, and in one, the occupational therapist was employed by the project. They were especially trained in the methodology applied to the study to uniform the data collection process [19]. Data collection at baseline took place between March 2013 and December 2015.

2.3. Study Outcome Variables

The outcome variables of the study were measured using the EQ-5D, a standardized instrument developed by the EuroQol Group as a measure of health-related quality of life. The EQ-5D consists of a descriptive system assessing five dimensions of health, namely mobility, self-care, usual activities, pain/discomfort and anxiety/depression, and a vertical visual analogue scale, EQ VAS. In the version of the descriptive system used in this study, the EQ-5D-5L, the five dimensions are assessed on a five-graded ordinal scale [23]. The respondent's scoring obtained on these dimensions can be converted to a single summary index number reflecting their preference compared to other health profiles, with a higher index score indicating better health profile. Respondents are then asked to rate their overall health on the day of the interview on the vertical visual analogue scale, EQ VAS, from 0 to 100, where 0 indicates the worst and 100 the best imaginable health. The EQ-5D has been tested extensively for validity and reliability [24,25]. In this study, we used separately both the index score obtained from the descriptive system and the ratings from the visual analogue scale, as two distinct but relevant measures of health-related quality of life among HA clients.

2.4. Other Study Variables

- Demographic data included age, sex and living conditions classified as "Living alone and having no close relationship with a partner", "Living alone and having a close relationship with a partner", "Living together with partner/family" and "Other".
- Cognitive impairment was assessed using the Montreal Cognitive Assessment (MoCA), a screening tool including eight domains: short-term memory, visuospatial abilities, executive functions, attention, concentration, working memory, language and orientation in time and place. The maximum score is 30 points, with a lower score indicating a higher level of cognitive impairment. MoCA has been tested for validity and reliability [26].
- Usability of the home environment was assessed by a revised version of Usability In My Home (UIMH) instrument [27]. The revised version consists of 11 self- reported questions on how satisfied the individual is with the home environment in different activities, such as cleaning, cooking, leisure activities etc. There are five response alternatives for each question, reaching from "very unsatisfied" (1 point) to "very satisfied" (5 points). Those who did not perform the activity described in the item were coded with a 0. Higher sum score indicates higher satisfaction with the home usability. UIMH has been tested for validity and reliability [27], and a further validation is currently ongoing [28]. Three aspects of usability have been identified; "Self-care aspects" (compromising 5 items: going to the toilet, personal hygiene, preparing meals, preparing snacks, and moving around the home with/without a mobility device, score 5–25 points), "Social aspects" (compromising 3 items: socializing with family and friends in the home, contacting others via telephone/Skype, and watching TV/listening to radio, score 3–15 points) and "Leisure/outdoor aspects" (compromising 3 items: entering the house, picking up the mail, and engaging in hobbies and leisure activities in the home, 3–15 points) [20].
- ADL dependence was assessed using the ADL-staircase [29] compromising nine activities; feeding, mobility, using the toilet, dressing, bathing, cooking, transportation, shopping, and cleaning, and shows how independent the individual is in these situations. The response alternatives are "independent without difficulty", "independent with difficulty", "partly dependent" or "dependent". The responses were summarized in a total score, 0–27, with a higher score indicating more problems. The ADL-staircase has been tested for validity and reliability [29].
- Participation frequency and satisfaction were assessed by means of eight study specific statements related to: (a) how frequently and (b) how satisfied they were with the following: having contact

with others in their home, helping others, doing something outside their home with others, and doing something outside their home alone.
- Fear of falling was assessed with a dichotomous yes/no question.

2.5. Analysis Design and Statistics

Data on participant age showed a non-normal distribution and was therefore split into three age categories: category 1 (age < 75 years), category 2 (age 75–90 years) and category 3 (age > 90 years). Category 2 (75–90 years) was used as the reference category in the regression analysis.

Living conditions were reclassified into two response alternatives. "Living alone and having no close relationship with a partner", "Living alone and having a close relationship with a partner" and "Other" were classified as "Living alone". "Living together with partner/family" was classified as "Living together". The scores from MoCA, ADL-staircase and UIMH (all three aspects) were reclassified from continuous to categorical variables. This was done to describe and visualize differences between groups. MoCA scores were classified into moderate, mild, or normal, and ADL-staircase and UIMH scores were divided into quartiles.

Descriptive statistics included independent T-tests and ANOVA examining the variables associations to health-related quality of life as measured separately by the EQ-5D-5L and the EQ VAS. Variables potentially associated with the outcomes (*p*-value below 0.2 at bivariate level) were then included in the multiple regression analysis by stepwise-forward selection, starting with variables that significantly differed between groups in the bivariate analysis. The independent variables were gender, age, living conditions, cognitive impairment, ADL dependence and usability of the home. We also tested the influence of items related to frequency and satisfaction with participation, and fear of falling but since they showed no potential association with the outcomes of interest, these variables were not included in the final analysis.

Associations with health-related quality of life were analyzed with multiple linear regression analysis using robust standard errors [30]. We used the EQ-5D-5L and the EQ VAS separately as the dependent variables of the regressions. Level of significance was set to $p < 0.05$. Post-estimation diagnostics was conducted by checking normality of residuals, the link test and measuring the information criteria of the models. Statistical analysis was performed using SPSS Statistics 25.0 (IBM Corporation, Armonk, NY, USA) and Stata 15.0 (Statacorp, College Station, TX, USA, 2015).

3. Results

The average self-rated health for the whole sample using the EQ VAS was 56.23 ± 20.89, while the mean EQ-5D-5L index score was 0.54 ± 0.25. The relationship between health-related quality of life and exposure variables at bivariate level are shown in Table 1. In the bivariate analysis, significant differences in self-rated health measured using the VAS score were found in relation to age, ADL dependence, and all aspects of usability in the home. Health-related quality of life, both the index score and the VAS score, were lower among the population younger than 75 years, those with the higher dependence in ADL, and among those with lower usability scores. Those living alone had a significantly higher index score, indicating a better health profile, but on average a similar score for self-rated health on the VAS.

The regression models are presented in Table 2. The final analysis included 174 and 178 participants, for the VAS score and the index score variables, respectively. Results show that the variance in health-related quality of life (both scores) was significantly associated with age. More specifically, participants aged under 75 years rated their health lower compared to the group of those aged between 75 and 90 years (*p* values equal to 0.016 and 0.009 for the VAS score and the index score respectively). ADL dependence was also significantly associated with health-related quality of life (both scores); the more dependent participants were in ADL, the worse their ratings on the EQ-5D (*p* values equal to 0.011 and 0.001 for the VAS score and the index score respectively).

Table 1. Bivariate associations between selected variables and EQ-5D measures.

Variable	EQ VAS				EQ-5D-5L Index			
	n (%)	Mean	SD ±	*p* *	*n* (%)	Mean	SD ±	*p* *
Gender								
Male	82 (36.61)	55.32	22.9	0.621	87 (37.02)	0.51	0.26	0.1732
Female	142 (63.39)	56.75	19.69		148 (62.98)	0.56	0.25	
Age								
<75 years	85 (37.95)	49.79	22.93	**0.001**	87 (37.02)	0.45	0.28	**0.001**
75–90 years	117 (52.23)	60.02	19.24		123 (52.34)	0.59	0.22	
>90 years	22 (9.82)	60.95	14.58		25 (10.64)	0.61	0.23	
Living conditions								
Alone	124 (55.61)	56.99	19.49	0.593	132 (56.41)	0.57	0.25	**0.048**
Together	99 (44.39)	55.48	22.58		102 (43.59)	0.5	0.26	
Cognitive impairment [1]								
Missing value	43 (19.20)	57.88	20.98	0.644	50 (21.28)	0.56	0.25	0.512
10–17	28 (12.50)	52.32	20.88		27 (11.49)	0.49	0.34	
18–25	97 (43.30)	57.3	20.7		102 (43.40)	0.54	0.23	
26–30	56 (25.00)	55.05	21.39		56 (23.83)	0.57	0.25	
ADL dependence [2]								
0–8 (Q_1)	68 (30.36)	63.18	19.11	**0.006**	69 (29.36)	0.66	0.16	**<0.001**
9–11 (Q_2)	42 (18.75)	56.57	19.86		45 (19.15)	0.58	0.22	
12–15 (Q_3)	64 (28.57)	52.38	19.76		66 (28.09)	0.51	0.24	
16–27 (Q_4)	50 (22.32)	51.42	23.26		55 (23.40)	0.41	0.31	
Usability-self-care aspect [3]								
0–14 (Q_1)	58 (25.89)	47.83	23.28	**<0.001**	62 (26.38)	0.34	0.28	**<0.001**
15–19(Q_2)	62 (27.68)	53.39	17.73		66 (28.09)	0.56	0.23	
20–22 (Q_3)	52 (23.21)	59.37	17.62		54 (22.98)	0.64	0.16	
23–25 (Q_4)	52 (23.21)	65.85	20.51		53 (22.55)	0.67	0.16	
Usability-social aspect [3]								
0–10 (Q_1)	58 (25.89)	50.21	24.31	**0.002**	62 (26.81)	0.47	0.28	**0.014**
11–12 (Q_2)	54 (24.10)	53.04	16.79		55 (26.38)	0.53	0.26	
13–14 (Q_3)	61 (27.23)	57.8	21.25		63 (25.11)	0.56	0.23	
15 (Q_4)	51 (22.77)	64.57	17.44		55 (21.70)	0.62	0.22	
Usability-outdoor/leisure aspect [3]								
0–5 (Q_1)	59 (26.34)	49.97	21.58	**0.002**	63 (26.81)	0.46	0.27	**0.001**
6–8 (Q_2)	58 (25.89)	54.19	22.96		62 (26.38)	0.51	0.28	
9–11 (Q_3)	57 (25.45)	57.18	17.18		59 (25.11)	0.57	0.22	
12–15 (Q_4)	50 (22.32)	64.9	18.79		51 (21.70)	0.64	0.18	
Total sample	224	56.23	20.89		235	0.54	0.25	

[1] As measured and categorized by Montreal Cognitive Assessment: moderate = 10–17, mild = 18–25, normal > 26.
[2] As measured by ADL-staircase. [3] As measured by Usability In My Home. * Figures in bold are significant.

Table 2. Models for estimation of factors associated with EQ-5D measures.

Variable	EQ VAS (N = 174; R^2 = 0.221)				EQ-5D-5L index (N = 178; R^2 = 0.3792)			
	β	Std error	*p* *	(95% CI)	β	Std error	*p* *	[95% CI]
Age < 75 (with reference to age 75–90)	−7.68	3.15	**0.016**	(−13.90; −1.45)	−0.093	0.035	**0.009**	[−0.163; −0.023]
Age > 90 (with reference to age 75–90)	3.18	5.55	0.568	(−7.78; 14.14)	0.021	0.044	0.633	[−0.066; 0.109]
Sex (female)	−0.66	3.17	0.835	(−6.91; 5.59)	0.013	0.037	0.731	[−0.060; 0.085]
Living conditions (together)	0.28	2.96	0.924	(−5.56; 6.12)	−0.017	0.033	0.598	[−0.083; 0.048]
ADL dependence	−0.79	0.31	**0.011**	(−1.40; −0.19)	−0.013	0.003	**0.001**	[−0.019; −0.008]
Cognitive impairment	−0.15	0.32	0.636	(−0.80; 0.49)	0.000	0.004	0.947	[−0.009; 0.008]
Usability (self−care aspect)	0.37	0.38	0.33	(−0.38; 1.12)	0.014	0.004	**0.001**	[0.006; 0.023]
Usability (social aspect)	1.64	0.61	**0.008**	(0.44; 2.83)	0.003	0.008	0.676	[−0.012; 0.018]
Usability (Leisure/outdoor aspect)	0.31	0.42	0.47	(−0.53; 1.14)	0.004	0.004	0.393	[−0.005; 0.012]
Constant	42.35	12.97	**0.001**	(16.74; 67.96)	0.403	0.132	**0.003**	[0.142; 0.663]

β = Beta Coefficient; * Figures in bold are significant.

Finally, participants who were more satisfied with the social aspect of usability in their homes also rated their health better than those who were less satisfied, using the VAS score (p = 0.008) Conversely, the social aspect of usability in the home was not significantly associated with health-related quality of life when using the EQ-5D index score as dependent outcome measure. Similarly, the self-care aspect of usability was significantly associated with the EQ-5D index score (p = 0.001) but not the VAS score.

4. Discussion

The aim of this study was to investigate factors associated with health-related quality of life among people applying for HA in Sweden. For this purpose, the study used multiple regression analysis to test the influence of variables already known from previous research to have an impact on health-related quality of life, including personal factors, activities, participation, and environmental factors.

The findings conclude that three variables were significantly associated with health-related quality of life measured with the VAS in this sample: age, ADL dependence, and the social aspect of usability in the home. Variables significantly associated with the health summary index were age, ADL dependence and the self-care aspect of usability.

Age in our data had a skewed distribution, and thus, was reclassified into three categories. The youngest age category covered a wide range, 25–75 years, and included 27.95% of the participants. The regression analysis showed that this age group rated their health significantly lower than the other older groups; a fact, which was not discovered when age was used as a continuous variable. Thus, categorization helped clarify the relationship between age and self-rated health in this case. In the general population, younger people rate their health higher than older people [5], but in this study, those aged under 75 years rate their health lowest compared to other groups.

A potential explanation might be that their expectations on life in general are similar to others in the same age, but when comparing their own life with others they feel much more restricted and impacted by their activity restrictions, in this study demonstrated by their higher level of ADL dependence, and thus perceive their health as lower. When it comes to the older age group, they might perceive their activity limitations and dependence as a natural consequence of ageing, and thus it has less impact on their self-rated health compared to those that are younger [20].

In our study, higher dependency in ADL was associated with low health-related quality of life. Moreover, the self-care aspect of usability was associated with a lower index score. This is in line with earlier studies focusing on older people [31,32]. A combination of functional limitations and physical demands may have negative influence on the ability to perform different activities, and therefore increase dependency in ADL [33] seems to lead to lower self-rated health [22]. An earlier study exploring the reasons for applying for a HA [34] revealed that among other things, increasing independence in ADL and not having to rely on family, friends, and home care services to perform different activities was a major argument for an application.

The social aspect of usability consisted of three items: socializing with friends and family in the home, contacting others via telephone/computer, and watching TV/listening to radio. The findings showed that participants, who were more satisfied with this aspect of usability, also rated their health higher. Social contact and possibility to engage in social activities is related to higher self-rated health in several previous studies [35–37]. Some of the participants in this study are considered as frail elderly and might have limited possibility to engage in social activities outside the home, which makes the own home and its usability more important.

A limitation of the study is the low representation of people with cognitive impairment, due to a higher frequency of missing data among this population. Cognitive impairment could not be assessed for 19.2% of the sample, due to fatigue, low mood, or disability among these participants. Thus, it is likely that the prevalence of cognitive impairments among persons applying for HA in Sweden might be higher than what actually is observed in our sample. In addition, only people able to communicate in Swedish were included in this study. This could be due to disability, such as aphasia or dementia related language difficulty, but also due to cultural reasons. Earlier research has showed that older immigrants might rate their health worse than citizens born and raised in Sweden [38]. This limitation as well might partially affect the generalizability of the results to the general Swedish population but does not affect the internal validity of our results and the consistency of the associations found.

It is also important to note that ADL dependence, as measured by the ADL-staircase, was calculated into sum-scores. This differs from the original version of the instrument [36] but has been previously applied in other studies [20,39]. Most important though, ratings of difficulty in activity performance are included in this study, potentially generating detailed information on the individuals' dependence in ADL. Still, further validation of the scale is needed [40].

This study had a cross-sectional design and only investigated factors associated with health-related quality of life among people applying for HA. However, given the aim of the HA, it is possible that both health-related quality of life and the strength of these associations would change after the HA. Future research could involve a longitudinal design exploring how implementations of HA and/or a rejection of applications affect health-related quality of life and associated factors.

5. Conclusions

To conclude, a variety of factors seem to be associated with how people applying for HAs in Sweden perceive their health-related quality of life. Our study suggests that applicants with higher dependence in ADL and other groups, such as applicants with less than 75 years and/or those living in homes which are not well suited for supporting the activity related to self-care (in other terms, people reporting lower usability scores in the self-care component), could be at risk of poorer self-rated health. Our finding suggests that care professionals, such as the occupational therapist working in the municipalities in charge of providing the HA grants, should pay special attention towards the situation of people with these characteristics, in order to reduce the observed gap in health-related quality of life both through personalization of care and the provision of integrative services. Such approach would be beneficial also in terms of prevention of a further health decline, as the literature strongly support that differences in health-related quality of life are strongly associated with further increase of disability and disease-related care burden.

Author Contributions: L.E. and A.M.F. designed the study. L.E. coordinated the data collection. L.B. and C.C. performed the statistical analyzes in dialogue with all authors. L.B. and B.T. took responsibility for writing the paper and all the authors interpreted data and critically revised the article. All the authors approved the final version of the manuscript.

Funding: This research was funded by the Swedish Research Council FORMAS and the Faculty of Medicine at Lund University. The study was based on data from a quasi-experimental study of housing adaptations in Sweden. It was accomplished within the context of the Centre for Ageing and Supportive Environment (CASE), Lund University, financed by the Swedish Research Council for Health, Working Life and Welfare, FORTE.

Acknowledgments: Sincere thanks are extended to all the participants in the study and the municipality occupational therapists for collecting data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fayers, P.M.; Sprangers, M.A.G. Understanding self-rated health. *Lancet* **2002**, *359*, 187–188. [CrossRef]
2. Perenboom, R.J.; Wijlhuizen, G.J.; Garre, Y.F.; Heerkens, Y.F.; van Meeteren, N.L. An empirical exploration of the relations between the health components of the International Classification of Functioning, Disability and Health (ICF). *Disabil. Rehabil.* **2012**, *34*, 1556–1561. [CrossRef] [PubMed]
3. DeSalvo, K.B.; Bloser, N.; Reynolds, K.; He, J.; Muntner, P. Mortality prediction with a single general self-rated health question. *J. Gen. Intern. Med.* **2006**, *21*, 267–275. [CrossRef] [PubMed]
4. Idler, E.L.; Benyamini, Y. Self-rated health and mortality: A review of twenty-seven community studies. *J. Health Soc. Behav.* **1997**, *38*, 21–37. [CrossRef] [PubMed]
5. Lubetkin, E.I.; Jia, H.; Franks, P.; Gold, M.R. Relationship among sociodemographic factors, clinical conditions, and health-related quality of life: Examining the EQ-5D in the U.S. general population. *Qual. Life Res.* **2005**, *14*, 2187–2196. [CrossRef] [PubMed]

6. Maharlouei, N.; Akbari, M.; Khabbaz Shirazy, M.; Yazdanpanah, D.; Lankarani, K.B. Factors associated with self-rated health status in Southwestern Iran: A population-based study. *Public Health* **2016**, *140*, 179–185. [CrossRef] [PubMed]
7. Kameyama, K.; Tsutou, A.; Fujino, H. The relationship between health-related quality of life and higher-level functional capacity in elderly women with mild cognitive impairment. *J. Phys. Ther. Sci.* **2016**, *28*, 1312–1317. [CrossRef] [PubMed]
8. Holmgren, M.; Lindgren, A.; de Munter, J.; Rasmussen, F.; Ahlström, G. Impacts of mobility disability and high and increasing body mass index on health-related quality of life and participation in society: A population-based cohort study from Sweden. *BMC Public Health* **2014**, *14*, 381. [CrossRef] [PubMed]
9. Törnvall, E.; Marcusson, J.; Wressle, E. Health-related quality of life in relation to mobility and fall risk in 85-year-old people: A population study in Sweden. *Ageing Soc.* **2016**, *36*, 1982–1997. [CrossRef]
10. Bernfort, L.; Eckard, N.; Husberg, M.; Alwin, J. A case of community-based fall prevention: Survey of organization and content of minor help services in Swedish municipalities. *Scand. J. Public Health* **2014**, *42*, 643–648. [CrossRef] [PubMed]
11. Pighills, A.C.; Torgerson, D.J.; Sheldon, T.A.; Drummond, A.E.; Bland, J.M. Environmental Assessment and Modification to Prevent Falls in Older People. *JAGS* **2011**, *29*, 26–33. [CrossRef] [PubMed]
12. Nachreiner, N.M.; Findorff, M.J.; Wyman, J.F.; McCarthy, T.C. Circumstances and Consequences of Falls in Community-Dwelling Older Women. *J. Womens Health* **2007**, *16*, 1437–1446. [CrossRef] [PubMed]
13. Johansson, K.; Josephsson, S.; Lilja, M. Creating possibilities for action in the presence of environmental barriers in the process of 'ageing in place'. *Ageing Soc.* **2009**, *29*, 49–70. [CrossRef]
14. Hwang, E.; Cummings, L.; Sixsmith, A.; Sixsmith, J. Impacts of Home Modifications on Aging-in-Place. *J. Hous. Elderly* **2011**, *25*, 246–257. [CrossRef]
15. Rantakokko, M.; Iwarsson, S.; Portegijs, E.; Viljanen, A.; Rantanen, T. Associations Between Environmental Characteristics and Life-Space Mobility in Community-Dwelling Older People. *J. Aging Health* **2015**, *27*, 606–621. [CrossRef] [PubMed]
16. Liu, J.Y.W. The severity and associated factors of participation restriction among community-dwelling frail older people: An application of the International Classification of Functioning, Disability and Health (WHO-ICF). *BMC Geriatr.* **2017**, *17*, 43. [CrossRef] [PubMed]
17. Lag om Bostadsanpassningsbidrag m.m. (SFS 1992:1547). Socialdepartementet. (Swedish Act on Housing Adaptation Grants. In Swedish). Available online: https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/lag-19921574-om-bostadsanpassningsbidrag-m-m_sfs-1992-1574 (accessed on 17 April 2017).
18. Boverket. *Bostadsanpassningsbidragen 2015*; Boverket: Karlskrona, Sweden, 2016.
19. Ekstam, L.; Carlsson, G.; Chiatti, C.; Nilsson, M.H.; Malmgren Fänge, A. A research-based strategy for managing housing adaptations: Study protocol for a quasi-experimental trial. *BMC Health Serv. Res.* **2014**, *14*, 602. [CrossRef] [PubMed]
20. Thordardottir, B.; Chiatti, C.; Ekstam, L.; Malmgren Fänge, A. Heterogeneity of Characteristics among Housing Adaptation Clients in Sweden—Relationship to Participation and Self-Rated Health. *Int. J. Environ. Res. Public Health* **2016**, *13*, 91. [CrossRef] [PubMed]
21. Carlsson, G.; Nilsson, M.H.; Ekstam, L.; Chiatti, C.; Malmgren Fänge, A. Falls and Fear of Falling among Persons Who Receive Housing Adaptations—Results from a Quasi-Experimental Study in Sweden. *Healthcare* **2017**, *5*, 66. [CrossRef] [PubMed]
22. World Health Organization (WHO); The World Bank. *World Report on Disability*; WHO: Geneva, Switzerland, 2011.
23. Herdman, M.; Gudex, C.; Lloyd, A.; Janssen, M.F.; Kind, P.; Parkin, D.; Bonsel, G.; Badia, X. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual. Life Res.* **2011**, *20*, 1727–1736. [CrossRef] [PubMed]
24. Van Reenen, M.; Janssen, B. EQ-5D-5L User Guide. EuroQol Research Foundation. Version 2.1. Available online: https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-5L_UserGuide_2015.pdf (accessed on 4 June 2017).
25. Conner-Spady, B.L.; Marshall, D.A.; Bohm, E.; Dunbar, M.J.; Loucks, L.; Al Khudairy, A.; Noseworthy, T.W. Reliability and validity of the EQ-5D-5L compared to the EQ-5D-3L in patients with osteoarthritis referred for hip and knee replacement. *Qual. Life Res.* **2015**, *24*, 1775–1784. [CrossRef] [PubMed]

26. Nasreddine, Z.S.; Phillips, N.A.; Bédirian, V.; Charbonneau, S.; Whitehead, V.; Collin, I.; Cummings, J.L.; Chertkow, H. The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool for Mild Cognitive Impairment. *J. Am. Geriatr. Soc.* **2005**, *53*, 695–699. [CrossRef] [PubMed]
27. Fänge, A.; Iwarsson, S. Physical housing environment: Development of a self-assessment instrument. *Can. J. Occup. Ther.* **1999**, *66*, 250–260. [CrossRef] [PubMed]
28. Chiatti, C.; Ekstam, L.; Malmgren Fänge, A. Validation of the Revised Version of the Usability in My Home Scale. Unpublished manuscript in progress. 2018.
29. Sonn, U.; Åsberg, K.H. Assessment of activities of daily living in the elderly. A study of a population of 76-year-olds in Gothenburg, Sweden. *Scand. J. Rehabil. Med.* **1991**, *23*, 193–202. [PubMed]
30. Chan, Y.H. Biostatistics 201: Linear Regression Analysis. *Singap. Med. J.* **2004**, *45*, 55–61.
31. Iwarsson, S.; Isacsson, Å. Quality of life in the elderly population: An example exploring interrelationships among subjective well-being, ADL dependence, and housing accessibility. *Arch. Gerontol. Geriatr.* **1997**, *26*, 71–83. [CrossRef]
32. Paskulin, L.; Vianna, L.; Molzahn, A.E. Factors associated with quality of life of Brazilian older adults. *Int. Nurs. Rev.* **2009**, *56*, 109–115. [CrossRef] [PubMed]
33. Iwarsson, S.; Isacsson, Å.; Lanke, J. ADL dependence in the elderly population living in the community: The influence of functional limitations and physical environmental demand. *Occup. Ther. Int.* **1998**, *5*, 173–193. [CrossRef]
34. Ekstam, L.; Malmgren Fänge, A.; Carlsson, G. Negotiating Control: From Recognizing a Need to Making a Decision to Apply for a Housing Adaptation. *J. Hous. Elderly* **2016**, *30*, 345–359. [CrossRef]
35. McGregor, M.J.; Slater, J.; Sloan, J.; McGrail, K.M.; Martin-Matthews, A.; Berg, S.; Plecash, A.; Sloss, L.; Trimble, J.; Mruphy, J.M. How's Your Health at Home: Frail Homebound Patients Reported Health Experience and Outcomes. *Can. J. Aging* **2017**, *36*, 273–285. [CrossRef] [PubMed]
36. Nummela, O.; Sulander, T.; Karisto, A.; Uutela, A. Self-rated health and Social Capital Among Aging People Across the Urban-Rural Dimension. *Int. J. Behav. Med.* **2009**, *16*, 189–194. [CrossRef] [PubMed]
37. Nyqvist, F.; Nygård, M.; Steenbeek, W. Social Capital and Self-rated Health Among Older People in Western Finland and Northern Sweden: A Multi-level Analysis. *Int. J. Behav. Med.* **2014**, *21*, 337–347. [CrossRef] [PubMed]
38. Emami, A.; Benner, P.E.; Lipson, J.G.; Ekman, S.L. Health as Continuity and Balance in Life. *West. J. Nurs. Res.* **2000**, *22*, 812–825. [CrossRef] [PubMed]
39. Thordardottir, B.; Ekstam, L.; Chiatti, C.; Malmgren Fänge, A. Factors associated with participation frequency and satisfaction among people applying for a housing adaptation grant. *Scand. J. Occup. Ther.* **2016**, *23*, 347–356. [CrossRef] [PubMed]
40. Iwarsson, S.; Horstmann, V.; Sonn, U. Assessment of dependence in daily activities combined with a self-rating of difficulty. *J. Rehabil. Med.* **2009**, *41*, 150–156. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

Parenting Desire and Minority Stress in Lesbians and Gay Men: A Mediation Framework

Anna Lisa Amodeo [1], Concetta Esposito [2], Vincenzo Bochicchio [3], Paolo Valerio [4], Roberto Vitelli [4], Dario Bacchini [1] and Cristiano Scandurra [4,*]

1 Department of Humanistic Studies, University of Naples Federico II, 80138 Napoli, Italy; amodeo@unina.it (A.L.A.); dario.bacchini@unina.it (D.B.)
2 SInAPSi Center, University of Naples Federico II, 80138 Napoli, Italy; concetta.esposito3@unina.it
3 Department of Humanistic Studies, University of Calabria, 87036 Rende, Italy; vincenzo.bochicchio@unical.it
4 Department of Neurosciences and Reproductive and Odontostomatological Sciences, University of Naples Federico II, 80138 Napoli, Italy; valerio@unina.it (P.V.); rvitelli@unina.it (R.V.); cristiano.scandurra@unina.it (C.S.)
* Correspondence: cristiano.scandurra@unina.it; Tel.: +39-334-152-3239

Received: 26 September 2018; Accepted: 19 October 2018; Published: 22 October 2018

Abstract: Despite the rapid increase in lesbian and gay (LG) people who desire and decide to become parents, LG childless individuals may encounter serious obstacles in the parenthood process, such as minority stress. Notwithstanding, the psychological processes by which prejudice events might affect the desire to become parents are still understudied. As an extension of the minority stress theory, the psychological mediation framework sheds light on these psychological processes, as it encompasses a more clinical view of stress. Within this framework, the current study aimed at assessing the role of prejudice events in affecting parenting desire in 290 childless Italian LG individuals (120 lesbians and 170 gay men), as well as the role of internalized heterosexism and sexual orientation concealment in mediating the relationship between prejudice events and parenting desire. The results suggest that only in lesbians prejudice events were negatively associated with parenting desire, and that sexual orientation concealment and internalized heterosexism were also negatively associated with parenting desire. Furthermore, sexual orientation concealment, and not internalized heterosexism, mediated the relationship between prejudice events and parenting desire in lesbians, but not gay men. The findings have important implications for clinical practice.

Keywords: parenting desire; lesbian; gay; minority stress; mediation

1. Introduction

Becoming parents represents a complex process that might be influenced by many social dimensions, such as economic issues, familial policies, local legislation, or housing conditions [1]. From a psychological point of view, a starting point in the parenthood path is the desire to become a parent, or rather "what one wants or would like to do" [2] (p. 10).

Despite the rapid increase in lesbian and gay (LG) people who desire and decide to become parents, differently from heterosexual people, LG childless individuals may encounter serious obstacles in the parenthood process due to their sexual minority status [3–5]. One of these is represented by minority stress, a specific stress caused by socially stigmatized status and associated with negative health outcomes [6,7]. Minority stress theory (MST) conceptualizes both distal and proximal stressors. Distal stressors are objective stressors (i.e., prejudice events, such as interpersonal violence, employment discrimination, mistreatment in healthcare setting, etc.), that are independent of the individual's perceptions or feelings. On the other hand, proximal stressors are subjective and internal stressors dependent on individual's perceptions, such as expectations of rejection, sexual orientation concealment, and internalized heterosexism, that is, the direction of negative social attitudes toward the self. MST assumes that both stressors assumes that objective stressors (i.e., prejudice events, such as violence and discrimination), as well as subjective stressors (i.e., expectations of rejection, sexual orientation concealment, and internalized homophobia), predict negative health outcomes, and that specific protective factors (e.g., resilience, community connectedness, social support, etc.) buffer this direct association [7,8]. Recently, some studies have demonstrated that minority stressors negatively affect parenting desire, thus resulting in a significant psychosocial barrier to parenthood [9–11].

A recent extension of MST—the psychological mediation framework (PMF)—places the accent on a more clinical and subjective view of the stress, thanks to the use of proximal stressors as mediators between prejudice events and health [12]. The PMF was postulated to analyze psychological paths linking stigma-related stressors to negative health outcomes. Thus, differently from MST, according to which stress is a mediator between social structure/status and health, the PMF looks at stress as the starting point of a mediating psychological chain leading to negative health outcomes. Psychological mediators are both the proximal minority stressors (e.g., internalized heterosexism and sexual orientation concealment) and some general psychological processes (e.g., interpersonal problems or emotion dysregulation). In summary, if MST considers both distal and proximal minority stressors as predictors of health outcomes, the PMF considers the proximal minority stressors as mediators between distal minority stressors and health.

Although recent studies have demonstrated the usefulness of approaching parenting dimensions as potential outcomes of the minority stress processes, going beyond the health outcomes [9–11], no previous studies have applied the PMF to parenting dimensions in LG individuals. As parenting desire represents the first link in the chain of the parenting process, and internalized heterosexism and sexual orientation concealment are the two main personal issues shaping parenting decisions [13], we were interested in analyzing the role of internalized heterosexism and sexual orientation concealment as mediators between prejudice events and parenting desire in childless LG individuals.

Furthermore, as gender is a fundamental dimension of parenthood [14], the difference between women and men was considered as a potential moderator. Indeed, although the old sexist stereotypes have been overcome in contemporary society, some different-sex couples still divide their childcare tasks or household chores based on gender norms, with women dedicated to these tasks and men employed outside the home [15]. On the contrary, most of the research on same-sex couples reported that they are more likely than different-sex couples to divide the labor fairly [16]. These generalizations must not lead to the perception of same-sex couples as more functioning than different-sex couples, but serve to show that gender might play a key role in the couple's dynamics and that it should be considered as a crucial factor in studies exploring parenthood dimensions. The article begins by providing an overview of the relationship between minority stressors and parenting desire in LG individuals. Then, it gives an overview of the internalized heterosexism and sexual orientation concealment as mediators, according to the PMF. Finally, as Italy represents the context of our study, it introduces a snapshot of the Italian context for LG individuals.

1.1. Minority Stressors and Parenting Desire in Lesbians and Gay Men

LG individuals have been found to be generally resilient in the face of stigma and to resist heterosexist social pressure [17,18]. In spite of this, they still experience prejudice events due to their stigmatized social status, thus experiencing high rates of minority stress and negative health outcomes [17,18]. When childless LG individuals become parents, they also become extremely visible and, because of this, minority stress might increase [11].

Recent studies have shown the association between minority stress and parenting dimensions. For instance, Bos et al. [10] found that lesbian mothers experiencing higher levels of prejudice events showed more parental stress and, at the same time, felt more pressed to justify their motherhood qualities to people than mothers experiencing fewer prejudice events. Similarly, mothers with higher levels of internalized heterosexism tended to defend their position as mothers more often than those with lower levels of internalized heterosexism. Another study, by Baiocco et al. [9], found that childless Italian LG adults with higher levels of internalized heterosexism were less likely to desire to marry, and to recognize some positive effects of same-sex legal recognition, than those with lower levels of internalized heterosexism. More recently, an Italian study by Scandurra et al. [11], in which MST was applied to a group of childless LG individuals, reported that prejudice events, sexual orientation concealment, and internalized heterosexism were negatively associated with parenting desire in lesbians, and that felt stigma negatively impacted parenting desire in gay men. Furthermore, support from family or significant others buffered the relationship between minority stressors and parenting desire.

These data seem to confirm that one of the obstacles that LG individuals may encounter in their parenthood path is heterosexism [4,5], which, in turn, leads to experiencing minority stress because sexual minority people do not match heterosexist expectations. Indeed, as suggested by Mezey [13], due to the heterosexist society in which LG individuals live, two fundamental personal and psychological dimensions that have to be considered in analyzing the parenthood process are internalized heterosexism and sexual orientation concealment.

LG individuals internalize societal values and messages that communicate that nonheterosexual orientations are immoral, that LG individuals cannot be good parents, and that children born to a same-sex couple will not grow up well [4,19,20]. Due to internalized heterosexism, childless LG individuals may come to question their ability to become parents, creating barriers to parenthood [13]. Sexual orientation concealment, which is the negative side of the coming out process, is also connected to internalized heterosexism. Revealing one's own nonheterosexual orientation to family members, friends, and colleagues, may have positive effects on the quality of relationships and mental health [21,22]. At the same time, it is often a difficult and pained process, which depends on several factors, such as age, class, race, and environment [19]. Indeed, this process implies the necessity to negotiate multiple identities in one's own environment, and this is particularly true for LG individuals who desire or want to become parents [13]. For instance, the fear that some LG individuals experience in coming out to their parents might represent a serious obstacle to the parenthood process [23]. Finally, despite similarities in the coming out process among lesbians and gay men, previous studies [24,25] have detected an interesting difference based on gender: many lesbians tend to perceive coming out as a necessary step to becoming mothers, while many gay men tend to regard coming out as a barrier to becoming fathers. Thus, for some gay men, coming out represents a break between being gay or being a father [26]. Substantially, it seems that this gender difference might be explained through the social expectations regarding becoming a parent, which are greater for women than for men [27]. For this reason, the individual's gender should be considered as a potential moderator in parenting studies.

1.2. Internalized Heterosexism and Sexual Orientation Concealment According to the Psychological Mediation Framework

As mentioned above, starting from a clinical view of stress and with the aim of extending the MST, the PMF was postulated as a theoretical framework to better understand psychological pathways that link minority stressors to negative health outcomes. Previous studies have empirically supported the validity of the PMF, both in LG [28,29] and transgender and gender nonconforming [30–32] people.

Specifically, and considering the main dimensions of the current study, there is evidence that internalized heterosexism and sexual orientation concealment mediate the relationship between prejudice events and health. For instance, Feinstein et al. [33] found that internalized heterosexism mediated the association between experiences of discrimination and both depression and social anxiety in both lesbians and gay men.

On the other hand, Ryan et al. [34] found that negative reactions to disclosure, which might be viewed as a form of prejudice event, were associated with depression and low self-esteem, and that autonomy need satisfaction following disclosure was a mediator between negative reactions and health. It is noteworthy that sexual orientation disclosure is not always beneficial, both because LG individuals could experience negative reactions from their interpersonal contexts and due to the inherently stressful process [35,36].

To our knowledge, the PMF has not yet been applied to parenting dimensions (e.g., parenting desire), although previous studies have considered these dimensions as potential outcomes of MST [9–11].

1.3. Italian Context for Lesbians and Gay Men

The Italian government has only recently recognized same-sex civil unions (law No. 76/2016), specifically in June 2016, and only after a scathing battle between conservative parties, supported by strongly religious groups, and moderate and left-wing parties. Before final approval of the original law, which foresaw the same benefits of different-sex marriage to same-sex couples, the so-called stepchild adoption (i.e., the possibility of adopting the biological child of a partner) was removed. This means that the only public body that can decide in favor of the adoption of a partner's biological child is the Supreme Court of Appeal, which has already provided this possibility to some Italian same-sex couples. Notwithstanding, contrary to different-sex couples, Italian same-sex couples cannot legally adopt children.

It is noteworthy that the social and political debate that led to law No. 76/2016 was specifically around the parental abilities of LG people, who were perceived by the opposition parties to be inadequate parents [11]. Along the same lines, a recent study by Lasio et al. [37], analyzing the speeches of Parliamentarians who expressed their opposition to LG parenthood, revealed that discourses were organized around a hegemonic model of gender, and that they contributed to reiterating old models of motherhood, maintaining the institutionalization of sexuality and reproduction within a patriarchal logic. These findings seem to be supported by some recent Italian studies, in which it has been reported that Italy is a strongly heteronormative [20,38–42] and genderist [43,44] society that still tends to equate women to mothers. Thus, social status differences continue to exist in Italy, although a great amount of progress has been made in the last few years.

Social status difference represents the most important dimension that leads to the development of minority stress, as it highlights the existence of social inequalities that, in turn, might be internalized by minority groups as a distinctive sign of their own identity. Indeed, previous studies underlined the presence of high levels of internalized heterosexism in Italian LG population [39], as well as in LG individuals who were in the reproductive age [9]. Unfortunately, despite the positive legal progress, Italian LG individuals still experience high rates of minority stressors [45], which, in turn, negatively affect parenting desire [11].

1.4. The Current Study

This study was aimed at filling a gap in the literature by applying the PMF to parenting desire in childless Italian LG individuals. Previous studies have already demonstrated that MST is a useful perspective for understanding psychosocial obstacles to parenting desire [9–11]. On the contrary, to our knowledge, although the PMF is an extension of MST, no previous studies have applied this framework to parenting desire.

Informed by MST, we first hypothesized that prejudice events could be negatively associated with parenting desire and, based on previous Italian studies highlighting the heteronormative social context for Italian LG individuals [20,38–42], that this association might be stronger in childless lesbian women than in childless gay men (Hypothesis 1). Second, informed both by the PMF and previous Italian studies [20,38–42], we hypothesized that the most subjective minority stressors (i.e., sexual orientation concealment and internalized heterosexism) would be negatively linked to parenting desire, with a stronger effect in lesbian women (Hypothesis 2). Third, informed by the PMF, we expected that both sexual orientation concealment and internalized heterosexism would act as mediators between prejudice events and parenting desire, and that these indirect effects would be dependent on an individual's gender (woman vs. man) (Hypothesis 3). Finally, as sociodemographic factors might influence both minority stressors and parenting desire [3,11,46,47], we considered the possible confounding effect of the following dimensions, age, education level, political orientation, and having a partner.

2. Materials and Methods

2.1. Participants

A total of 290 childless LG individuals (120 lesbians and 170 gay men), recruited online, participated in the study. The total sample ranged in age from 18 to 50 years (lesbians, M = 28.25, SD = 6.14; gay men, M = 33.11, SD = 9.64). Participants could participate in the online survey only if they self-identified with lesbian or gay sexual orientation, were of the Italian age of consent (i.e., 18-years old), did not already have children, and had lived in Italy for at least 10 years. Nineteen participants were excluded due to uncompleted questionnaires. Demographic characteristics for both the total sample and the sample differentiated by gender are reported in Table 1.

Table 1. Sociodemographic characteristics of participants (N = 290).

Characteristics	Total (N = 290) n (%) or M ± SD	Lesbians (n = 120) n (%) or M ± SD	Gay men (n = 170) n (%) or M ± SD	p-Value
Age	31.10 ± 8.69	28.25 ± 6.14	33.11 ± 9.64	<0.001
Education				
≤ High school	116 (40)	57 (47.5)	59 (34.7)	0.028
≥ College	174 (60)	63 (52.5)	111 (65.3)	
Political orientation *				
Left-wing	226 (77.9)	87 (72.5)	139 (81.8)	0.061
Centrist	64 (22.1)	33 (27.5)	31 (18.2)	
Stable partner				
Yes	172 (59.3)	71 (59.2)	101 (59.4)	0.967
No	118 (40.7)	49 (40.8)	69 (40.6)	

Note: M = Mean; SD = Standard Deviation; Group differences in age were assessed through the Student's t-test; Group differences in all other characteristics were assessed through the χ^2 test. * No participant declared to be right-wing.

2.2. Procedures

Participants answered the questionnaires at the end of 2016, when the law on same-sex civil unions was yet to be approved. The survey was disseminated through the main social networks (e.g., Facebook), posting the advertisement on many LG groups. Furthermore, some Italian non-governmental organization (NGO) representatives involved in the defense of lesbian, gay, bisexual, transgender, and queer rights helped us to carry out a snowball sampling recruitment, asking their affiliates to answer the questionnaire and to disseminate the survey to their personal contacts. No rewards were provided to participate in the survey.

The study was designed with regard to all of the principles of the Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects. According to the EU Regulation 2016/679 "General Data Protection Regulation" (GDPR) and the Code of Conduct for Research in Psychology of the Italian Association of Psychology approved on 27th March 2015, the participants gave their informed consent before completing the questionnaire. Furthermore, in accordance with Italian legislation, privacy was guaranteed through a secure gateway, accessible only to the Principal Investigator, who removed all of the IP addresses before sharing any data with the other researchers.

2.3. Measures

2.3.1. Sociodemographic Characteristics

Sociodemographic variables included gender (male, female, other with specification required), sexual orientation, age, level of education (≤high-school and ≥college), political orientation (left-wing, centrist, right-wing), and actual stable partner (yes/no).

2.3.2. Parenting Desire

A single-item question based on the work of Baiocco and Laghi [48] was used to determine information about parenting desire: "Would you like to have a baby?". The response options were dichotomous (yes/no). In the current sample, 95 (79.2%) lesbians and 123 (72.4%) gay men responded positively to the question.

2.3.3. Prejudice Events

The Experiences of Discrimination Scale (EDS) [49,50] is an eight-item measure for assessing four prejudice events: verbal abuse (e.g., "I heard jokes or unpleasant or derogatory comments about my sexual orientation"), unequal treatment (e.g., "Because of my sexual orientation I have not been able to get something important to me, for example a grant, a job"), avoidance (e.g., "It happened to me that some people have avoided me because of my sexual orientation"), and victimization (e.g., "I was physically assaulted because of my sexual orientation"). The response options ranged from "never" to "often" on a five-point Likert scale. Higher scores indicate more prejudice events. The internal consistency reliability was α = 0.84.

2.3.4. Internalized Heterosexism

The Measure of Internalized Sexual Stigma for Lesbians and Gay Men (MISS-LG) [39] is a 17-item questionnaire that assesses the negative attitudes of LG people towards homosexuality and specific aspects of themselves. A final score is achieved through three dimensions: (1) identity, or rather the enduring propensity to have negative self-attitudes (e.g., "I'd prefer to be heterosexual"); (2) social discomfort, that is the fear of publicly identifying as a LG individual, disclosure at work and in one's personal life, and negative beliefs about the religious, political, and moral acceptability of nonheterosexual orientation (e.g., "I'm careful of what I wear and what I say to avoid showing my homosexuality"); and (3) sexuality, or rather negative attitudes towards other LG individuals, and negative evaluation of same-sex relationships and sexual behaviors in LG people (e.g., "I don't believe in love between homosexuals"). The response options ranged from "I disagree" to "I agree" on a five-point Likert scale. The internal consistency reliability of the entire scale was α = 0.80 for the lesbian form and α = 0.81 for the gay man form. We used the total score of the scale, as done by other authors [51]. Higher scores indicate greater internalized heterosexism.

2.3.5. Sexual Orientation Concealment

The Outness Inventory (OI) [39,52] is an 11-item scale for assessing the degree of openness about an LG individual's sexual orientation with 11 individuals or groups (i.e., mother, father, siblings, extended family, new and old straight friends, work peers and supervisors, members and leaders of religious communities, and strangers). This measure included three scales: family, world, and religion. Overall outness was calculated as the subscales average. The response options ranged from "person definitely does not know about your sexual orientation status" to "person definitely knows about your sexual orientation status, and it is openly talked about" on a seven-point Likert scale. Scores were recoded so that higher scores indicated greater sexual orientation concealment. In the current study, the internal consistency reliability for the entire scale was α = 0.80.

2.4. Statistical Analyses

Analyses were carried out using IBM SPSS statistics version 21 (IBM Corp., Armonk, New York, NY, USA). The PROCESS macro version 3.1 [53] was used to test the study's hypothesized model, which included both a moderation (i.e., the individual's gender) and a mediation (i.e., sexual orientation concealment and internalized heterosexism) component.

A total effect moderation model (model template 59 in the PROCESS macro) was first specified. In this model, the effect of prejudice events (independent variable) on parenting desire (binary dependent variable; direct effect), and on sexual orientation concealment and internalized heterosexism (mediators), as well as the effect of both sexual orientation concealment and internalized heterosexism on parenting desire, were supposed to be conditional (or moderated) by the individual's gender (woman vs. man); however, estimation of the coefficients in the statistical model revealed no significant interaction between prejudice events and gender on both sexual orientation concealment and internalized heterosexism (both *ps* > 0.05). Since nonsignificant interactions also influenced the estimate of the hypothesized indirect effects (which were necessarily conditional with these interactions in the model), along with all inferential tests thereof, they were trimmed in the final moderated mediation model (Figure 1), and were thus expressed by only three interaction terms: (1) gender by prejudice event interaction on parenting desire; (2) gender by sexual orientation concealment interaction on parenting desire; and (3) gender by internalized heterosexism interaction on parenting desire (model template 15 in PROCESS macro).

Significant conditional direct and indirect effects were probed by using the pick-a-point approach. As a final step, in order to quantify the indirect effects as a function of the moderator, and to provide inferential tests for those conditional indirect effects, indices of moderated mediation were estimated for each of the hypothesized moderated mediation paths, as recommended by Hayes [53]. Bias-corrected bootstrap confidence intervals, based on 5000 resamples, were used as indicators of effect size. Confidence intervals that did not contain zero indicated a significant indirect effect via the specific mediator.

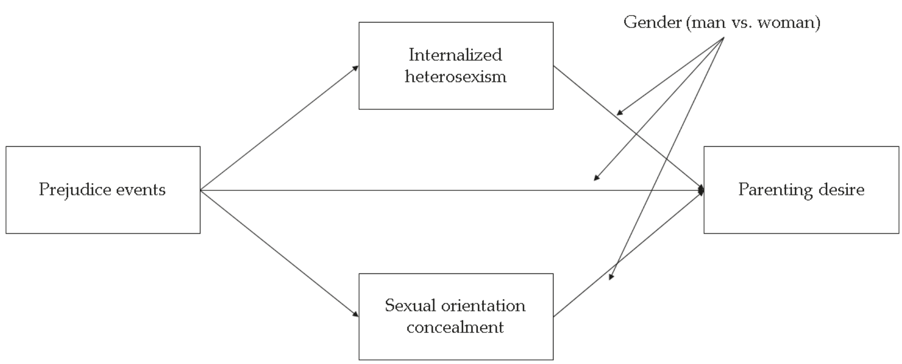


Figure 1. The hypothesized moderated mediation model. For reasons of simplification, control variables were not reported in the figure.

3. Results

3.1. Descriptive Statistics and Preliminary Analyses

The means, standard deviations, and bivariate correlations between all variables, separately for lesbians and gay men, are reported in Table 2. The results highlighted a significant positive association only between parenting desire and internalized heterosexism in lesbians; no significant associations resulted for gay men. Age was negatively correlated with internalized heterosexism, and positively associated with sexual orientation concealment, in both LGs. Having a stable partner was associated with internalized heterosexism in lesbians, whilst sexual orientation concealment was associated with the absence of a stable partner in both samples.

Table 2. Descriptive statistics and bivariate correlations between minority stressors, parenting desire, and sociodemographic characteristics.

	1	2	3	4	5	6	7	8
1. Prejudice events	-	−0.04	0.12	0.05	−0.14	−0.07	−0.04	−0.02
2. Internalized heterosexism	0.18 *	-	−0.39 ***	0.04	−0.22 **	−0.12	−0.03	0.21 **
3. Sexual orientation concealment	−0.01	−0.45 ***	-	0.07	0.21 **	−0.02	−0.13	−0.20 *
4. Parenting desire	−0.16	0.18 *	0.14	-	−0.15 *	−0.01	−0.08	−0.08
5. Age	−0.23 *	−0.24 **	0.19 *	−0.12	-	0.07	−0.09	−0.15
6. Education (≤ High school)	−0.06	−0.13	0.09	−0.08	0.38 ***	-	−0.04	−0.13
7. Political orientation (left-wing)	−0.09	−0.05	−0.05	0.04	−0.10	−0.16	-	−0.14
8. Stable partner (yes)	0.09	0.12	−0.26 **	−0.20 *	0.01	0.01	0.06	-

Note: Lesbians scores are below diagonal and gay men scores are above diagonal;*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

As a further preliminary step, we examined gender differences on prejudice events, sexual orientation concealment, and internalized heterosexism by performing a multivariate analysis of variance, including age, education, political orientation, and stable partner as covariates (MANCOVA). The analysis showed significant differences between lesbians and gay men, with Wilks' lambda = 0.92, $p < 0.001$, F (3282) = 8.37, $\eta^2 = 0.08$. Specifically, gay men reported higher levels of prejudice events—F (1284) = 16.59, $p < 0.001$, $\eta^2 = 0.06$—and internalized heterosexism—F (1284) = 4.20, $p < 0.05$, $\eta^2 = 0.01$—than lesbians did. No significant difference was found based on sexual orientation concealment.

3.2. Associations between Minority Stressors and Parenting Desire

The results of the regression analysis are presented in Table 3.

Table 3. Test of the moderated mediation effect of prejudice events on parenting desire through sexual orientation concealment and internalized heterosexism.

Predictors	Sexual Orientation Concealment	Internalized Heterosexism	Parenting Desire
	B	B	B
Age	0.03 ***	−0.01 **	−0.04 **
Education (≤High school)	−0.04	−0.08	−0.14
Political orientation (left-wing)	0.30	0.06	0.30
Stable partner (yes)	−0.56 ***	0.15 **	−0.62 *
Prejudice events	0.27 *	0.02	0.05
Internalized heterosexism			0.24
Sexual orientation concealment			0.12
Prejudice events X gender			−1.23 *
Internalized heterosexism X gender			2.17 *
Sexual orientation concealment X gender			0.69 *
R^2	0.11	0.07	
−2LogLikelihood			292.07 $\Delta\chi^2$ (11) = 32.98, $p = 0.005$

Note: Unstandardized coefficients. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

The effects of prejudice events (Hypothesis 1), sexual orientation concealment, and internalized heterosexism (Hypothesis 2) on parenting desire were all moderated by gender (all interaction terms were significant). The pick-a-point approach revealed that prejudice events significantly reduced parenting desire in lesbians, but not in gay men (Figure 2), with Bs = −1.18 and 0.05, *ps* = 0.01 and 0.86, 95% CIs (−2.08, −0.28) and (−0.52, 0.63), respectively. Both sexual orientation concealment and internalized heterosexism were positively associated with parenting desire in lesbians, but not in gay men (Figures 3 and 4, respectively), with Bs = 0.81 and 0.12, *ps* = 0.01 and 0.36, 95% CIs (0.20, 1.42) and

(−0.14, 0.39) for sexual orientation concealment, respectively; and Bs = 2.41 and 0.24, *ps* = 0.003 and 0.55, 95% CIs (0.78, 4.04) and (−0.55, 1.04) for internalized heterosexism, respectively.

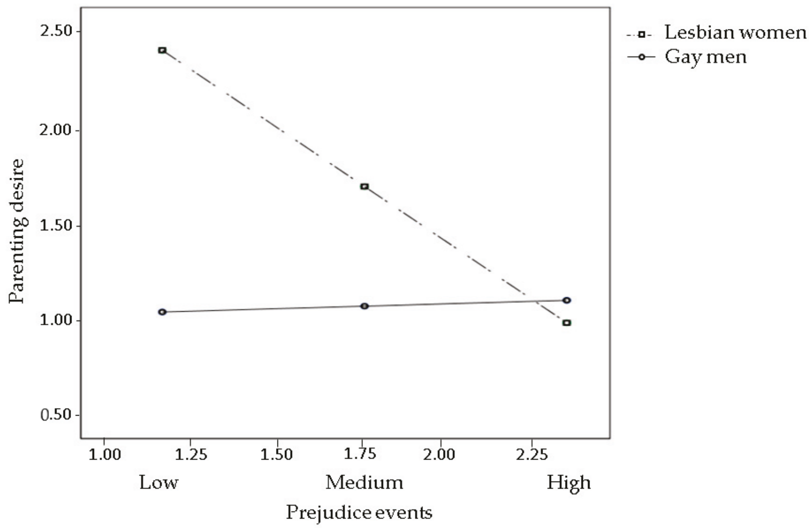


Figure 2. Interaction between prejudice events and gender on parenting desire. Estimated log odds of parenting desire are reported for combinations of prejudice events (low, medium, and high) and individual's gender (lesbian women vs. gay men).

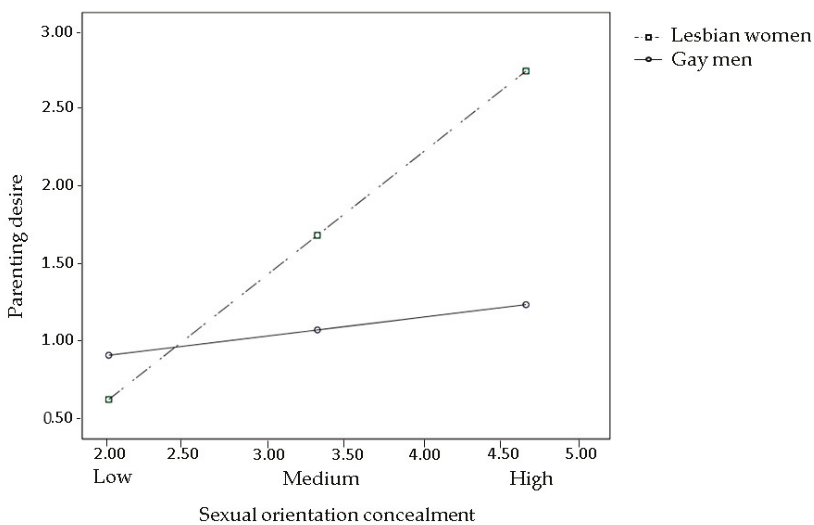


Figure 3. Interaction between sexual orientation concealment and gender on parenting desire. Estimated log odds of parenting desire are reported for combinations of sexual orientation concealment (low, medium, and high) and individual's gender (lesbian women vs. gay men).

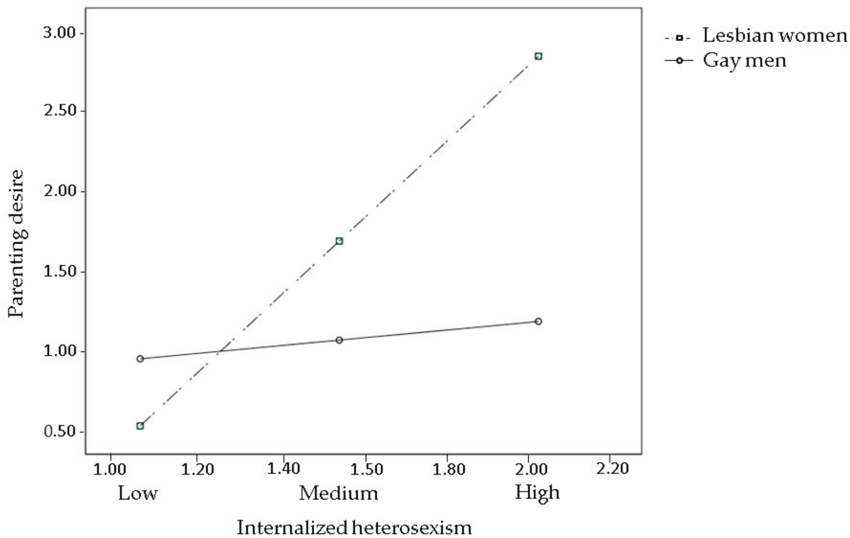


Figure 4. Interaction between internalized heterosexism and gender on parenting desire. Estimated log odds of parenting desire are reported for combinations of internalized heterosexism (low, medium, and high) and individual's gender (lesbian women vs. gay men).

With respect to moderated mediation effects (Hypothesis 3), we found only one indirect effect of prejudice events on parenting desire through sexual orientation concealment, which involved lesbians, but not gay men. The index of moderated mediation confirmed the significance of this effect (Table 4).

Table 4. Indirect effects of prejudice events on parenting desire through sexual orientation concealment and internalized heterosexism conditional on gender.

	Moderated Mediation Index	
Mediators	**B**	**95% C.I.**
Sexual orientation concealment	0.19	0.001, 0.63
Internalized heterosexism	0.05	−0.20, 0.40

3.3. Control Variables

Only age and having a stable partner were found to significantly predict outcomes in our study. Age was negatively associated with internalized heterosexism and parenting desire, whereas it was positively linked to sexual orientation concealment. Having a stable partner was associated with higher internalized heterosexism and lower sexual orientation concealment. Furthermore, having a stable partner was negatively associated with parenting desire.

4. Discussion

Informed by previous studies that used MST to explore parenting dimensions in LG individuals [9–11], as well as those that highlighted the heteronormative context for Italian LG individuals [20,38–42], the current study intended to apply part of the PMF as an extension of MST to parenting desire to a group of childless Italian LG individuals. We found partial support for this operation. Indeed, our results highlighted both a direct and an indirect effect (through sexual orientation concealment) of prejudice events on parenting desire, but only in lesbians. Of note, the direction of this effect changed depending on whether the association was direct or indirect. Specifically, we found that prejudice events were negatively associated with parenting desire, but that this effect became positive when it was mediated by sexual orientation concealment. Thus, our findings shed light on clinical practice, highlighting the need for structuring clinical interventions to ameliorate the negative effects of minority stressors on parenting desire.

In support of the first hypothesis of this study, we found that prejudice events were negatively associated with parenting desire only in lesbians and not in gay men. This means that, unlike gay men, prejudice events reduce the likelihood of lesbians desiring to become mothers. This finding might be read through the lens of MST, as well as through evidence that the Italian context is a highly heteronormative social context. To this end, previous studies have already demonstrated that discrimination episodes negatively impact parenting desire in lesbians, but not in gay men [11], and that this gender difference is due to the double stigma that many lesbians experience as both lesbians and women [54]. Furthermore, experiencing a prejudice event due to one's own sexual orientation and/or gender identity ends up emphasizing a status difference, as what is affected is a specific identity dimension. It is then likely that, in a heteronormative context, such as the Italian one [20,38–42], a lesbian woman experiencing a prejudice event will perceive this difference in status more than a gay man, because parenthood seems to still be a woman's prerogative [9,11]. Thus, a lesbian woman might feel herself to be less adequate than a gay man to become a parent, and this ends up decreasing the levels of parenting desire.

Regarding the second hypothesis, that sexual orientation concealment and internalized heterosexism negatively affect parenting desire in lesbians, but not in gay men, our findings were significant but, contrary to our expectations, they suggested that both subjective minority stressors were positively associated with parenting desire in lesbians and not in gay men, increasing the level of parenting desire in lesbians, rather than decreasing it. This unexpected finding might be read through the lens of MST and the related social pressure that some minority people could feel to conform to social expectations. Thus, it is probably that higher levels of internalized heterosexism, in addition to questioning one's own parenting abilities [13], might lead to a need to conform to social expectations, including becoming a mother in a context that still promotes motherhood as a key value for female identity [48]. Similarly, in line with the finding discussed above—that internalized heterosexism increases parenting desire in lesbians—the fact that sexual orientation concealment increased parenting desire only in lesbians, and not in gay men, might be read through the hypothesis that concealing one's own sexual orientation would lead lesbians to conform to societal expectations that tend to equate women to mothers. Notwithstanding these interpretative hypotheses, this finding should be qualitatively investigated to better understand how internalized stigma and sexual orientation concealment could lead to an increase in parenting desire.

Finally, regarding the third hypothesis (the moderated mediation model) that both sexual orientation concealment and internalized heterosexism would mediate the relationship between prejudice events and parenting desire, and that gender would moderate these indirect effects, our findings provided partial support. Indeed, only sexual orientation concealment, and not internalized heterosexism, acted as a significant mediator between prejudice events and parenting desire, and only in lesbians, increasing, rather than decreasing, their level of parenting desire. Thus, in the presence of sexual orientation concealment, prejudice events ended up increasing parenting desire in lesbians. As suggested by previous studies [24,25], unlike in gay men, the visibility of lesbian motherhood in previous decades, together with the social expectations that a woman will become a mother [27], might lead lesbians to perceive coming out as a mandatory step to becoming mothers. Thus, as suggested by our findings, experiencing prejudice events as a lesbian woman might lead to concealing one's own sexual orientation, which, in turn, increases the desire to become a mother, as becoming a mother would increase one's visibility as a normative woman, fitting the heteronormative equivalence of women as mothers.

4.1. Limitations and Suggestions for Future Research

This study had significant limitations that might affect the generalizability of the results to the general Italian LG population. First, the cross-sectional design of the study did not allow certain inferences to be made concerning the causal relationships between variables. Indeed, findings have to be interpreted as associations which do not necessarily prove causality between variables. Future longitudinal studies should assess the cause-and-effect relationships between minority stressors and parenting desire, analyzing potential changes over time.

A second limitation concerns the use of a single-item question to assess parenting desire. Although the same question had been used in a previous study [46], a more composite measure on parenting desire should be used in future research. Similarly, a qualitative investigation on parenting desire, and on its relationship with minority stressors, is recommended in order to achieve a better understanding of the psychological processes underlying these dimensions.

A third limitation was due to a sample selection bias. Indeed, participants were mostly recruited through social networks and Italian NGOs engaged in the defense of LG rights, thus not representing the general LG population. This may even explain why no right-wing individuals participated in the survey. Finally, in line with the previous limit, a last limitation was the absence of diverse ethnic groups of LG individuals, which prevented us from reading the findings from an intersectionality perspective. To this end, it is plausible to hypothesize that sociocultural differences between Caucasians and non-Caucasians living in Italy exist, both in terms of minority stress—that is usually more pervasive in multiple marginalized individuals [55]—and parenting process.

4.2. Implications for Clinical Practice

Clinicians might encounter LG individuals for many reasons, including LG individuals or couples who want to become parents or are questioning about a parenting plan. The results of this study might provide helpful suggestions for clinical practice with childless LG individuals, despite its limitations. Indeed, the PMF is a conceptual framework that allows an understanding of the negative effects of stress on individual psychological dimensions.

Our findings suggest that, unlike in gay men, parenting desire might be influenced by stigmatizing episodes in lesbians, and that this effect is partly explained by the action of sexual orientation concealment. Furthermore, although internalized heterosexism was not a significant mediator between prejudice events and parenting desire, it did increase the level of parenting desire in lesbians, probably because internalizing negative societal values about being lesbian leads to conforming to social expectations that tend to equate women to mothers. Thus, our results might shed light on some clinical implications above all for lesbians, both if they are single or in couple.

Our data should lead clinicians to pay particular attention to minority stress in clinical settings [56,57], and especially to the detrimental effects that stigma, and in particular proximal stressors (i.e., internalized heterosexism and sexual orientation concealment), might have on parenting desire. To this end, clinicians might assume the function of alleviating the emotional impact that internal stressors might have on individuals' desire, thoughts, or beliefs, helping clients to unhook their self-representation from the social dialectic that still tends to construct a negative image of the LG population, in particular as parents. This seems particular true for lesbians, who live a greater pressure than gay men to conform to social expectations related to parenthood. Indeed, although we usually interpret parenting desire or plan as a sign of comfort towards the self, due to an inherently generative project, our data suggest that the converse is also possible, as both internalized heterosexism and sexual orientation concealment increase, rather than decrease, parenting desire. This means that parenting desire might also be a sign of internalized heterosexism or sexual orientation concealment, rather than a manifestation of a positive identity. Thus, clinicians should deeply explore internal dynamics related to proximal stressors and be cautious in automatically interpreting parenting desire as a sign of comfort.

Furthermore, as both internalized heterosexism and sexual orientation concealment depends on the social stigma that burdens on LG people, clinicians should help clients to develop or increase awareness of stigma, allowing them to perceive oppression as a societal problem that afflict LG individuals as a class of people, rather than as something that concerns them as specific individuals.

Finally, clinicians should also consider group approach as an alternative to individual or couple approach, as the group activates mirroring processes that productively allow participants to reshape self-image [58] and increase the self-empowerment processes [59]. Indeed, LG individuals or couples experiencing high levels of internalized heterosexism might negotiate their negative self-image with peers, ameliorating the impact of stigma on health. In the same vein, the group might also be helpful for LG individuals or couples who tend to conceal their sexual orientation to family, friends, or colleagues. Therefore, sharing one's own identity to others living similar experiences in a protected and secure setting might allow an elaboration of feelings such as shame and self-hatred, thus facilitating identity or couple affirmation processes. In turn, as suggested by our findings, this clinical work could lighten the weight of social expectations related to parenthood, in particular in women.

5. Conclusions

This study provided support for applying the PMF to parenthood dimensions—in particular, parenting desire—in childless LG individuals. Indeed, notwithstanding the limitations of the research, and the nonconfirmation of all of the hypotheses, this study sheds light on the psychological processes that cause stress to increase or decrease the desire to become a parent. Future studies should thoroughly apply the PMF to parenthood dimensions, in order to explore the role of both general psychological processes (e.g., interpersonal problems or emotion dysregulation) and protective factors (e.g., resilience or community connectedness).

Author Contributions: A.L.A., V.B., and C.S. conceived and designed the experiments; C.E., R.V., D.B. and C.S. analyzed data; V.B., P.V. and R.V. contributed materials and analysis tools; D.B. and C.S. supervised the activity planning and execution; A.L.A., C.E. D.B., and C.S. wrote the paper.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mills, M.; Rindfuss, R.R.; McDonald, P.; Velde, E. Why do people postpone parenthood? Reasons and social policy incentives. *Hum. Reprod. Update* **2011**, *17*, 848–860. [CrossRef] [PubMed]
2. Riskind, R.G.; Patterson, C.G. Parenting intentions and desires among childless lesbian, gay, and heterosexual individuals. *J. Fam. Psychol.* **2010**, *24*, 78–81. [CrossRef] [PubMed]
3. Gato, J.; Santos, S.; Fontaine, A.M. To have or not to have children? That is the question. Factors influencing parental decisions among lesbians and gay men. *Sex Res. Soc. Policy* **2017**, *14*, 310–323. [CrossRef]
4. Gianino, M. Adaptation and transformation: The transition to adoptive parenthood for gay male couples. *J. GLBT Fam. Stud.* **2008**, *4*, 205–243. [CrossRef]
5. Goldberg, A.E.; Smith, J.Z. Stigma, social context, and mental health: Lesbian and gay couples across the transition to adoptive parenthood. *J. Couns. Psychol.* **2011**, *58*, 139–150. [CrossRef] [PubMed]
6. Meyer, I.H. Prejudice, social stress, and mental health in lesbian, gay and bisexual populations: Conceptual issues and research evidence. *Psychol. Bull.* **2003**, *129*, 674–697. [CrossRef] [PubMed]
7. Meyer, I.H. Prejudice and discrimination as social stressors. In *The Health of Sexual Minorities: Public Health Perspectives on Lesbian, Gay, Bisexual and Transgender Populations*; Meyer, I.H., Northridge, M.E., Eds.; Springer: New York, NY, USA, 2007; pp. 242–267.
8. Frost, D.M. Social stigma and its consequences for the socially stigmatized. *Soc. Personal. Psychol. Compass* **2011**, *5*, 824–839. [CrossRef]
9. Baiocco, R.; Argalia, M.; Laghi, F. The desire to marry and attitudes toward same-sex family legalization in a sample of Italian lesbians and gay men. *J. Fam. Issues* **2014**, *35*, 181–200. [CrossRef]

10. Bos, H.M.W.; van Balen, F.; van den Boom, D.C.; Sandfort, T.G.M. Minority stress, experience of parenthood and child adjustment in lesbian families. *J. Reprod. Infant. Psychol.* **2004**, *22*, 291–304. [CrossRef]
11. Scandurra, C.; Bacchini, D.; Esposito, C.; Bochicchio, V.; Valerio, P.; Amodeo, A.L. The influence of minority stress, gender, and legalization of civil unions on parenting desire and intention in lesbian women and gay men: Implications for social policy and clinical practice. *J. GLBT Fam. Stud.* **2018**, 1–25. [CrossRef]
12. Hatzenbuehler, M.L. How does sexual minority stigma "get under the skin"? A psychological mediation framework. *Psychol. Bull.* **2009**, *135*, 707–730. [CrossRef] [PubMed]
13. Mezey, N.J. How lesbians and gay men decide to become parents or remain childfree. In *LGBT-Parent Families. Innovations in Research and Implications for Practice*; Goldberg, A.E., Allen, K.R., Eds.; Springer: New York, NY, USA, 2013; pp. 59–70.
14. Woodward, L.G.; Fergusson, D.M.; Horwood, L.J. Gender differences in the transition to early parenthood. *Dev. Psychopathol.* **2006**, *18*, 275–294. [CrossRef] [PubMed]
15. Lachance-Grzela, M.; Bouchard, G. Why do women do the lion's share of housework? A decade of research. *Sex Roles* **2010**, *63*, 767–780. [CrossRef]
16. Farr, R.H.; Patterson, C.J. Coparenting among lesbian, gay, and heterosexual couples: Associations with adopted children's outcomes. *Child Dev.* **2013**, *84*, 1226–1240. [CrossRef] [PubMed]
17. Hatzenbuehler, M.L.; Pachankis, J.E. Stigma and minority stress as social determinants of health among lesbian, gay, bisexual, and transgender youth. *Pediatr. Clin. N. Am.* **2016**, *63*, 985–997. [CrossRef] [PubMed]
18. Institute of Medicine (IOM). *The Health of Lesbian, Gay, Bisexual, and Transgender People: Building a Foundation for Better Understand*; The National Academies Press: Washington, DC, USA, 2011.
19. Goldberg, A.E. *Lesbian and Gay Parents and Their Children: Research on the Family Life Cycle*; American Psychological Association: Washington, DC, USA, 2010; ISBN 978-1-4338-0536-3.
20. Pacilli, M.G.; Taurino, A.; Jost, J.T.; van der Toorn, J. System justification, right-wing conservatism, and internalized homophobia: Gay and lesbian attitudes toward same-sex parenting in Italy. *Sex Roles* **2011**, *65*, 580–595. [CrossRef]
21. Shilo, G.; Savaya, R. Effects of family and friend support on LGB youths mental health and sexual orientation milestones. *Fam. Relat.* **2011**, *60*, 318–330. [CrossRef]
22. Vaughan, M.D.; Waehler, C.A. Coming out growth: Conceptualizing and measuring stress-related growth associated with coming out to others as a sexual minority. *J. Adult. Dev.* **2010**, *17*, 94–109. [CrossRef]
23. Brown, S.; Smalling, S.; Groza, V.; Ryan, S. The experiences of gay men and lesbians in becoming and being adoptive parents. *Adopt. Q.* **2009**, *12*, 229–246. [CrossRef]
24. Brinamen, C.F.; Mitchell, V. Gay men becoming fathers: A model of identity expansion. *J. GLBT Fam. Stud.* **2008**, *4*, 521–541. [CrossRef]
25. Mallon, G.P. *Gay Men Choosing Parenthood*; Columbia University Press: New York, NY, USA, 2004; ISBN 9780231117975.
26. Bergman, K.; Rubio, R.J.; Green, R.J.; Padrón, E. Gay men who become fathers via surrogacy: The transition to parenthood. *J. GLBT Fam. Stud.* **2010**, *6*, 111–141. [CrossRef]
27. Hicks, S. Maternal men: Perverts and deviants? Making sense of gay men as foster careers and adopters. *J. GLBT Fam. Stud.* **2006**, *2*, 93–114. [CrossRef]
28. Liao, K.Y.H.; Kashubeck-West, S.; Weng, C.Y.; Deitz, C. Testing a mediation framework for the link between perceived discrimination and psychological distress among sexual minority individuals. *J. Couns. Psychol.* **2015**, *62*, 226–241. [CrossRef] [PubMed]
29. Schwartz, D.R.; Stratton, N.; Hart, T.A. Minority stress and mental and sexual health: Examining the psychological mediation framework among gay and bisexual men. *Psychol. Sex Orientat. Gend. Divers.* **2016**, *3*, 313–324. [CrossRef]
30. Breslow, A.S.; Brewster, M.E.; Velez, B.L.; Wong, S.; Geiger, L.; Soderstrom, B. Resilience and collective action: Exploring buffers against minority stress for transgender individuals. *Psychol. Sex Orientat. Gend. Divers.* **2015**, *2*, 253–265. [CrossRef]
31. Scandurra, C.; Bochicchio, V.; Amodeo, A.L.; Esposito, C.; Valerio, P.; Maldonato, N.M.; Bacchini, D.; Vitelli, R. Internalized transphobia, resilience, and mental health: Applying the Psychological Mediation Framework to Italian transgender individuals. *Int. J. Environ. Res. Public Health* **2018**, *15*, 508e. [CrossRef] [PubMed]
32. Timmins, L.; Rimes, K.A.; Rahman, Q. Minority stressors and psychological distress in transgender individuals. *Psychol. Sex Orientat. Gend. Divers.* **2017**, *4*, 328–340. [CrossRef]

33. Feinstein, B.A.; Goldfried, M.R.; Davila, J. The relationship between experiences of discrimination and mental health among lesbians and gay men: An examination of internalized homonegativity and rejection sensitivity as potential mechanisms. *J. Consult. Clin. Psychol.* **2012**, *80*, 917–927. [CrossRef] [PubMed]
34. Ryan, W.S.; Legate, N.; Weinstein, N. Coming out as lesbian, gay, or bisexual: The lasting impact of initial disclosure experiences. *Self Identity* **2015**, *14*, 549–569. [CrossRef]
35. Igartua, K.J.; Gill, K.; Montoro, R. Internalized homophobia: A factor in depression, anxiety, and suicide in the gay and lesbian population. *Can. J. Commun. Ment. Health* **2003**, *22*, 15–30. [CrossRef] [PubMed]
36. Legate, N.; Ryan, W.S.; Weinstein, N. Is coming out always a "good thing"? Exploring the relations of autonomy support, outness, and wellness for lesbian, gay, and bisexual individuals. *Soc. Psychol. Pers. Sci.* **2012**, *3*, 145–152. [CrossRef]
37. Lasio, D.; Congiargiu, N.; De Simone, S.; Serri, F. Gender fundamentalism and heteronormativity in the political discussion about lesbian and gay parenthood. *Sex Res. Soc. Policy* **2018**, 1–12. [CrossRef]
38. Amodeo, A.L.; Picariello, S.; Valerio, P.; Scandurra, C. Empowering transgender youths: Promoting resilience through a group training program. *J. Gay Lesbian Ment. Health* **2018**, *1*, 3–19. [CrossRef]
39. Lingiardi, V.; Baiocco, R.; Nardelli, N. Measure of Internalized Sexual Stigma for Lesbians and Gay Men: A new scale. *J. Homosex.* **2012**, *59*, 1191–1210. [CrossRef] [PubMed]
40. Scandurra, C.; Braucci, O.; Bochicchio, V.; Valerio, P.; Amodeo, A.L. "Soccer is a matter of real men?" Sexist and homophobic attitudes in three Italian soccer teams differentiated by sexual orientation and gender identity. *Int. J. Sport Exerc. Psychol.* **2017**, 1–17. [CrossRef]
41. Scandurra, C.; Mezza, F.; Bochicchio, V.; Valerio, P.; Amodeo, A.L. LGBT elders health from the minority stress perspective. Literature review & research recommendations. *Psicol. Della Salut.* **2017**, *2*, 70–96. [CrossRef]
42. Scandurra, C.; Picariello, S.; Valerio, P.; Amodeo, A.L. Sexism, homophobia and transphobia in a sample of Italian pre-service teachers: The role of socio-demographic features. *J. Educ. Teach.* **2017**, *43*, 245–261. [CrossRef]
43. Vitelli, R.; Scandurra, C.; Pacifico, R.; Selvino, M.S.; Picariello, S.; Amodeo, A.L.; Valerio, P.; Giami, A. Trans identities and medical practice in Italy: Self-positioning towards gender affirmation surgery. *Sexologies* **2017**, *26*, 43–51. [CrossRef]
44. Vitelli, R.; Scandurra, C.; Pacifico, R.; Selvino, M.S.; Picariello, S.; Amodeo, A.L.; Valerio, P.; Giami, A. Identifications de genre trans et pratiques médicales en Italie: Positionnement personnel envers les forms de chirurgie de confirmation de genre. *Sexologies* **2017**, *26*, 199–204. [CrossRef]
45. Lingiardi, V.; Carone, N.; Morelli, M.; Baiocco, R. "It's a bit too much fathering this seed": The meaning-making of the sperm donor in Italian lesbian mother families. *Reprod. Biomed. Online* **2016**, *33*, 412–424. [CrossRef] [PubMed]
46. Roy, R.N.; Schumm, W.R.; Britt, S.L. *Transition to Parenthood*; Springer: New York, NY, USA, 2014; ISBN 978-1-4614-7768-6.
47. Scandurra, C.; Amodeo, A.L.; Valerio, P.; Bochicchio, V.; Frost, D.M. Minority stress, resilience, and mental health: A study of Italian transgender people. *J. Soc. Issues* **2017**, *73*, 563–585. [CrossRef]
48. Baiocco, R.; Laghi, F. Sexual orientation and the desires and intentions to become parents. *J. Fam. Stud.* **2013**, *19*, 90–98. [CrossRef]
49. Bartos, S.E.; Baban, A. Predictors of emotional distress in Romanian gay men. *Psychol. Health* **2010**, *25*, 194–195. [CrossRef]
50. Montano, A.; Andriola, E. *Parlare di Omosessualità a Scuola. Riflessioni e Attività per la Scuola Secondaria*; Talking about Homosexuality in School; Reflections and Activities for the Secondary School; Edizioni Erickson: Trento, Italy, 2011; ISBN 9788861378278.
51. Pistella, J.; Tanzilli, A.; Ioverno, S.; Lingiardi, V.; Baiocco, R. Sexism and attitudes toward same-sex parenting in a sample of heterosexuals and sexual minorities: The mediation effect of sexual stigma. *Sex Res. Soc. Policy* **2018**, *15*, 139–150. [CrossRef]
52. Mohr, J.J.; Fassinger, R.E. Measuring dimensions of lesbian and gay male experience. *Meas. Eval. Couns. Dev.* **2000**, *33*, 66–90. [CrossRef]
53. Hayes, A.F. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*, 2nd ed.; Guilford Press: New York, NY, USA, 2018; ISBN 9781462534654.

54. Connolly, C.M. A feminist perspective of resilience in lesbian couples. *J. Fem. Fam. Ther.* **2006**, *18*, 137–162. [CrossRef]
55. Balsam, K.F.; Molina, Y.; Beadnell, B.; Simoni, J.; Walters, K. Measuring multiple minority stress: The LGBT People of Color Microaggressions Scale. *Cultur. Divers. Ethnic Minor. Psychol.* **2011**, *17*, 163–174. [CrossRef] [PubMed]
56. Scandurra, C.; Amodeo, A.L.; Bochicchio, V.; Valerio, P.; Frost, D.M. Psychometric characteristics of the Transgender Identity Survey in an Italian sample: A measure to assess positive and negative feelings towards transgender identity. *Int. J. Transgend.* **2017**, *18*, 53–65. [CrossRef]
57. Amodeo, A.L.; Vitelli, R.; Scandurra, C.; Picariello, S.; Valerio, P. Adult attachment and transgender identity in the Italian context: Clinical implications and suggestions for further research. *Int. J. Transgend.* **2015**, *16*, 49–61. [CrossRef] [PubMed]
58. Amodeo, A.L.; Picariello, S.; Valerio, P.; Bochicchio, V.; Scandurra, C. Group psychodynamic counselling with final-year undergraduates in clinical psychology: A clinical methodology to reinforce academic identity and psychological wellbeing. *Psychodyn. Pract.* **2017**, *23*, 161–180. [CrossRef]
59. Scandurra, C.; Picariello, S.; Scafaro, D.; Bochicchio, V.; Valerio, P.; Amodeo, A.L. Group psychodynamic counselling as a clinical training device to enhance metacognitive skills and agency in future clinical psychologists. *Eur. J. Psychol.* **2018**, *14*, 444–463. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

Depression and Its Association with Health-Related Quality of Life in Postmenopausal Women in Korea

Hyejin Park [1] and Kisok Kim [2,*]

[1] Department of International Healthcare Administration, Daegu Catholic University, Gyeongsan 38430, Korea; hjpark@cu.ac.kr
[2] College of Pharmacy, Keimyung University, Daegu Metropolitan 42601, Korea
* Correspondence: kimkisok@kmu.ac.kr; Tel.: +82-53-580-5932

Received: 28 September 2018; Accepted: 21 October 2018; Published: 23 October 2018

Abstract: Menopause is associated with depressive symptoms that can significantly affect a woman's quality of life. The objective of this study was to evaluate the association between depression and health-related quality of life (HRQoL) in postmenopausal women. In this cross-sectional descriptive study, participants (n = 3860) were selected from the 2013–2015 Korea National Health and Nutrition Examination Survey (KNHANES). The sociodemographic characteristics, medical history of depression, and EQ-5D scores of the participants were obtained from the KNHANES dataset. Age, educational level, and income were associated with HRQoL in these participants. Moreover, depression exerted a considerable influence on HRQoL in postmenopausal women. The adjusted odds ratios in participants with depression for the EQ-5D dimensions were as follows: 5.52 (95% CI = 4.04–7.55, $p < 0.001$) for anxiety/depression, 3.86 (95% CI = 2.78–5.36, $p < 0.001$) for usual activities, and 2.52 (95% CI = 1.68–3.78, $p < 0.001$) for self-care. Our findings suggest a strong association between depression and HRQoL. Hence, preventing the onset or exacerbation of depression may significantly improve quality of life in postmenopausal women.

Keywords: demographic factor; EQ-5D; health-related quality of life; depression

1. Introduction

The prevalence of depression is increasing worldwide [1]. Indeed, depression is one of the most common mental disorders among elderly people (prevalence, 1–16%) [2]. Along with somatic illness, functional disability, cognitive impairment, and lack of social contacts, female gender is also associated with depressive disorders in elderly populations [3]. Epidemiological and clinical studies have shown that the prevalence of depression is higher in women than in men [4] by up to twofold [5]. This difference is thought to be related to changes in endocrines that control the reproductive system [6]. Postmenopausal women experience considerable biological and psychological changes, including a decreased level of estrogen, which may be related to depression [7,8]. Estrogen interacts with its receptors in the limbic area of the brain, which is important for the regulation of emotions, cognition, and behavior [9,10].

Depression reduces an individual's mental and physical health [11,12] and is associated with a diminished quality of life (QoL) [13,14]. Depression is also associated with several functional disturbances and significant reductions in several aspects of QoL, including social functioning [15]. In addition, depressive symptoms have various effects on physical and mental health-related quality of life (HRQoL) of elderly individuals [16]. In older adults, lowered QoL has been reported to be greatly dependent on reduced physical function. Inability to perform activities of daily living or instrumental activities of daily living has been known to be associated with decreased QoL [17,18]. Besides, depression has been found to be significantly correlated with functional disability [19]. It is known that the effect of depression on functional disability may partially be due to deteriorated physical activity and social interactions of depressed elderly individuals [20]. Depressive symptoms and disorders are frequent causes of emotional and physical suffering and are associated with elevated risks of disability in diverse areas of functioning [21].

HRQoL is influenced by sociodemographic factors, such as gender, age, educational level, and income [22]. In addition, low socioeconomic status increases the risk of depression [23]. Therefore, it is important to analyze the relationship between depression and HRQoL in postmenopausal women according to sociodemographic factors.

Few studies have addressed depression in relation to HRQoL in postmenopausal Korean women using a population-based sample. Therefore, in this report, we investigated the association between depression and HRQoL in postmenopausal women using data from the Korea National Health and Nutrition Examination Survey (KNHANES). We also evaluated the prevalence of depression according to participants' sociodemographic characteristics and the relationship between depression-related morbidity and the five dimensions of EuroQoL. We postulate that the prevalence of depression and HRQoL may be significantly influenced by sociodemographic factors. Additionally, the prevalence of depression could be associated with HRQoL in postmenopausal women in Korea.

2. Methods

2.1. Study Population

This study used data from the KNHANES 2013–2015, which included a health and nutrition survey and a medical examination. The KNHANES sample was chosen using a stratified multistage cluster sampling method with proportional allocation based on the National Census Registry. Face-to-face interviews using a structured questionnaire were conducted by trained interviewers. From this sample, postmenopausal women were selected for inclusion in the present study. Of the postmenopausal women, 3860 provided data with no missing variables. The study protocol was approved by the Korean Ministry of Health and Welfare (# 2013-07CON-03-4C, 2013-12EXP-03-5C) and was conducted in accordance with the Ethical Principles for Medical Research Involving Human Subjects, as defined by the Helsinki Declaration. All participants in this study provided written informed consent.

2.2. Source of Data and Variable Definitions

Participants from selected census blocks provided information on their age, educational level, income, and medical history of depression. Height and weight were measured with the participants dressed in light clothing with no shoes to determine body mass index (BMI). BMI was calculated as weight (kg) divided by the square of height (m^2). The participants were classified as underweight (BMI $<$ 18.5), normal (18.5 $\leq$ BMI $<$ 22.9), overweight (23.0 $\leq$ BMI $<$ 24.9), or obese (BMI $\geq$ 25.0) according to the WHO definitions for Asian populations. In this study, HRQoL was treated as the dependent variable that was influenced by depression as an independent variable. The depression criterion was a self-reported history of physician-diagnosed depression (DSM-IV). In this study, we present the prevalence of current physician-diagnosed depression. As confounding factors, educational level and income were used as indicators of socioeconomic status. Educational level was classified as less than a middle school graduate, middle school graduate, or high school or higher. Income was calculated by dividing household income by the square root of the number of members in the household according to the Organization for Economic Co-operation and Development (OECD) method. Income was categorized into quartiles based on the income of the participant's age group. HRQoL was assessed using the EQ-5D questionnaire. The EQ-5D is a self-reported, descriptive health status instrument with five health dimensions: mobility, self-care, pain/discomfort, usual activities, and anxiety/depression. Each dimension has three levels, namely "no problems", "some problems", and "severe or extreme problems" [24]. The five dimensions of the ED-5D are then converted into EQ-5D index scores using Korean specific preference weight [25]. Average scores of the EQ-5D index ranged from −0.17 to 1, where 1 indicates no problem in any of the five dimensions, zero indicates death, and negative values indicate health statuses worse than death. The Korean versions of the EQ-5D tools have been validated in a previous study [26]. The kappa value of EQ-5D in dimensions between test and retest was 0.32–0.64, and the intraclass correlation coefficient of EQ-5D was 0.61 [26]. For the purpose of this study, the levels were used as an overall measure of the perceived HRQoL.

2.3. Statistical Analyses

Differences in categorical variables between groups were evaluated by the Mantel–Haenszel chi-squared test. The presence of a linear trend was evaluated by defining a linear contrast in the linear regression models. Logistic regression models were used to estimate the odds ratio (OR) and 95% CIs for abnormal (disability) versus normal (no problem) in the EuroQol categories among participants who reported having depression compared with those who did not. Statistical analyses were conducted using SAS version 9.4 (SAS Institute, Cary, NC, USA).

3. Results

Table 1 shows the prevalence of depression according to the demographic characteristics of the participants. The prevalence of depression was in the range of 2.7–7.2%. Educational level and income level were significantly associated with depression. The prevalence of depression decreased significantly with increasing income and educational level ($p < 0.01$).

Table 1. Prevalence of depression in postmenopausal women according to demographic characteristics.

Variable	*n*	Depression (%)	*p*-Value [a]
Age (years)			
<60	1303	46 (3.5)	0.092
60–69	1255	68 (5.4)	
≥70	1302	64 (4.9)	
BMI (kg/m^2)			
<18.5	92	4 (4.4)	0.133
18.5–22.9	1391	57 (4.1)	
23.0–24.9	990	43 (4.3)	
≥25.0	1387	74 (5.3)	
Education			
≤Elementary school	2126	112 (5.3)	0.007
Middle school	601	31 (5.2)	
≥High school	1133	35 (3.1)	
Income			
Quartile 1 (lowest)	919	66 (7.2)	<0.001
Quartile 2	980	45 (4.6)	
Quartile 3	976	40 (4.1)	
Quartile 4 (highest)	985	27 (2.7)	

Note: [a] *p* is determined by Mantel–Haenszel chi-squared test.

The mean EQ-5D scores of the participants are listed in Table 2. The EQ-5D score decreased with increasing age (p for trend < 0.001). In addition, the EQ-5D scores increased with increasing BMI (0.90, 0.89, 0.86, and 0.83 for obese, overweight, normal weight, and underweight, respectively). The EQ-5D scores also increased with increasing educational level (p for trend < 0.001) and income level ($p < 0.001$).

Table 3 shows the participants' age, BMI, and EQ-5D score according to depression category. The mean age, BMI, and EQ-5D score of participants with depression were 64.7, 24.2, and 0.88, respectively. The mean age and BMI were not significantly different between participants with and without depression, whereas the EQ-5D score was significantly lower in the participants with depression (0.77 and 0.89, respectively; $p < 0.001$).

Table 2. Mean EQ-5D scores of postmenopausal women according to demographic characteristics.

Variable	n	EQ-5D	p for Trend [a]
Age (years)			
<60	1303	0.94 ± 0.11	<0.001
60–69	1255	0.89 ± 0.15	
≥70	1302	0.81 ± 0.20	
BMI (kg/m^2)			
<18.5	92	0.83 ± 0.22	0.119
18.5–22.9	1391	0.90 ± 0.16	
23.0–24.9	990	0.89 ± 0.15	
≥25.0	1387	0.86 ± 0.17	
Education			
≤Elementary school	2126	0.84 ± 0.19	<0.001
Middle school	601	0.92 ± 0.12	
≥High school	1133	0.94 ± 0.10	
Income			
Quartile 1 (lowest)	919	0.85 ± 0.19	<0.001
Quartile 2	980	0.88 ± 0.17	
Quartile 3	976	0.89 ± 0.16	
Quartile 4 (highest)	985	0.91 ± 0.15	

Note: [a] p is determined by linear trend test.

Table 3. Descriptive characteristics of the participants by depression category.

Variable	All	Non-Depression	Depression
n	3860	3682	178
Age (years)	64.7 ± 9.1	64.7 ± 9.1	65.2 ± 8.4
BMI (kg/m^2)	24.2 ± 3.3	24.2 ± 3.3	24.3 ± 3.1
EQ-5D	0.88 ± 0.17	0.89 ± 0.16	0.77 ± 0.23 *

Note: Data are mean ± standard deviation. * $p < 0.0001$ by t-test.

Table 4 shows the ORs for disability. The adjusted ORs for disability for each EuroQol category were significantly related to the prevalence of depression after adjusting for age (model 1) or for all other potential covariates (model 2) ($p < 0.001$). Compared to participants without depression, the adjusted ORs for mobility, self-care, usual activity, pain/discomfort, and anxiety/depression were 2.49 (95% CI, 1.79–3.47), 2.52 (95% CI, 1.68–3.78), 3.86 (95% CI, 2.78–5.36), 2.14 (95% CI, 1.57–2.92), and 5.52 (95% CI, 4.04–7.55), respectively, in those with depression ($p < 0.001$) (model 2).

Table 4. Adjusted odds ratios (OR, 95% CIs) for disability for the EuroQol categories in postmenopausal women.

Dimension	Non-Depression (n = 3682)	Depression (n = 178)	p-Value
EuroQoL-mobility			
Model 1	1.00 (reference)	2.69 (1.94–3.72)	<0.001
Model 2	1.00 (reference)	2.49 (1.79–3.47)	<0.001
EuroQoL-self care			
Model 1	1.00 (reference)	2.66 (1.78–3.97)	<0.001
Model 2	1.00 (reference)	2.52 (1.68–3.78)	<0.001
EuroQoL-usual activities			
Model 1	1.00 (reference)	4.01 (2.91–5.54)	<0.001
Model 2	1.00 (reference)	3.86 (2.78–5.36)	<0.001

Table 4. *Cont.*

Dimension	Non-Depression (n = 3682)	Depression (n = 178)	p-Value
EuroQoL-pain/discomfort			
Model 1	1.00 (reference)	2.24 (1.64–3.05)	<0.001
Model 2	1.00 (reference)	2.14 (1.57–2.92)	<0.001
EuroQoL-anxiety/depression			
Model 1	1.00 (reference)	5.86 (4.30–7.99)	<0.001
Model 2	1.00 (reference)	5.52 (4.04–7.55)	<0.001

Note: Model 1 was adjusted for age; model 2 was adjusted for age, body mass index (BMI), educational level, and income.

4. Discussion

In this large population-based study, we investigated the association between depression and HRQoL using the EQ-5D. Our results indicate a significant association between depression and educational level and income level in postmenopausal Korean women. The relationship between depression and sociodemographic characteristics, such as gender, age, marital status, income level, and educational level has been investigated worldwide in various populations [27]. In many of these studies, the educational level and income of the participants were significantly related to the prevalence of depressive symptoms. The association between educational level and depression may be caused by reduced access to information about risk factors [28,29]. Alternatively, a poor education could be associated with low income, which could affect physical and mental health [30].

Our results showed that age, educational level, and income are associated with HRQoL. This is consistent with two previous studies of adult populations [31,32]. The sociodemographic determinants of HRQoL in the general population have been documented in Sweden, the Netherlands, Norway, the United Kingdom, the United States, and Japan [33–38]. In these studies, sex, age, educational attainment, income, and chronic disease had significant impacts on participants' QoL.

In this study, depression was significantly related to HRQoL in postmenopausal women. Previous studies have found robust association between depression and low HRQoL in specific population subgroups [39–42]. In particular, several studies worldwide have explored the relationship between depressive symptoms or a depressive disorder and HRQoL in older adults. A recent study on Portuguese elderly population showed that older adults with depression symptoms had a higher probability of reporting lower levels of HRQoL after adjustment for sex, age, region, and number of noncommunicable diseases [43]. In a similar context, a review paper by Sivertsen et al. indicated a clear and consistent association between depression and lower HRQoL in older adults in clinical and community settings [44]. Based on 19 cross-sectional studies, this review found a significant association between severity of depression and lower HRQoL in older adults, regardless of assessment instruments used for HRQoL. Furthermore, in 10 longitudinal studies, a depressive disorder and a higher depressive symptom score were consistently associated with lower HRQoL, and this association was found to be stable over time. Participants with a depressive disorder at baseline had lower HRQoL at follow-up than participants without depression, and the severity of depressive symptoms at baseline had a significant effect on any improvement in HRQoL at follow-up.

Thus far, there have only been a few studies that have investigated this relationship in a population-based sample of postmenopausal women. Depression at any time of life, including postmenopausal period, is known to negatively impact QoL measures as well as somatic complaints [45]. Community-based longitudinal studies have reported that the risk of depression is significant increased during the menopause transition compared with premenopausal and that the prevalence of depression in some premenopausal women could be an important source of variability in measures of QoL at this stage of a woman's life [45–47]. In our study, the ORs for disability in the five dimensions were significantly increased among postmenopausal women with depression. Anxiety/depression and usual activities exerted the greatest influence on HRQoL among postmenopausal women with depression. These results suggest that in addition to anxiety/depression, limitations on one's usual activities are also important targets for improving HRQoL in postmenopausal women.

This study has several limitations. First, the results only indicate associations and cannot be used to establish causal relationships due to the cross-sectional design of this study. Second, self-reports of current depression state and sociodemographic variables may lead to misclassification and recall bias. Third, the use of EQ-5D to assess QoL may also have limitations related to the reliability and objectivity of the findings. However, this study also has several strengths. To our knowledge, this is the first population-based study to assess the association between depression and HRQoL among postmenopausal Korean women using nationally representative data. This study analyzed a large population based on systematic sampling, which enhances the generalizability of the findings. Investigation of factors related to HRQoL in specific subgroups is an important concern for health policymakers and for the development of appropriate interventions to improve individuals' QoL. Our results suggest that health policymakers should focus on the effects of sociodemographic factors and depression on the HRQoL of postmenopausal women in Korea.

5. Conclusions

In this population-based study, HRQoL among postmenopausal women was associated with educational level and income level. In addition, postmenopausal women with depression had a significantly lower HRQoL. Depression significantly increased the ORs for disability in all five dimensions of the EQ-5D. Our results indicate a significant association between depression and HRQoL in postmenopausal Korean women. Therefore, prevention and management of depression are important for improving the QoL of postmenopausal women.

Author Contributions: H.P. contributed to data analysis and interpretation, statistical analysis, and drafting the manuscript. K.K. contributed to design of the study, drafting and revision of the manuscript, and supervision of the study. All authors have read and approved the final manuscript.

Funding: This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (Nos. NRF-2018R1D1A1B07050009, NRF-2015R1D1A1A01056617).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hidaka, B.H. Depression as a disease of modernity: Explanations for increasing prevalence. *J. Affect. Disord.* **2012**, *140*, 205–214. [CrossRef] [PubMed]
2. Domènech-Abella, J.; Mundó, J.; Leonardi, M.; Chatterji, S.; Tobiasz-Adamczyk, B.; Koskinen, S.; Ayuso-Mateos, J.L.; Haro, J.M. The association between socioeconomic status and depression among older adults in Finland, Poland and Spain: A comparative cross-sectional study of distinct measures and pathways. *J. Affect. Disord.* **2018**, *241*, 311–318. [CrossRef] [PubMed]
3. Djernes, J.K. Prevalence and predictors of depression in populations of elderly: A review. *Acta Psychiatr. Scand.* **2006**, *113*, 372–387. [CrossRef] [PubMed]
4. Leach, L.S.; Christensen, H.; Mackinnon, A.J.; Windsor, T.D.; Butterworth, P. Gender differences in depression and anxiety across the adult lifespan: The role of psychosocial mediators. *Soc. Psychiatry Psychiatr. Epidemiol.* **2008**, *43*, 983–998. [CrossRef] [PubMed]
5. Kessler, R.C.; McGonagle, K.A.; Zhao, S.; Nelson, C.B.; Hughes, M.; Eshleman, S.; Wittchen, H.U.; Kendler, K.S. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: Results from the National Comorbidity Survey. *Arch. Gen. Psychiatry* **1994**, *51*, 8–19. [CrossRef] [PubMed]
6. Noble, R.E. Depression in women. *Metabolism* **2005**, *54*, 49–52. [CrossRef] [PubMed]
7. Steiner, M.; Dunn, E.; Born, L. Hormones and mood: From menarche to menopause and beyond. *J. Affect. Disord.* **2003**, *74*, 67–83. [CrossRef]
8. Freeman, E.W.; Sammel, M.D.; Liu, L.; Gracia, C.R.; Nelson, D.B.; Hollander, L. Hormones and menopausal status as predictors of depression in women in transition to menopause. *Arch. Gen. Psychiatry* **2004**, *61*, 62–70. [CrossRef] [PubMed]

9. Westberg, L.; Eriksson, E. Sex steroid-related candidate genes in psychiatric disorders. *J. Psychiatry Neurosci.* **2008**, *33*, 319–330. [PubMed]
10. Ostlund, H.; Keller, E.; Hurd, Y.L. Estrogen receptor gene expression in relation to neuropsychiatric disorders. *Ann. N. Y. Acad. Sci.* **2003**, *1007*, 54–63. [CrossRef] [PubMed]
11. Prince, M.; Patel, V.; Saxena, S.; Maj, M.; Maselko, J.; Phillips, M.R.; Rahman, A. No health without mental health. *Lancet* **2007**, *370*, 859–877. [CrossRef]
12. Luppa, M.; Sikorski, C.; Luck, T.; Ehreke, L.; Konnopka, A.; Wiese, B.; Weyerer, S.; König, H.H.; Riedel-Heller, S.G. Age- and gender-specific prevalence of depression in latest-life–systematic review and meta-analysis. *J. Affect. Disord.* **2012**, *136*, 212–221. [CrossRef] [PubMed]
13. Kessler, R.C.; Bromet, E.J. The epidemiology of depression across cultures. *Annu. Rev. Public Health* **2013**, *34*, 119–138. [CrossRef] [PubMed]
14. Ferrari, A.J.; Charlson, F.J.; Norman, R.E.; Patten, S.B.; Freedman, G.; Murray, C.J.; Vos, T.; Whiteford, H.A. Burden of depressive disorders by country, sex, age, and year: Findings from the global burden of disease study 2010. *PLoS Med.* **2013**, *10*, e1001547. [CrossRef] [PubMed]
15. Ravindran, A.V.; Matheson, K.; Griffiths, J.; Merali, Z.; Anisman, H. Stress, coping, uplifts, and quality of life in subtypes of depression: A conceptual frame and emerging data. *J. Affect. Disord.* **2002**, *71*, 121–130. [CrossRef]
16. Ritti-Dias, R.M.; Cucato, G.G.; de Matos, L.D.N.J.; Cendoroglo, M.S.; Nasri, F.; Wolosker, N.; Costa, M.L.M.; Gazelato de Mello Franco, F. Depression and cancer were independently associated with quality of life in Brazilian older people. *Australas. J. Ageing.* (in press). [CrossRef] [PubMed]
17. Akosile, C.O.; Mgbeojedo, U.G.; Maruf, F.A.; Okoye, E.C.; Umeonwuka, I.C.; Ogunniyi, A. Depression, functional disability and quality of life among Nigerian older adults: Prevalences and relationships. *Arch. Gerontol. Geriatr.* **2018**, *74*, 39–43. [CrossRef] [PubMed]
18. Akosile, C.O.; Anukam, G.O.; Johnson, O.E.; Fabunmi, A.A.; Okoye, E.C.; Iheukwumere, N.; Akinwola, M.O. Fear of falling and quality of life of apparently-healthy elderly individuals from a Nigerian population. *J. Cross Cult. Gerontol.* **2014**, *29*, 201–209. [CrossRef] [PubMed]
19. Geerlings, S.W.; Beekman, A.T.F.; Deeg, D.J.H.; Twisk, J.W.R.; Van-Tilburg, W. The longitudinal effect of depression on functional limitations and disability in older adults: An eight-wave prospective community-based study. *Psychol. Med.* **2001**, *31*, 1361–1371. [CrossRef] [PubMed]
20. Pennix, B.W.J.H.; Geerlings, S.W.; Deg, D.J.H.; van-Eijk, J.T.M.; van-Tilburg, W.; Beekman, A.T.F. Minor and major depression and the risk of death in older persons. *Arch. Gen. Psychiatry* **1999**, *56*, 889–895. [CrossRef]
21. Gureje, O.; Kola, L.; Afolabi, E.; Olley, B.O. Determinants of quality of life of elderly Nigerians: Results from the Ibadan Study of Ageing. *Afr. J. Med. Med. Sci.* **2008**, *37*, 239–247. [PubMed]
22. Jiang, Y.; Hesser, J.E. Associations between health-related quality of life and demographics and health risks. Results from Rhode Island's 2002 behavioral risk factor survey. *Health Qual. Life Outcomes* **2006**, *4*, 14. [CrossRef] [PubMed]
23. Lorant, V.; Deliège, D.; Eaton, W.; Robert, A.; Philippot, P.; Ansseau, M. Socioeconomic inequalities in depression: A meta-analysis. *Am. J. Epidemiol.* **2003**, *157*, 98–112. [CrossRef] [PubMed]
24. Dritsaki, M.; Petrou, S.; Williams, M.; Lamb, S.E. An empirical evaluation of the SF-12, SF-6D, EQ-5D and Michigan Hand Outcome Questionnaire in patients with rheumatoid arthritis of the hand. *Health Qual. Life Outcomes* **2017**, *15*, 20. [CrossRef] [PubMed]
25. Nam, H.S.; Kim, K.Y.; Kwon, S.S.; Koh, K.W.; Poul, K. *EQ-5D Korean Valuation Study Using Time Trade of Method*; Centers for Disease Control and Prevention: Seoul, Korea, 2007.
26. Lee, S.I. *Validity and Reliability Evaluation for EQ-5D in Korea*; Centers for Disease Control and Prevention: Seoul, Korea, 2011.
27. Bromet, E.; Andrade, L.H.; Hwang, I.; Sampson, N.A.; Alonso, J.; de Girolamo, G.; de Graaf, R.; Demyttenaere, K.; Hu, C.; Iwata, N.; et al. Cross-national epidemiology of DSM-IV major depressive episode. *BMC Med.* **2011**, *9*, 90. [CrossRef] [PubMed]
28. Turgunova, L.; Laryushina, Y.; Turmukhambetova, A.; Koichubekov, B.; Sorokina, M.; Korshukov, I. The incidence of depression among the population of central Kazakhstan and its relationship with sociodemographic characteristics. *Behav. Neurol.* **2017**, *2017*, 2584187. [CrossRef] [PubMed]
29. Çakıcı, M.; Gökçe, Ö.; Babayiğit, A.; Çakıcı, E.; Eş, A. Depression: Point-prevalence and risk factors in a North Cyprus household adult cross-sectional study. *BMC Psychiatry* **2017**, *17*, 387. [CrossRef] [PubMed]

30. Lahelma, E.; Martikainen, P.; Laaksonen, M.; Aittomäki, A. Pathways between socioeconomic determinants of health. *J. Epidemiol. Community Health* **2004**, *58*, 327–332. [CrossRef] [PubMed]
31. Singh, K.; Kondal, D.; Shivashankar, R.; Ali, M.K.; Pradeepa, R.; Ajay, V.S.; Mohan, V.; Kadir, M.M.; Sullivan, M.D.; Tandon, N.; et al. Health-related quality of life variations by sociodemographic factors and chronic conditions in three metropolitan cities of South Asia: The CARRS study. *BMJ Open* **2017**, *7*, e018424. [CrossRef] [PubMed]
32. Rezaei, S.; Hajizadeh, M.; Kazemi, A.; Khosravipour, M.; Khosravi, F.; Rezaeian, S. Determinants of health-related quality of life in Iranian adults: Evidence from a cross-sectional study. *Epidemiol. Health* **2017**, *39*, e2017038. [CrossRef] [PubMed]
33. Burström, K.; Johannesson, M.; Diderichsen, F. Swedish population health-related quality of life results using the EQ-5D. *Qual. Life Res.* **2001**, *10*, 621–635. [CrossRef] [PubMed]
34. Terwindt, G.M.; Ferrari, M.D.; Tijhuis, M.; Groenen, S.M.; Picavet, H.S.; Launer, L.J. The impact of migraine on quality of life in the general population: The GEM study. *Neurology* **2000**, *55*, 624–629. [CrossRef] [PubMed]
35. Wahl, A.K.; Rustøen, T.; Rokne, B.; Lerdal, A.; Knudsen, Ø.; Miaskowski, C.; Moum, T. The complexity of the relationship between chronic pain and quality of life: A study of the general Norwegian population. *Qual. Life Res.* **2009**, *18*, 971–980. [CrossRef] [PubMed]
36. Kind, P.; Dolan, P.; Gudex, C.; Williams, A. Variations in population health status: Results from a United Kingdom national questionnaire survey. *BMJ* **1998**, *316*, 736–741. [CrossRef] [PubMed]
37. Johnson, J.A.; Coons, S.J.; Ergo, A.; Szava-Kovats, G. Valuation of EuroQOL (EQ-5D) health states in an adult US sample. *Pharmacoeconomics* **1998**, *13*, 421–433. [CrossRef] [PubMed]
38. Hisashige, A.; Mikasa, H.; Katayama, T. Description and valuation of health-related quality of life among the general public in Japan by the EuroQol. *J. Med. Invest.* **1998**, *45*, 123–129. [PubMed]
39. Ghimire, S.; Baral, B.K.; Pokhrel, B.R.; Pokhrel, A.; Acharya, A.; Amatya, D.; Amatya, P.; Mishra, S.R. Depression, malnutrition, and health-related quality of life among Nepali older patients. *BMC Geriatr.* **2018**, *18*, 191. [CrossRef] [PubMed]
40. Thau, A.J.; Rohn, M.C.H.; Biron, M.E.; Rahmatnejad, K.; Mayro, E.L.; Gentile, P.M.; Waisbourd, M.; Zhan, T.; Hark, L.A. Depression and quality of life in a community-based glaucoma-screening project. *Can. J. Ophthalmol.* **2018**, *53*, 354–360. [CrossRef] [PubMed]
41. Akioyamen, L.E.; Genest, J.; Shan, S.D.; Inibhunu, H.; Chu, A.; Tu, J.V. Anxiety, depression, and health-related quality of life in heterozygous familial hypercholesterolemia: A systematic review and meta-analysis. *J. Psychosom. Res.* **2018**, *109*, 32–43. [CrossRef] [PubMed]
42. Fu, T.; Cao, H.; Yin, R.; Zhang, L.; Zhang, Q.; Li, L.; Feng, X.; Gu, Z. Depression and anxiety correlate with disease-related characteristics and quality of life in Chinese patients with gout: A case-control study. *Psychol. Health Med.* **2018**, *23*, 400–410. [CrossRef] [PubMed]
43. de Sousa, R.D.; Rodrigues, A.M.; Gregório, M.J.; Branco, J.D.C.; Gouveia, M.J.; Canhão, H.; Dias, S.S. Anxiety and depression in the Portuguese older adults: Prevalence and associated factors. *Front. Med.* **2017**, *4*, 196. [CrossRef] [PubMed]
44. Sivertsen, H.; Bjørkløf, G.H.; Engedal, K.; Selbæk, G.; Helvik, A.S. Depression and quality of life in older persons: A review. *Dement. Geriatr. Cogn. Disord.* **2015**, *40*, 311–339. [CrossRef] [PubMed]
45. Wariso, B.A.; Guerrieri, G.M.; Thompson, K.; Koziol, D.E.; Haq, N.; Martinez, P.E.; Rubinow, D.R.; Schmidt, P.J. Depression during the menopause transition: Impact on quality of life, social adjustment, and disability. *Arch. Women's Ment. Health* **2017**, *20*, 273–282. [CrossRef] [PubMed]
46. Freeman, E.W.; Sammel, M.D.; Boorman, D.W.; Zhang, R. Longitudinal pattern of depressive symptoms around natural menopause. *JAMA Psychiatry* **2014**, *71*, 36–43. [CrossRef] [PubMed]
47. Bromberger, J.T.; Matthews, K.A.; Schott, L.L.; Brockwell, S.; Avis, N.E.; Kravitz, H.M.; Everson-Rose, S.A.; Gold, E.B.; Sowers, M.; Randolph, J.F., Jr. Depressive symptoms during the menopausal transition: The study of women's health across the nation (SWAN). *J. Affect. Disord.* **2007**, *103*, 267–272. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

Spatial Analysis of Socio-Economic and Demographic Factors Associated with Contraceptive Use among Women of Childbearing Age in Rwanda

Faustin Habyarimana [1,2,*] and Shaun Ramroop [1]

1 School of Mathematics, Statistics and Computer Sciences, University of KwaZulu-Natal, Pietermaritzburg Campus, Private Bag X01, Scottsville 3209, South Africa; ramroops@ukzn.ac.za
2 College of Education, University of Rwanda, PO BOX 5039 Kigali, Rwanda
* Correspondence: singizwabb@gmail.com; Tel.: +27-033-260-5612

Received: 7 August 2018; Accepted: 25 October 2018; Published: 27 October 2018

Abstract: Contraceptive use is considered as essential for protecting women's health and rights, influencing fertility and population growth, and helping to promote economic development. The main objective of this study was to analysis the factors and spatial correlates of contraceptive use among women of childbearing age. The 2015 Rwanda Demographic and Health Survey (RDHS) data were used to identify the factors associated with contraceptive use in Rwanda. A Bayesian geo-additive model was used in order to account for fixed effects, nonlinear effects, spatial and random effects inherent in the data. The overall prevalence of use of any contraceptive method among married women of childbearing age in Rwanda was 52.7%. A woman's age, wealth quintile, level of education, working status, number of living children, and exposure to the media was found to increase contraceptive use. The findings from the study also found disparities in contraceptive use at provincial and district level, where prevalence was higher in districts of Northern provinces and lower in districts of western provinces. The findings of this study suggest that exposure to information on contraceptive use in health centres, empowerment of women to access quality contraceptive-use services and religions to play an important role in explaining and informing their adherents on the importance of using a contraceptive method.

Keywords: family planning; Bayesian; spatial; RDHS; Rwanda

1. Introduction

Contraceptive use is considered as essential for protecting women's health and rights, influencing fertility and population growth, and helping to promote economic development, especially in sub-Saharan Africa. Contraceptive use helps to avoid maternal deaths by preventing unwanted pregnancies and abortions. It also helps to determine the number of children in the family and enhances adequate child spacing [1,2]. Some contraceptive methods such as condoms help, not only for family planning purposes, but also to prevent sexually transmitted infections, for instance, HIV, among others.

The various studies in the literature point out that satisfying a woman's need for contraception, particularly in family planning, may considerably reduce fertility [3,4]. However, young women confront many problems in contraceptive method use, mainly in family planning services, for instance, fear of side effects, cost, and lack of enough information [5]. In general, contraception use and family planning are essential for improving the health of the population. Although many United Nations member countries, especially those in the developed world, have strong family planning programmes, this is not the case in some sub-Saharan African countries, some sub-regions of Asia and Latin America and in the Caribbean, where, regardless of a rise in the contraceptive-prevalence rate, a large number of mainly poor, uneducated women and those with limited access to family planning services, continue to have unmet needs for contraceptives [6,7].

The prevalence of any type of contraception-use method among married women in Rwanda has improved in the last fifteen years. It was 17.4% among married women in 2005, tripled (52%) between 2005 and 2010, and was 53% in 2014 [8]. This may be attributable to the Government of Rwanda's commitment to strengthen the health sector, especially family planning services. The maternal and child mortality rate in Rwanda has also tremendously declined over a period of 20 years. As shown by the Rwanda Demographic and Health Survey reports, in 2000, 2005, 2010 and 2015 there were 1 071, 750, 476 and 210 deaths per 100,000 live births respectively for maternal and for child mortality [8]. In neighboring countries, especially in East Africa, in the last 15 years contraceptive use has generally increased. The prevalence of current contraceptive use among married women in Uganda was 26% in 2004, 34% in 2010 and 38% in 2016; in Tanzania it was 26.4% in 2005, 34.4% in 2010 and 38.4% in 2015 and 2016; in Kenya, it was 39% in 2003, 46% in 2008 and 2009 and 58% in 2014; and in Burundi it was 27.7% in 2010 and 29% in 2016 and 2017 [9–12]. Identifying determinants of contraceptive use among women of reproductive age is essential in order to formulate adequate health programmes, policies and strategies and possible interventions that can promote the well-being of children and maternal health in general. We now consider the problem of contraceptive use.

Various studies in the literature assessed the determinants of modern contraceptive (pill, IUD, injections, condoms, female sterilization, male sterilization, implants/norplant, lactational amenorrhea, standards days methods) use only [13–15] and did not include traditional methods (withdrawal, periodic abstinence) which may sometimes be the most accessible. The current study addresses this problem by combining all types of contraceptive-use methods among married women of childbearing age, nationwide. In addition, the RDHS data set has inherent nonlinear, spatial and random effects that needed to be accounted for. We consider these aspects to capture the heterogeneity of the data. Consequently, the main objectives of the current study were to address these problems in their entirety, by analyzing the factors associated with any type of contraceptive use and mapping the spatial distribution of use at district level in Rwanda. These objectives were achieved by applying structured spatial modelling that accounted for fixed, nonlinear, spatial and random effects. To the best of our knowledge, there is no study in the literature that assessed the determinants of contraceptive use among women in Rwanda, using a structured spatial model. It was expected that the findings from the present study would help policy makers and other public health institutions in Rwanda to visualize the spatial distribution of the use of any type of contraceptive method among women of reproductive age at district level. Therefore, it would help them to improve the current strategies when targeting the districts of low prevalence of contraceptive use.

2. Materials and Methods

2.1. Source of Data

The current study used data from the 2014/15 Rwanda Demographic and Health Survey (RDHS). The sampling used in this survey was a two-stage stratified method; in the first stage, 492 primary sampling areas were selected, whose 113 were selected from urban areas and 379 from rural areas. The sampling was done with probability proportional to the number of households in the village. In the second stage, systematic sampling was used for all households existing in the selected village and 26 households were selected in each village. More details on sampling techniques and data collection can be found in [8].

2.1.1. Outcome Variable

The different types of contraceptive methods used are the modern (the pill, IUD, injection, male condom, female condom, female sterilization, male sterilization, implants/Norplant, lactational amenorrhea (LAM), the standard days method, periodic abstinence), traditional (withdrawal and periodic abstinence). In the survey, women were asked whether they used any of these methods. In this study, a woman is considered as a current contraceptive user if she uses at least one of the above-mentioned methods and as a non-contraceptive user if otherwise, and this was coded as "1" and "0" for contraceptive use and non-contraceptive use respectively.

2.1.2. Independent Variables

The independent variables used in the current study have been used elsewhere [6,16–19]. The various socio-economic and demographic factors considered include current age of the respondent (continuous); the age at first cohabitation (continuous), level of education (no education, primary, secondary, tertiary); have heard about family planning in the last 12 months from newspapers/magazines (yes or no); have heard about family planning in the last 12 months on television (yes or no); have been visited by a family planning worker during the last 12 months (yes or no); husband or partner desire for more children compared to the wife or partner (same, husband wants more, husband wants fewer, do not know); wealth quintile of her household (poorest, poorer, middle, richer, richest); currently working (yes or no); province of residence (Kigali, South, West, North, East); place of residence (urban, rural); religious affiliation (Catholic, Protestant, Seventh Day Adventist, other); husband or partner's education level (no education, primary, secondary, tertiary); decision-making for using contraceptive (mainly respondent, mainly husband or partner, joint decision, other); person who usually decides on healthcare (respondent alone, respondent and husband/partner, respondent and other person, husband/partner alone, someone else, other); number of living children (0, 1, 2, 3, 4 and more); and age of the husband or partner.

2.2. Statistical Model

Let y_{ijm} denote the contraceptive use status of woman, i from stratum j and cluster m, with $i = 1, 2, 3, \ldots, 6847$, $j = 1, 2, 3, \ldots, 60$ and $m = 1, 2, 3, \ldots, 492$. The outcome variable is defined as a dichotomous variable such that $y_{ijm} = 1$ if the women i is currently using any type of contraceptive method and $y_{ijm} = 0$ if the women i is not currently using any type of contraceptive method. The contraceptive use status among women of reproductive age is a binary outcome in the current study and hence it is assumed to follow a Bernoulli distribution:

$$y_{ijm} \sim Bernoulli(p_{ij}) \tag{1}$$

where p_{ij} is the probability that a woman i from district j is currently using a contraceptive method and $1 - p_{ij}$ is the probability that a woman i from district j is not currently using a contraceptive method. Therefore, the use of contraceptive methods among women of reproductive age can be associated with the explanatory variables using an appropriate link function from a generalized linear models approach as follows:

$$logit\ (p_{ij}) = \log\left(\frac{p_{ij}}{1 - p_{ij}}\right) = W'_{ij}\beta \tag{2}$$

Model (2) is known as binary logistic regression, where W'_{ij} is the vector of explanatory variables and β is the vector of coefficient parameters. However, classical Generalized Linear model (GLM) has a rigid assumption that all observations are independent; but this assumption is sometimes not satisfied, as some observations may have, for instance, spatial dependence, or may have nonlinear effects. Hence, there is a need to include nonlinear and spatial variability in model (2) and it is given by:

$$logit\ (p_{ij}) = W'_{ij}\beta + \sum_{k=1}^{q} f_k\left(x_{ijk}\right) + f_{spat}(s_j) \tag{3}$$

where β_i is the vector of fixed effect corresponding to categorical variables, f_k is the appropriate smooth function of continuous variables such as mother's current age and mother's age at first cohabitation and $f_{spat}(s_j)$ are the parameters of random effects, which capture unobserved spatial heterogeneity at district s_j.

2.2.1. Parameter Estimation

In the current study, the parameter estimation is done based on a full Bayesian analysis framework, where the appropriate prior distributions have to be assigned to all unknown parameters. In this study, diffuse priors $p(\beta) \propto const$ are assigned to all fixed regression parameters and the second-order Gaussian random walk priors were assigned to non-parametric continuous covariates [20]. The structured spatial effects s_i were modeled through a Gaussian Markov random field specified as an intrinsic conditional autoregressive (ICAR) prior distribution [21].

2.2.2. Posterior and Inferences

Posterior distributions are the distribution of the parameters after observing the data and are obtained by updating the prior distribution with observed data. A full Bayesian inference is based on the analysis of the posterior distribution of the model parameters. In the current study, the Bayesian posterior marginal distribution was estimated, based on integrated nested Laplace Approximation [22] and R-INLA was used for inferential analysis.

2.2.3. Model Selection

The model goodness of fit was assessed based on the Deviance Information Criteria (DIC) that states that the smaller the value of the DIC the better is the model fit [23,24] and the DIC value is given by:

$$\text{DIC} = \overline{D} + \text{pD} \tag{4}$$

where $\overline{D}$ is the posterior mean of the deviance and pD is the number of effective parameters in the model that penalize the complexity of the model. The DIC takes both the model fit (summarized by $\overline{D}$) and the model complexity (captured by pD) into consideration when comparing models.

2.2.4. Model Building

The present study fitted and examined the following models:

Model 1: $logit\ (p_{ij}) = W'_{ij}\beta$
Model 2: $logit(p_{ij}) = W'_{ij}\beta + f_1(current\ age_i) + f_2(age\ at\ first\ cohabitation_i)$
Model 3: $logit(p_{ij}) = W'_{ij}\beta + f_{spatj}(district_j),\ j = 1, 2, \ldots, 30$
Model 4: $logit\ (p_{ij}) = W'_{ij}\beta + f_1(current\ age_i) + f_2(age\ at\ first\ cohabitation_i) + f_j(district_j),\ j = 1, 2, \ldots, 30$

Model 1 is classical logistic regression, where all categorical variables (women's education, women's region affiliation, women's working status, women's province of residence, have heard about family planning on radio in last 12 months, have heard about family planning in newspapers/magazines in last 12 months, visited by a family planning worker in the last 12 months, visited a health facility in the last 12 months, currently residing with husband or not, number of living children, wealth quintile of the household, husband desires children) and (current age of the woman and woman's age at first cohabitation) were considered as fixed effects and assumed to have a linear effect on the outcome variable.

In model 2, categorical variables listed earlier in model 1 were assumed to have a linear effect on the response variable, whereas continuous variables were modeled non-parametrically and model 2 is commonly known as an additive logistic regression model. In model 3, all predictor variables were modeled as fixed effects and structured random effects as structured spatial effects that cover the unobserved covariates which are essential within the districts. Model 4 is an extension of model 2, including structured spatial effects and is known as a structured additive regression model.

3. Results

3.1. Descriptive Analysis

Table 1 presents descriptive statistics of the participants. The overall prevalence of any contraceptive use among married women of reproductive age in Rwanda was 52.7%, of whom 46.8% use modern methods and 5.9% use traditional methods. The average current age of women was 32.8 years, with a minimum age of 16 and a maximum age of 49 years. The average age of first cohabitation was 21 years, the minimum age was 10 and maximum age was 95. Most of the respondents were from rural areas (83%) and 17.0% were from urban areas. It is also observed in the same table that 55.0% of respondents had heard about family planning on the radio, 7.1% on TV and 6.1% in newspapers or magazines. Table 1 shows that 29.2% of the respondents were visited by a family planning planner in the 12 months prior to the survey, 71.8% of the respondents visited health facilities in the 12 months prior to the survey and 46.3% were told about family planning at a health facility. The majority of the respondents (86.8%) were working during the period of the survey. It can also be observed that most of the respondents had primary education (70.5%) and most of the women had two living children (45.2%). It is also observed that 38.2% of the respondents were Catholic, 45.2% were Protestant, 3.2% were from religions other than Catholic, Protestant or Seventh Day Adventist and 12.8% were Seventh Day Adventist.

Table 1. Descriptive statistics of the participants' details.

Variable	Categories	% or M or Range
Current contraceptive use	Yes No	52.7 47.3
Respondent current age	Continuous	M = 32.81 (Minimum = 16, maximum = 49)
Age at first cohabitation	Continuous	M = 21.09 (Minimum = 10, Maximum = 95)
Place of residence	Urban Rural	17 83
Respondent education level	No education Primary Secondary Tertiary	16.5 70.5 10.3 2.7
Wealth index	Poorest Poorer Middle Rich Richest	18.8 21.2 20.9 19.9 19.3
Heard about family planning on the radio in last few months	Yes No	55 45
Heard about family planning on the TV in last few months	Yes No	7.1 92.8
Heard about family planning from newspapers/magazines in the last few months	Yes No	6.1 93.9
Visited by family planning worker in the last 12 months	Yes No	29.2 70.8
Visited health facility in the last 12 months	Yes No	71.8 28.2

Table 1. *Cont.*

Variable	Categories	% or M or Range
At health facility, told of family planning	Yes	46.3
	No	53.7
Respondent currently working	Yes	86.8
	No	13.2
Husband/partner education level	No education	16.9
	Primary	70.1
	Secondary	9.2
	Tertiary	3.5
	Do not know	0.2
Number of living children	0	5.2
	1	17.9
	2	45.2
	3	17.4
	4 or more	37.4
Religion	Catholic	38.2
	Protestant	45.6
	Seventh Day Adventist	12.8
	Others	3.2

It is observed from Table 2 that modern contraceptive methods were the most (46.9%) used among married women of childbearing age in Rwanda and traditional method use was 5.9%.

Table 2. Prevalence of contraceptive use by type.

Current use by Method Type	N (%)
No method	3237 (47.3)
Traditional method	402 (5.9)
Modern method	3208 (46.8)

The association between current contraceptive use among married women or women living with their partners, and various potential factors, was tested by the Chi-square statistical test at 5% level of significance and the results are summarized in Table 3. Education level of the respondents was significantly associated with current use of contraceptive methods (p-value < 0.0001). It can be seen from the table that the prevalence of contraceptive use was 56.8%, 53.5%, 53.7% and 47.3% among women with tertiary education, secondary and primary education level respectively. The working status of the woman was significantly associated with current use of contraceptive methods among women of reproductive age (p-value < 0.0001). It can be observed from the table that the prevalence of contraceptive use was 53.6% among women who were working at the time of the survey. Hearing about family planning on the radio, from newspapers or magazines and on the TV was each significantly associated with contraceptive use (p-value < 0.0001) and the prevalence of contraceptive use in this regard was 55.8%, 63.9% or 60.1% among the women respectively. Religion was significantly associated with contraceptive use among women. The prevalence of contraceptive use was 56.9% among Catholic women, 47.5% among Protestant women, 58.9% among Seventh Day Adventist women and 53.4% among women from religions other than Catholic, Protestant and Seventh Day Adventist. The issue of education and its influence on contraceptive use is also a salient finding of the current research. The education of the husband was significantly associated with current contraceptive use among women (p-value = 0.018). The prevalence of contraceptive use was 48.1%, 53.7%, 53.0% and 54.2% among women married to husbands with no education, primary education, secondary education and tertiary education respectively.

Table 3. The prevalence of contraceptive use among women of childbearing age by category.

Indicator Variable	Category	Contraceptive Use		p-Value
		No	Yes	
Woman's education level	No education Primary Secondary Tertiary	596(52.7%) 2233(46.3%) 329(46.5%) 79(43.2%)	535(47.3%) 2592(53.7%) 379(53.5%) 104(56.8%)	<0.001
Respondent currently working	Yes No	2755(46.4%) 481(53.0%)	3183(53.6%) 426(47.0%)	<0.0001
Heard about family planning on the radio in the last few months	Yes No	1666(44.2%) 1571(51.0%)	2100(55.8%) 1509(49.0%)	<0.0001
Heard about family planning from newspapers/magazines in the last 12 months	Yes No	150(36.1%) 3086(48.0%)	265(63.9%) 3344(52.0%)	<0.0001
Heard about family planning on TV in the last few months	Yes No	195(39.9%) 3042(47.9%)	294(60.1%) 3314(52.1%)	0.001
Woman's religious affiliation	Catholic Protestant Seventh Day Adventist Other	1129(43.1%) 1637(52.5%) 361(41.1%) 103(46.6%)	1489(56.9%) 1482(47.5%) 517(58.9%) 118(53.4%)	<0.0001
Visited by family planning worker in the last 12 months	Yes No	813(40.7%) 2424(50.0%)	1184(59.3%) 2426(50.0%)	<0.0001
Visited health facility in the last 12 months	Yes No	2379(48.4%) 858(44.5%)	2539(51.6%) 1071(55.5%)	0.004
At health facility, told about family planning	Yes No	1081(47.5%) 1298(49.1%)	1194(52.5%) 1345(50.9%)	0.265
Household wealth index	Poorest Poorer Middle Rich richest	667(51.7%) 730(50.4%) 654(45.8%) 599(44.1%) 586(44.4%)	622(48.3%) 719(49.6%) 774(54.2%) 760(55.9%) 735(55.6%)	<0.0001
Place of residence	Urban Rural	516(44.3%) 2720(47.9%)	648(55.7%) 2962(52.7%)	0.028
Number of living children	0 1 2 3	349(98.3%) 607(49.6%) 498(42.7%) 1137(4.3%)	6(1.7%) 618(50.4%) 867(57.3%) 1427(55.7%)	<0.0001
Husband/partner desires children	Same number More Fewer Do not know	430(51.9%) 1886(45.5%) 563(46.0%) 358(55.2%)	398(48.1%) 2261(54.5%) 660(54.0%) 290(44.8%)	<0.0001
Husband/partner education	No education Primary Secondary Tertiary Don't know	601(51.9%) 2222(46.3%) 295(47.0%) 111(45.2%) 6(50.0)	558(48.1%) 2579(53.7%) 333(53.0) 132(54.2%) 6(50.0%)	0.018

We now consider the model fit comparisons. The findings from Table 4 reveal that model 4 has smaller DIC compared to the DIC of model 1, model 2 and model 3. Therefore based on the principles of the DIC (that state that the smaller the DIC the better the model fit [23]), model 4 was found to be the better model fit, and it is therefore used in the final analysis in this study.

Table 4. Model comparison based on Deviance Information Criteria (DIC).

Statistics	Model 1	Model 2	Model 3	Model 4
DIC	8488.64	8560.34	8465.45	8462.39
$\overline{D}$	8428	8486.14	8388	8362.12
pD	29.92	37.10	38.73	50.133

3.2. Fixed Effect

In the multivariate structured geo-additive model, any variable that was statistically significant in cross-tabulation at 5% level of significance were included in the analysis and the final analysis retained only those variables that were at 5% level of significance.

The results from fixed effects are summarized in Table 5. It is observed that the contraceptive use among women of reproductive age in Rwanda increases with higher wealth and education levels, exposure to mass media and number of living children. The education level of the woman is highly associated with contraceptive use among women of reproductive age. It was observed from the table that a woman with tertiary education was 79.3% (OR = 1.7934, 95% CI: 1.2355, 2.61300) more likely to use contraceptive methods compared to woman with no education. A woman with secondary education was 1.34 (OR = 1.3371, CI: 1.0853, 1.7331) more likely to use contraceptive methods compared to a woman without education, while a woman with primary education was 1.17 (OR = 1.1735, CI: 1.0125, 1.3599) more likely to use contraceptive methods compared to a woman without education.

Table 5. Summary of fixed effects of factors associated with women of childbearing age.

Variable	Posterior Estimate of the Mean	Posterior Standard Error	Odds Ratio	95% Credible Interval (CI)
Intercept	−5.3585	0.4743	0.0047	(0.0017 0.0112)
Woman's education (No education = reference)			1.0000	
Tertiary	0.5841	0.1909	1.7934	(1.2355 2.61300)
Secondary	0.3157	0.1193	1.3371	(1.0853 1.7331)
Primary	0.1600	0.0752	1.1735	1.0125 1.3599)
Wealth quintile (Poorest = reference)			1.0000	
Richest	0.3598	0.1023	1.4330	(1.1727 1.7517)
Richer	0.2998	0.0887	1.3496	(1.1340 1.6063)
Middle	0.2201	0.0857	1.2462	(1.0533 1.4745)
Poorer	0.0703	0.0834	1.0728	(0.9107 1.2636)
Visited health facility in the last 12 months (No = reference)				
Yes	0.1960	0.0598	1.2165	(1.0820 1.3681)
Currently residing with husband (Living = reference)				
Stays elsewhere	−0.5811	0.0882	0.5593	(0.4702 0.6647)
Number of living children (4 and more = reference)				
3	−0.1217	0.0837	0.8854	(0.7512 1.0435)
2	−0.3055	0.0935	0.7368	(0.6131 0.8848)
1	−0.8057	0.1176	0.4468	(0.3544 0.5625)
0	−4.9371	0.4299	0.0072	(0.0029 0.0156)
Religion (Others = reference)				
Catholic	−03459	0.0585	0.7047	(0.6308 0.7936)
Protestant	0.1338	0.0854	1.1432	(0.967 1.3523)
Seventh Day Adventist	0.0211	0.1513	1.0213	(0.7600 1.3761)
Respondent currently working (No = reference)				
Yes	0.2394	0.0795	1.2705	(1.0868 1.4850)
Heard about family planning on the radio in the last few months (No = reference)				
Yes	0.1109	0.0554	1.1173	(1.0021 1.2455)

Table 5. *Cont.*

Variable	Posterior Estimate of the Mean	Posterior Standard Error	Odds Ratio	95% Credible Interval (CI)
Visited by family planning worker in the last 12 months (Yes = reference)				
No	0.1960	0.0598	1.21653	(1.0812 1.3681)
Heard about family planning from newspapers/magazines (No = reference)				
Yes	0.2884	0.1197	1.3343	(1.0564 1.6896)
Province (West = reference)				
Kigali	0.0548	0.1047	1.0563	(0.4702 1.5038)
South	−0.1224	0.1066	0.8848	(0.8601 1.2969)
North	0.3580	0.1141	1.4305	(1.1436 1.7898)
East	0.1704	0.1051	1.1858	(0.9649 1.4579)
Husband desires children (Same number = reference)				
More	0.2622	0.0818	1.2998	(1.1069 1.5258)
Fewer	−0.0456	0.1153	0.9554	(0.7619 1.1979)
Do not know	0.2073	0.0964	1.2304	(1.0182 1.4866)

The wealth of the family was found to be a significant predictor of women's contraceptive use. Women from richest (OR = 1.4330, CI: 1.1727, 1.7517), richer (OR = 1.3496, CI: 1.1340, 1.6063) and middle (OR = 1.2462, CI: 1.0533, 1.4745) families were more likely to use contraceptive methods than women from poorest families.

Considering health facilities, a woman who visited a health facility within 12 months prior to the survey was 1.2165 (OR = 1.2165, 95% CI: 1.0882, 1.3681) more likely to use any type of contraceptive method compared with a woman who did not visit a health facility within 12 months.

It was also observed from the results that a woman who was not residing with a husband or partner at the time of the survey was 0.559 (OR = 0.5593, 95% CI: 0.4702, 0.6647) times less likely to use contraceptive methods compared to a woman who was residing with a husband or partner. It was also observed that contraceptive use increases with the number of living children in the family. A woman who had no living child was 0.0072 (OR = 0.0072, 95% CI: 0.0029, 0.0156) less likely to use a contraceptive method compared with a woman who had four or more living children. A woman who had one living child was 0.447 (OR = 0.4468, 95% CI: 0.3544, 0.5625) less likely to use a contraceptive method compared to a woman who had four or more living children. A woman who had two living children was 0.737 (OR = 0.7368, 95% CI: 0.6131, 0.8848) less likely to use a contraceptive method compared to a woman who had four or more living children. A Catholic woman was 0.7047 (OR = 0.7047, 95% CI: 0.6308, 0.7936) less likely to use a contraceptive method compared to women from other religions other than Protestant and Seventh Day Adventist. A working woman was 1.27 (OR = 1.2705, 95% CI: 1.0868, 1.4850) more likely to use a contraceptive method than a non-working woman. A woman who had heard about family planning on the radio was 1.1173 (OR = 1.1173, 95% CI: 1.0021, 1.2455), on TV was 1.2200 (OR = 1.2200, 95%CI: 1.0812, 1.3681) and from newspapers/magazines was 1.3343 (OR = 1.3343, 95%CI: 1.0564, 1.6896) more likely to use a contraceptive method than a woman who had not heard about family planning within 12 months. Contraceptive use is higher in the Northern Province compared to other provinces. A woman from Northern Province was 1.43 (OR = 1.4305, CI: 1.1436, 1.7898) more likely to use a contraceptive method than a woman from the Western Province, but a woman from the Southern Province, Kigali and East did not show any statistical association. A woman whose husband desired more children than her was 1.2998 (OR = 1.2998, 95% CI: 1.1069, 1.5258) more likely to use a contraceptive method than a woman whose husband desired the same number of children. A woman who did not know whether her husband or partner desired the same, less or more children was 1.230 (OR = 1.2304, 95% CI: 1.0182, 1.4866) more likely to use a contraceptive method than a woman whose husband desired the same number of children.

3.3. Non-Linear Effect

The current study also considered the non-linear effects from continuous variables on contraceptive use and results from the structured multivariate model are summarized in Figures 1 and 2. It is observed from Figure 1 that contraceptive use among women of reproductive age decreases with age. It can be observed from Figure 2 that contraceptive use follows an inverse U-shape, where the contraceptive use increases with the woman's age at first cohabitation, up until close to 25 years old and afterward decreasing with age.

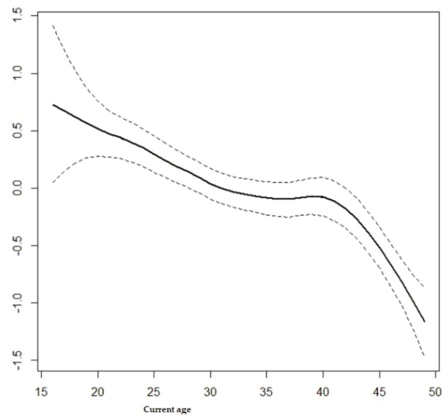


Figure 1. Effects of woman's current age on the use of contraceptive methods.

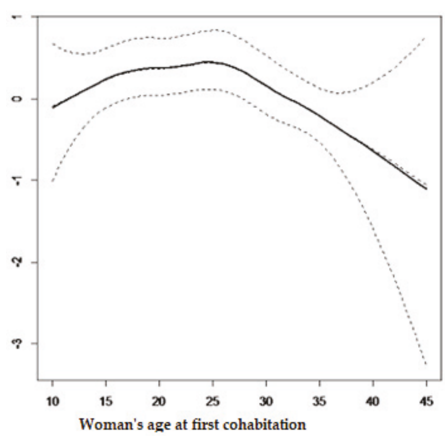


Figure 2. Effects of woman's age at first cohabitation.

3.4. Spatial Effects

The present study found positive and negative structured spatial effects on contraceptive use among women of reproductive age in Rwanda. Figure 3 presents structured spatial effects on contraceptive use among women of childbearing age. The light blue colour shows low contraceptive use while dark blue shows high contraceptive use. The numbers (1-30) shown in the map correspond to the districts codes indicated in Table 6.

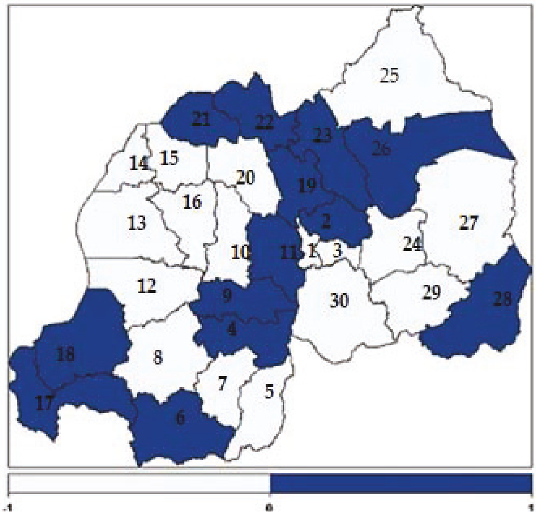


Figure 3. Structured spatial effects on contraceptive use among women of childbearing age.

Table 6. Districts of Rwanda and their codes used in RDHS 2014/15.

Code	District	Code	District	Code	District	Code	District
1	Nyarugenge	9	Ruhango	17	Rusizi	24	Rwamagana
2	Gasabo	10	Muhanga	18	Nyamasheke	25	Nyagatare
3	Kicukiro	11	Kamonyi	19	Rulindo	26	Gatsibo
4	Nyanza	12	Karongi	20	Gakenke	27	Kayonza
5	Gisagara	13	Rubavu	21	Musanze	28	Kirehe
6	Nyaruguru	14	Rubavu	22	Burera	29	Ngoma
7	Huye	15	Nyabihu	23	Gicumbi	30	Bugesera
8	Nyamagabe	16	Ngororero				

4. Discussion

The main objective of this study was to identify the factors associated with contraceptive use among women of childbearing age in Rwanda and to identify and map the possible spatial distribution of contraceptive use at district level. Figure 3 shows that contraceptive use was unequally distributed among districts, being higher in districts of the Northern Province and lower in districts of the Western Province. This finding is similar to the findings in [8]. In the dark blue coloured district a great number of women use contraceptive methods and in the light blue coloured districts a low number of women use contraceptive methods. Figure 3 shows that contraceptive use among women in the dark blue coloured areas was higher, mostly in four districts of the Northern Province (Musanze, Burera, Rulindo, Gicumbi), in four districts from the Southern Province (Kamonyi, Ruhango, Nyanza and Nyaruguru), in two districts from the Eastern Province (Kirehe and Gatsibo), in two districts from Western Province (Nyamasheke and Rusizi) and in Gasabo district in Kigali City [8].

The level of wealth of the family was found to be a significant predictor of contraceptive use among women of reproductive age. The findings from this study indicated that contraceptive use increases with greater wealth of the household. This finding is similar to that of [8,25,26]. This may be due to the fact that the wealthier families can easily access health centre facilities, mass media and education, among other factors, all of which are well known for increasing the use of contraceptives among women.

Exposure to the use of contraceptives was found to play an essential role in its use as well as in family planning. Similar findings were also found in studies by [15,19,27,28]. This may be due to the fact that, for instance, when a woman visits a health centre or is visited by health centre workers or family planning field workers, she is told about the importance of using contraceptive methods. This may explain why a woman in such cases was more likely to use a contraceptive method.

The current findings show a strong association between the number of living children in the family and use of contraceptive methods. It was observed from the results that contraceptive use increases with increasing numbers of living children. This was also found by [16,29–32]. A woman who has many children tends to limit births because of financial burdens of paying school fees and providing health care, among other factors.

The findings from this study pointed towards a significant relationship between religious affiliation and the use of a contraceptive method among women of reproductive age. This was also found in the studies by [33]. In this study, the use of contraceptives was highest among Protestant women, followed by Seventh Day Adventists. Among Catholic women there was not a significant association with contraceptive use compared to women from religions other than Protestant and Seventh Day Adventist. However, in some studies such as [34,35], a significant relationship was found.

The findings from the current study also highlighted a strong association between the education level of the women and contraceptive use. The results showed that the use of any type of contraceptive method increases with women's level of education i.e., educated women are more frequent users of contraceptive use than uneducated women. This result was found in other similar studies [26,27,36–38]. However, studies by [26] in Mali found no significant association between education level of the women and contraceptive use. The significant association is not surprising because education is known as a powerful factor associated with women's empowerment, knowledge about their body and reproductive physiology and other maternal and child health information. The educated women are most likely to perceive the advantage of having few children and its effect on the family or their individual economic productivity.

The results from the current study also revealed a strong association between working status of the women and the use of contraceptive methods among those of childbearing age in Rwanda. The women that were working at the time of the survey were more likely to use contraceptives compared to women who were not working over this period. This was found elsewhere in similar studies [33]. This may be due to the fact that working status sometimes empowers women, not only financially, but also to access contraception information, maternal and child healthcare and this empowerment promotes contraceptive use, especially for family planning.

As was expected, this study revealed a strong, significant association between contraceptive use and marital status. It was found that women who resided with their husband/partner were more likely to use any type of contraceptive methods than women who did not reside with their husband or partner. This was found in other studies by [27].

In light of the findings of the current study, the use of any contraceptive method reduces with increasing age, up to 25 years old, and thereafter, increases up to 42 years old, but thereafter decreases once more. In many similar studies in the literature it was also found that the use of contraceptives reduces with the increasing age of the women [32,36,39]. The low contraceptive use among women from the less than 30-year-old age group may be due to the fact that they are usually newly married and interested in having children in the first years of their marriage. The higher prevalence rate of contraceptive use among women aged between 30 and 42 years old may indicate that most of the women in this age group have reached their desired number of children. It is not surprising that after the age of 42 years, the use of contraceptive methods decreases, likely because many women at this age may not be sexually active. This may be due to the fact that a decrease in fecundity correlates with the onset of menopause.

The findings from the current study revealed that the age of the women at first cohabitation was a very important factor associated with the use of contraceptives among women of childbearing age. This was found elsewhere by [39]. Figure 2 showed that cohabitating at an early age increases the use of contraceptive methods. It was found that contraceptive use increases in the first cohabitation age group of 10 to 25 years old. This may be due to the fact that at this age a higher number of women are not yet married and may not desire children.

5. Study Limitations

The analysis in this study was mainly based on a cross-sectional study and this may not draw the causal relationship or effect between contraceptive use and independent variables; only associations can be drawn from this study. Therefore, a longitudinal study is suggested for future work in order to identify relevant trends and patterns over time. Furthermore, the study used on married woman only and not on unmarried women. Another possible area for future study is to differentiate types of birth control, as not all have equal effectiveness.

6. Conclusions

The findings of this study suggest that improvement of exposure to contraceptive-use information in health centres, empowerment of women to access quality contraceptive-use services and religious affiliation all play a significant role in explaining and informing their adherents on the importance of using a contraceptive method. These findings also highlighted the districts with lower numbers of women using contraceptive methods and this can help policy makers and other related public health institutions to design specific programmes targeting these districts in order to improve the health status and living conditions of these women.

Author Contributions: F.H. acquired data, performed analysis and drafted the manuscript, S.R. revised and edited the manuscript and finally all the authors approved the manuscript.

Acknowledgments: The authors acknowledge the National Institute of Statistics of Rwanda (NISR) [Rwanda], the Ministry of Health (MOH) [Rwanda], and ICF International for the data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kayode, G.A.; Adekanmbi, V.T.; Uthman, O.A. Risk factors and a predictive model for under-five mortality in Nigeria: Evidence from Nigeria demographic and health survey. *BMC Pregnancy Childbirth* **2012**, *12*, 10. [CrossRef] [PubMed]
2. Palloni, A.; Pinto Aguirre, G.; Lastiri, S. The effects of breast-feeding and the pace of childbearing on early childhood mortality in Mexico. *Bull. Pan. Am. Health Organ.* **1994**, *28*, 93–111. [PubMed]
3. Westoff, C.F. *Unmet Need at the End of the Century*; DHS Comparative Reports No. 1; ORC Macro. Inc.: Calverton, MD, USA, 2001.
4. Sinding, S.W.; Ross, J.A.; Rosenfield, A.G. Seeking common ground: Unmet need and demographic goals. *Int. Fam. Plan. Perspect.* **1994**, *1*, 23–32. [CrossRef]
5. Blanc, A.K.; Tsui, A.O.; Croft, T.N.; Trevitt, J.L. Patterns and trends in adolescents' contraceptive use and discontinuation in developing countries and comparisons with adult women. *Int. Perspect. Sex. Reprod. Health* **2009**, *35*, 63–71. [CrossRef] [PubMed]
6. Cleland, J.; Bernstein, S.; Ezeh, A.; Foundes, A.; Glasier, A.; Innis, J. Family planning: The unfinished agenda. *Lancet* **2006**, *368*, 1810–1827. [CrossRef]
7. Singh, S.; Darroch, J.E. *Adding it up: Costs and Benefits of Contraceptive Services*; Guttmacher Institute and UNFPA: New York, NY, USA, 2012.
8. National Institute of Statistics of Rwanda (NISR) [Rwanda]; Ministry of Health(MoH) [Rwanda]; ICF International. *Rwanda Demographic and Health Survey 2014–15*; NISR: Rockville, MD, USA; MoH: Rockville, MD, USA; ICF International: Rockville, MD, USA, 2015.
9. KNBS. *Kenya Demographic and Health Survey 2008–09*; KNBS: Nairobi, Kenya, 2010.

10. Ministry of Health, Community Development; Gender, Elderly and Children (MoHCDGEC) [Tanzania Mainland]; Ministry of Health (MoH) [Zanzibar], National Bureau of Statistics (NBS); Office of the Chief Government Statistician (OCGS); ICF. *Tanzania Demographic and Health Survey and Malaria Indicator Survey (TDHS-MIS) 2015–16*; MoHCDGEC: Dar es Salaam, Tanzania; Rockville, MD, USA; MoH: Dar es Salaam, Tanzania; Rockville, MD, USA; NBS: Dar es Salaam, Tanzania; Rockville, MD, USA; OCGS: Dar es Salaam, Tanzania; Rockville, MD, USA; ICF: Dar es Salaam, Tanzania; Rockville, MD, USA, 2016.
11. UBOS; ICF International Inc. *Uganda Demographic and Health Survey 2011*; Uganda Bureau of Statistics: Kampala, Uganda; ICF International Inc.: Claverton, UK, 2012.
12. Ministère à la présidence chargé de la Bonne Gouvernance et du Plan[Burundi] (MPBGP); Ministère de la santé Publique et de la lutte contre le sida [Burundi] (MSPLS); Institut de Statistiques et d'Etudes Economiques du Burundi (ISTEEBU), et ICF International. *Triosième Enquête Démographique et de Santé*; ISTEEBU, MSPLS, et ICF International: Bujumbura, Burundi, 2017.
13. Wang, W.; Alva, S.; Winter, R.; Burgert, C. *Contextual Influences of Modern Contraceptive Use among Rural Women in Rwanda and Nepal, DHS Analytical Studies No 41*; ICF International: Calverton, MD, USA, 2013.
14. Brunie, A.; Tolley, E.E.; Ngabo, F.; Wesson, J.; Chen, M. Getting to 70%: Barriers to modern contraceptive use for women in Rwanda. *Int. J. Gynecol. Obstet.* **2013**, *123*, e11–e15. [CrossRef] [PubMed]
15. Frederic, T.M.; Phoibe, K.; Ntaganira, J. Assessment of Knowledge, Attitudes, and Practice on Contraceptive Use among Women Attending Family Planning Services in Some Health Centers of Muhima District Hospital, Rwanda. *Open Sci. J.* **2017**, *2*. [CrossRef]
16. Adeyemi, A.S.; Olugbenga-Bello, A.I.; Adeoye, O.; Salawu, M.; Aderinoye, A.; Agbaje, M. Contraceptive prevalence and determinants among women of reproductive age group in Ogbomoso, Oyo state, nigeria. *Open Access J. Contracept.* **2016**, *7*, 33–41. [PubMed]
17. Hailu, T.G. Determinants and Cross-Regional Variations of Contraceptive Prevalence Rate in Ethiopia: A Multilevel Modeling Approach. *Am. J. Math. Stat.* **2015**, *5*, 95–110.
18. Nonvignon, J.; Novignon, J. Trend and determinants of contraceptive use among women of reproductive age in Ghana. *Etude De La Popul. Afr.* **2014**, *28*, 956–967. [CrossRef]
19. Osmani, A.K.; Reyer, J.A.; Osmani, A.R.; Hamajiman, N. Factors influencing contraceptive use among women in Afghanistan: Secondary analysis of Afghanistan Health Survey 2012. *Nagoya J. Med Sci.* **2015**, *77*, 551–561. [PubMed]
20. Wecker, W.E.; Ansley, C.F. The signal extraction approach to nonlinear regression and spline smoothing. *J. Am. Stat. Assoc.* **1983**, *78*, 81–89. [CrossRef]
21. Besag, J.; Green, P.J. Spatial statistics and Bayesian computation. *J. R. Stat. Soc. Ser. B (Methodol.)* **1993**, *55*, 25–37.
22. Rue, H.; Martino, S.; Chopin, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **2009**, *71*, 319–392. [CrossRef]
23. Spiegelhalter, D.J.; Best, N.G.; Carlin, B.P.; Van der Linde, A. Bayesian measures of model complexity and fit. *J. R. Stat. Soc. B* **2002**, *64*, 583–639. [CrossRef]
24. Best, N.; Richardson, S.; Thomson, A. A comparison of Bayesian spatial models for disease mapping. *Stat. Methods Med. Res.* **2005**, *14*, 35–59. [CrossRef] [PubMed]
25. Okech, T.C.; Wawire, N.W.; Mburu, T.K. Contraceptive use among women of reproductive age in Kenya's city slums. *Int. J. Bus. Soc. Sci.* **2011**, *2*, 22–43.
26. Solanke, B.L. Factors influencing contraceptive use and non-use among women of advanced reproductive age in Nigeria. *J. Heal. Popul. Nutr.* **2017**, *36*, 1. [CrossRef] [PubMed]
27. Nyarko, S.H. Prevalence and correlates of contraceptive use among female adolescents in Ghana. *BMC Women's Health* **2015**, *15*, 60. [CrossRef] [PubMed]
28. Tekelab, T.; Melka, A.S.; Wirtu, D. Predictors of modern contraceptive methods use among married women of reproductive age groups in Western Ethiopia: A community based cross-sectional study. *BMC Women's Health* **2015**, *15*, 52. [CrossRef] [PubMed]
29. Lasee, A.; McCormick, J.B. Demographic and socio-economic determinants of contraceptive use in a low income community of Karachi. *J.-Pak. Med. Assoc.* **1996**, *46*, 228–230. [PubMed]
30. Hussain, N. Demographic, Socio-Economic and Cultural Factors Affecting Knowledge and Use of Contraception Differentials in Malda District, West Bengal. *J. Community Med. Health Educ.* **2011**, *1*, 2. [CrossRef]

31. Anasel, M.G.; Mlinga, U.J. Determinants of contraceptive use among married women in Tanzania: Policy implication. *Etude De La Popul. Afr.* **2014**, *28*, 978–988. [CrossRef]
32. Khraif, R.; Salam, A.A.; Elsegaey, I.; Ajumah, A. Dynamics of contraceptive use: A study of King Saud University women staff, Riyadh. *Middle East Fertil. Soc. J.* **2017**, *22*, 18–26. [CrossRef]
33. Nketiah-Amponsah, E.; Arthur, E.; Abuosi, A. Correlates of contraceptive use among Ghanaian women of reproductive age (15–49 years). *Afr. J. Reprod. Health* **2012**, *16*, 154–169.
34. Clements, S.; Madise, N. Who is being served least by family planning providers? A study of modern contraceptive use in Ghana, Tanzania and Zimbabwe. *Afr. J. Reprod. Health* **2004**, *8*, 124–136. [CrossRef] [PubMed]
35. Kayembe, P.K.; Fatuma, A.B.; Mapatano, M.A.; Mambu, T. Prevalence and determinants of the use of modern contraceptive methods in Kinshasa, Democratic Republic of Congo. *Contraception* **2006**, *74*, 400–406. [CrossRef] [PubMed]
36. Achana, F.S.; Jackson, E.F.; Welaga, P.; Awine, T.; Asuo-Mante, E.; Oduro, A.; Awoonor-William, J.K.; Phillips, J.F. Spatial and socio-demographic determinants of contraceptive use in the Upper East region of Ghana. *Reprod. Health* **2015**, *12*, 29. [CrossRef] [PubMed]
37. Cleland, J.; Phillips, J.F.; Amin, S.; Kamal, G.M. *The Determinants of Reproductive Change in Bangladesh*; The World Bank: Washington, DC, USA, 1994.
38. Stephenson, R.; Baschieri, A.; Clements, S.; Hennink, M.; Madise, N. Contextual influences on modern contraceptive use in sub-Saharan Africa. *Am. J. Public Health* **2007**, *97*, 1233–1240. [CrossRef] [PubMed]
39. Gayawan, E.; Arogundade, E.D.; Adebayo, S.B. Possible determinants and spatial patterns of anaemia among young children in Nigeria: A Bayesian semi-parametric modelling. *Int. Health* **2014**, *6*, 35–45. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Review

The Epidemiology of Suicide in Young Men in Greenland: A Systematic Review

Hannah Sargeant [1], Rebecca Forsyth [1] and Alexandra Pitman [1,2,*]

1 UCL Division of Psychiatry, London W1W 7NF, UK; hannah.sargeant.11@ucl.ac.uk (H.S.); rebecca.forsyth.14@ucl.ac.uk (R.F.)
2 Camden & Islington NHS Foundation Trust, London NW1 0PE, UK
* Correspondence: a.pitman@ucl.ac.uk; Tel.: +44-207-679-9038

Received: 30 September 2018; Accepted: 27 October 2018; Published: 1 November 2018

Abstract: Suicide is the leading cause of death among young men aged 15–29 in Greenland, but few epidemiological studies have described this problem. We aimed to summarise descriptive epidemiological studies of suicide in young men in Greenland compared with other demographic groups in Denmark and Greenland to inform future suicide prevention strategy. We searched PubMed, PsycINFO, and Embase using an agreed search strategy to identify English-language papers describing suicide epidemiology in Greenlandic men aged 15–29. We followed PRISMA guidelines in screening and appraising eligible publications. Eight articles fulfilled inclusion criteria of 64 meeting search criteria. Findings covering 1970–2011 supported a dramatic rise in suicide rates in Greenlandic men aged 15–24 from 1976, who remained the highest-ranking demographic group over 1976–2011 compared with men and women of all age groups in Denmark and Greenland. Highest rates recorded were almost 600 per 100,000 per year in men aged approximately 20–23 over 1977–1986. No studies described suicide epidemiology after 2011, and no studies described risk factors for suicide in young men. Given the very high suicide rates recorded for young men over 1976–2011, such studies will be essential for informing the development and evaluation of appropriate preventive interventions.

Keywords: suicide; premature mortality; young men; Greenland; Denmark

1. Introduction

Internationally suicide is the second leading cause of mortality in young men after accidental deaths [1]. Mortality indicators such as potential years of life lost (PYLL) demonstrate the economic and social cost of suicide in young men, which has become a serious public health problem over the last 70 years [1]. Despite widespread international recognition of the problem [1], and growing media concern, research is lacking on the mediators of suicide risk in young men and the interventions to mitigate them. International studies suggest that individual-level risk factors for suicide in men include psychiatric illness, substance misuse, lower socioeconomic status, rural residence, and single marital status [1]. However, there is a clear need for epidemiological studies of young men in specific regions, to understand the local sociocultural influences on their suicide risk, and the development of appropriate responses. The suicide rate in Arctic communities is considered a major public health concern [2], particularly in adolescents and young men [3]. Indigenous communities in the circumpolar north face a range of geocultural and economic hardships, including the challenges of inadequate housing and access to health care, in the context of global forces eroding local traditions [2]. Greenland has attracted particular media concern over suicide rates in young men, but few epidemiological studies have described this problem. A large (2,150,000 km^2) and remote island near the continent of North America, Greenland is a former colony of Denmark, largely populated by Inuit people [4]. According to suicide data published by Statistics Greenland, suicide accounts for 8% of total deaths in Greenland and is the leading cause of death among young men aged 15–29 [5]. However, due to its territorial status, Greenland's suicide rates are subsumed within those for Denmark, such that international rankings mask the problem. Broad comparison of the population suicide rates published by Statistics Greenland for 2011 (83 per 100,000) [5] greatly exceed those published by the World Health Organisation (WHO) for Guyana, the country with the highest population suicide rates internationally that year (32.5 per 100,000) [6], or for the 2015 global average of 10.7 per 100,000 [7]. International comparisons of suicide rates in young men identify particularly high rates in Eastern Europe and Japan, suggesting that suicide risk is higher for young men in countries undergoing transition or rapid social change [8]. Societies in transition experience changes in cultural norms, family cohesion, economic pressures, substance use, and migration patterns. The complex interaction of these variables requires investigation at the local level.

Greenland proposed a Greenlandic suicide prevention strategy in 2004, identifying young men as a high risk group, but acknowledging a lack of research into the aetiology of suicide in this group [9]. Although suicide prevention was included in the 2007–2012 Greenlandic public health programme, it was omitted in the 2013–2019 version [10]. However, Greenland participates in a US-led 2015 Arctic Council project, RISING SUN (Reducing the Incidence of Suicide in Indigenous Groups: Strengths United through Networks) to share expertise in suicide prevention in Arctic communities [2]. The success of such policy efforts relies in part on having a clear understanding of the problem epidemiologically. Suicide data collection in Greenland has a broken history, and policy-makers lack a clear picture of recent trends and high-risk groups. The earliest systematic suicide data collection in Greenland was by one physician from 1891–1930 [11], but there is then a gap until 1951 when annual reports by the country's Chief Medical Officer commenced [12]. Whilst actual numbers of suicides are published by Statistics Greenland for 1990–2013 [5], comparison of patterns with other Arctic areas is difficult without clear presentation of rates, both temporally and by age group. To inform future Greenlandic suicide prevention strategies and highlight gaps in evidence, it is important to understand how suicide rates in young men in Greenland compare with those for young men in Denmark, and with men in other age groups in Greenland.

We aimed to conduct a systematic review of research studies describing the epidemiology of suicide in young men (aged 15–29) in Greenland compared with young men in Denmark, and Greenlandic men in other age groups, using international evidence published up until 2018. In synthesising these findings we aimed to test the hypothesis that suicide rates in young men in Greenland are greater than those for these other groups. Our objectives were to describe historic and recent temporal trends in suicide rates in young men, comparing rates with those for other age and gender groups within Greenland and with other young men in Denmark, and identifying specific risk factors for suicide in young men. In conducting our review we identified a dramatic rise in suicide rates in Greenlandic men aged 15–24 from 1976, such that young men supplanted older men as the highest-risk demographic group for suicide over the period 1976–2011. Our search criteria did not identify any studies describing suicide epidemiology in young men from 2011 onwards, nor specific risk factors for suicide at any point, so recent patterns remain unclear.

2. Materials and Methods

2.1. Search Strategy

In developing our search criteria we decided to exclude research on self-harm due to wide international variations in definitions based on degree of suicidal intent [13], although acknowledging that self-harm is a key risk factor for suicide in men in high-income countries [14]. We chose the age range 15–29 to cover WHO definitions of youth (aged 15–24 years), teenagers (aged 15–19 years), and young adults (aged 20–24 years) [1], as well as the upper limit of under 30 used in the Greenlandic suicide prevention strategy [9]. We registered our review protocol on PROSPERO register of systematic reviews, and used the PRISMA checklist to guide the design and reporting of our study [15].

We searched the online database, PubMed, using the MeSH terms "Suicide" and "Greenland" and the equivalent keywords Suicid* and Greenland, to give the final search criteria (("Suicide" [Mesh]) OR Suicid*) AND ("Greenland" [Mesh] OR Greenland). We used no date restrictions on publication period. The search was repeated on Embase, and with slight variation on PyscINFO as the MeSH term "Greenland" was not available. No restrictions were applied to age-range or gender, as the predicted relatively limited number of articles on this topic made it feasible to apply age and gender exclusions post-search. We conducted the search on 11 October 2017 and screened titles and abstracts for eligibility, followed by full text review. We conducted secondary searching of the reference lists of identified articles, and of the publications lists of identified authors, and emailed international experts in the field to identify additional references. We repeated the search in PubMed on 15 August 2018 to check for any recent studies.

2.2. Inclusion and Exclusion Criteria

We included articles:

- reporting primary quantitative research involving any length of follow-up
- with titles/abstracts mentioning suicide, fatal or non-fatal suicide attempts, or suicidal ideation
- with titles or abstracts mentioning the Greenlandic population or Inuit population
- specifying age-ranges covering men aged 15–29
- specific to the Greenlandic population (those born in Greenland with at least one Greenland-born parent)
- published in English

We excluded articles:

- specific to Inuit populations in other countries
- presenting data on the age-range 15–29 but not specific to men
- focusing on non-fatal suicide attempt or suicidal ideation but not suicide

2.3. Study Selection

Two authors (AP and HS) independently screened all citations identified in the search for eligibility, comparing included/excluded articles for agreement. For those identified for full text review, two authors (AP and HS/RF) reviewed each manuscript independently to rate eligibility. The third independent reviewer resolved any discrepancies arising during title, abstract or full-text review.

2.4. Data Extraction

We developed a data extraction form based on STROBE criteria for observational studies [16]. For each article the following data were extracted: design, objectives, setting, participants, outcomes, summary measures relevant to research question (e.g., absolute numbers of suicides, suicide rates, standardised mortality ratios, other risk estimates), bias and limitations (as a marker for quality), and interpretation. For each paper included in the study, two authors (AP and HS/RF) independently reviewed each paper, rating quality (including an assessment of study and outcome bias) and summarising findings, including all summary measures presented. Each reviewer completed a data extraction form independently, and compared findings, updating entries in a master data extraction form. Again, a third independent reviewer resolved any discrepancies. As a team we discussed the potential for publication bias, and the potential for selective reporting within studies.

Due to the mixture of summary measures presented, the differing age-groupings, and the discontinuous time periods covered, we used a narrative approach to synthesise temporal findings, taking into account study quality. We used the PRISMA statement to structure findings [17].

3. Results

3.1. Studies Identified

Eight articles fulfilled inclusion criteria, of 64 meeting search criteria (Figure 1). Included studies were published between 1979 and 2015, with characteristics summarised in Table 1. Studies reported suicide data specific to young men over 1970–2011. There was great heterogeneity of time periods studied, ranging from rates for single years, to an aggregated period of 41 years (1970–2011) [18]. There was also great heterogeneity of measures presented (absolute numbers; suicide rates per 100,000 inhabitants; suicide rates per 100,000 person-years), prohibiting meta-analysis. Whilst one study used a case-control design [19], the remaining seven presented cross-sectional data derived from national suicide registers (based on death certificates and police reports), or from these primary sources themselves. Six of those seven aggregated data for the periods of interest, but one presented data for seven specific years over 1970–1995 [20], although but only for West Greenland where suicide rates are lower than for East Greenland or the capital Nuuk [21]. Seven studies reported suicide data on Greenlandic individuals of all ages, from which we were able to extract selected findings relating to men aged 15–29. One study concentrated solely on young people aged 15–30 [22]. Two only reported data only from West Greenland [20] or Nuuk [19].

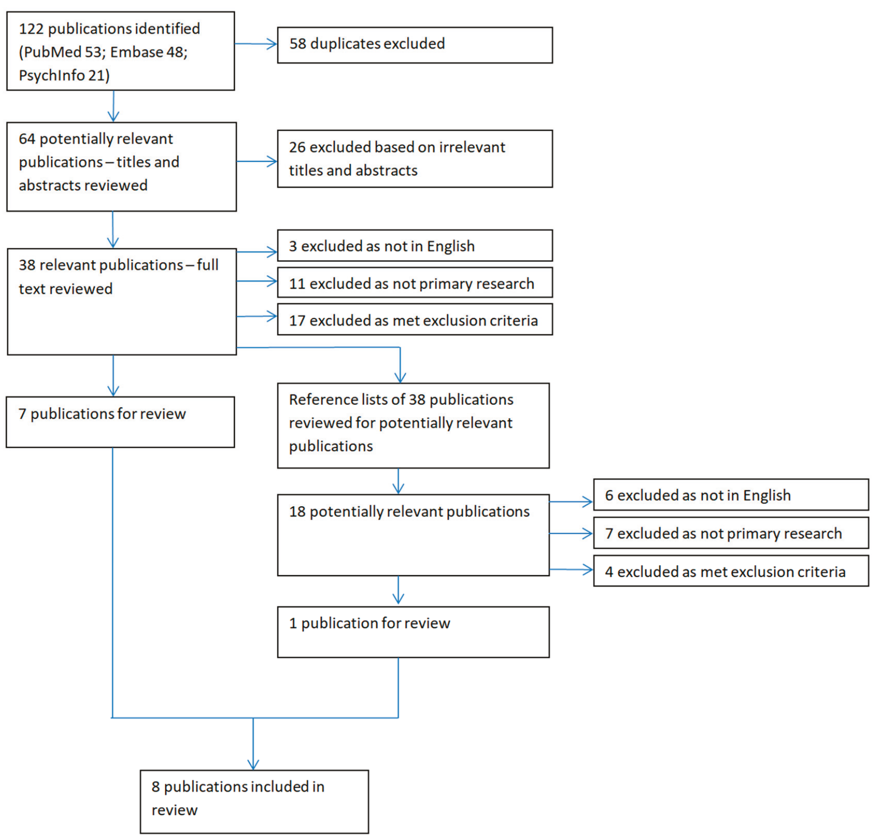


Figure 1. Flow diagram summarising the study selection process.

Table 1. Studies included in systematic review (in date order of publication).

Author and Title	Design, Setting and Objectives	Participants	Variables	Results (Those Relevant to Research Question)	Bias and Limitations	Interpretation
Grove and Lynge, 1979 Suicide and attempted suicide in Greenland. A controlled study in Nuuk [19].	Case-control study. Capital of Greenland, Nuuk, in 1972 and 1973. Aimed to identify risk factors for suicide and suicide attempt in cases of fatal and non-fatal suicide attempt compared to general population never suicidal controls.	**Cases**: all Greenlanders in Nuuk who died by suicide or attempted suicide in 1972 and 1973 (n = 1576 males + 1697 females), with attempted suicide defined as that requiring hospital admission. **Controls**: Nuuk residents admitted to hospital for a somatic disease or pregnancy, matched by gender and age group for each case, but with no history of suicide attempt.	Information on suicide cases collected from available records "coming to the attention of any authority in the district of Nuuk." i.e., death certificates, and police reports. Information on suicide attempt cases was additionally collected from hospital files, police reports, and crime registers to capture unvalidated measures of family composition, childhood, atmosphere in parental home, alcohol consumption, criminal records, exposure to attempted or completed suicide, marital status, inter-racial marriage, psychiatric history, obstetric history, and method of suicide attempt. Age groupings used: 15–19; 20–24; 25–39; ≥40.	Measures presented for young men: absolute numbers of suicide deaths. Findings: 12 suicides recorded in Nuuk during 1972 and 1973: 10 were in men. Age-specific data identified 4 suicides in men aged 15–24 over this period (compared with 2 for women), and 16 suicide attempts in men of the same age. Suicide rates for 4 age-groups were estimated for both genders combined (300 cases of suicide per 100,000 inhabitants in men and women aged 15–19; 173 cases per 100,000 inhabitants in men and women aged 20–24). Suicide rates for men and women aged 25–39 were 226 per 100,000, whilst those for men and women of 40 and over were 54. Risk factors were not separated out by age group or gender. Ratios of suicide attempts to suicide deaths were not presented, but on direct calculation were: 5:3 in men aged 15–19 (compared with 10:1 in similar aged women) 11:1 in men aged 20–24 (compared with 12:1 in similar aged women). 10:5 in men aged 25–39 (compared with 11:0 in similar aged women).	Outcomes were separated out for men and women only in relation to absolute numbers of suicide deaths. Unclear whether method of identifying suicide cases in Nuuk in 1972–1973 was comprehensive. Presented age-specific suicide rates per 100,000 inhabitants (of Nuuk) per year, using census mid-estimates of the population denominator in each age band. Findings from the population of Nuuk may not be generalizable to the whole Greenlandic the population due to urban-rural differences. Estimates of the general population denominator for 1973 were based on mid-point interpolations between 1970 and 1976 census data. These were therefore unreliable as mid-point population estimates would fail to take into account changing migration and fertility patterns.	Little can be concluded about suicide epidemiology in young men in the capital of Greenland as absolute numbers of suicides in men of different age groups were too low to attempt meaningful comparisons. Estimates of age-specific suicide rates were not specific to men, and used an unreliable estimate of the denominator. Direct calculation of the ratios of suicide attempts to suicide deaths indicate that the ratio increases from 5:3 for men aged 15–19 to 11:1 in men aged 20–24 and then falls to 10:5 for men aged 25–39.
Lynge, 1985 Suicide in Greenland [12]	Aggregated data from annual cross-sectional studies of suicide cases. Whole of Greenland during the period 1974–1984. To describe suicide rates in Greenland.	**Cases**: All those aged 15 and above who were born in Greenland and died by suicide in 1974–1984, as identified from death certificates and police reports (n = 318).	Data on age, gender and place of residence of those dying by suicide, based on police reports and death certificates, as well as unvalidated measures of known motives and whether alcohol was involved in the suicide. Age groupings used: 15–19; 20–24; 25–39; ≥40	Measures presented for young men: absolute numbers of suicide deaths; suicide rates per 100,000 inhabitants per year. Findings: 318 total suicides in men and women over this period. 256 were in men, with the highest absolute numbers of suicides (n = 97) in men in the age group 20–24. Age and gender-specific suicide rates per 100,000 inhabitants per year were highest in men age 20–24 (387 per 100,000). The suicide rate in men aged 15–19 (151 per 100,000) was higher than that in men aged 40 and over (73 per 100,000) but lower than that in men aged 20–24 (162 per 100,000). The suicide rates for women were lower than for men in all age groups (44 per 100,000 in women aged 15–19; 94 per 100,000 in women aged 20–24).	Author acknowledges that due to non-systematic collection of death certificates, some cases of suicide may have been missed. Data for an aggregated period obscured year on year changes. Presented crude suicide rates for 100,000 inhabitants per year.	For the period 1974–1984 suicide rates in young men aged 20–24 were the highest of all age and gender groups in Greenland at 387 per 100,000 inhabitants per year, compared with 151 per 100,000 for men aged 15–19, 162 per 100,000 for men aged 25–39, and 73 per 100,000 in men over 40. The gender gap in suicide rates over this period was greatest in the age group 20–24 at 387 *versus* 94 per 100,000. Year on year rates were not presented so no information was provided on temporal trends.

Table 1. *Cont.*

Author and Title	Design, Setting and Objectives	Participants	Variables	Results (Those Relevant to Research Question)	Bias and Limitations	Interpretation
Thorslund, 1990 Inuit Suicides in Greenland [23]	Aggregated data from annual cross-sectional studies of suicide cases. Whole of Greenland from 1977–1986. To describe the epidemiology of suicide among Greenlandic Inuit.	**Cases**: All recorded suicides in Greenland from 1977–1986 among Greenland-born individuals (n = 403), using death certificates and police reports.	Data on age, gender, occupation, marital status, parental status, psychiatric history, substance misuse history, and circumstances of the death, history of suicidal behaviour, derived from public files, death certificates, municipal welfare office records, and police reports. Age groupings used: histogram presented data for a continuous measure of age from 10 to 60	Measures presented for young men: average suicide rates per 100,000 population per year. Findings: 403 suicides analysed over a 10 year period, of which 326 were of men (of any age). Visual plots of average suicide rates per 100,000 population per year showed that suicide risk increased sharply from the ages of approximately 15–17, peaking at approximately age 20, falling from approximately age 23, and with further peaks in the mid-30s and the 50s. The great disparity between male and female suicide rates from the age of about 17 diminished by approximately age 27. Data for risk factors, including occupation, was not separated out by gender or age group.	Little information was presented regarding the reliability of routine data. Data for an aggregated period obscured year on year changes. Presented crude suicide rates for 100,000 inhabitants, as an average per year using data over a 10 year period 1977–1986. Data for risk factors was not presented by gender or age group.	Over the 10 year period 1977–1986, men aged approximately 20 to 23 had the highest suicide rates in the Greenlandic population, greatly exceeding those for women of the same age.
Thorslund, 1991Suicide Among Inuit Youth in Greenland 1977–1986 [22]	Aggregated data from annual cross-sectional studies of suicide cases. Whole of Greenland from 1977–1986. To describe risk factors for suicide among youths aged 15–30 in Greenland.	**Cases**: All suicides in Greenland from 1977–1986 of youths aged 15–30 (n = 287), using death certificates and police reports. **Controls**: n = 320 randomly selected young people living in Greenland, sampled via postal questionnaire in 1988.	Data collected on place of residence (town/village), occupation, presence of alcohol in blood, and personal circumstances prior to death, based on public files (police reports, death certificates and social service departments), Public Databases (Public Health Department and Central Bureau of Statistics) and from questionnaires (for controls). Age groupings used: 15–30.	Measures presented for young men: crude average number of suicides per year per 100,000 inhabitants (for both genders); proportions of suicide deaths by gender and occupational group (aged 15–30); proportions of suicide deaths by gender and recent psychosocial stressor (all ages). Findings: In a sub-sample of youth dying by suicide aged 15–30 in 1982–1986 for which occupational data were presented (n = 184), the majority were men (n = 144; 78%). Of this sub-sample, 26% of men were hunters/fishermen *versus* 0% of women (and 15% of male controls), 21% were unskilled workers vs 13% of women (and 37% of male controls), 24% were unemployed vs 38% of women (and 3% of male controls), and 5% were white collar workers vs 10% of women (and 13% of male controls). No statistical tests were presented for these comparisons. For men and women dying by suicide aged 15–30 from 1977–1986, in over 30% of cases there were reports of personal problems with a spouse prior to the suicide, with the text indicating that this was more common in women (but not whether this was statistically significant). The text also indicated that in 20% of male suicides there was a history of rejection from parents or friends "or other shameful situations" prior to the suicide.	Data for an aggregated period obscured year on year changes. Little information presented regarding the reliability of routine data. Very little information provided on the characteristics of the control group used in the comparison of proportions of men and women dying by suicide by occupational group (derived from a survey of Greenlandic residents in 1988). No statistical tests presented for comparisons between men and women aged 15–30, or between cases and controls. Data on risk factors were not presented by gender. Presented crude average number of suicides per year per 100,000 inhabitants (for both genders).	From 1982–1986, 78% of suicides among youth aged 15–30 were in men. These men primarily worked in traditional hunting/fishing jobs, unskilled jobs, or were unemployed. On the basis of cases of suicide in those aged 15–30 from 1977–1986, rejection by parents of friends or "other shameful situations" were implicated in 20% of male suicide cases aged 15–30.

Table 1. *Cont.*

Author and Title	Design, Setting and Objectives	Participants	Variables	Results (Those Relevant to Research Question)	Bias and Limitations	Interpretation
Leineweber & Arensman 2003 Culture Change and Mental Health: The Epidemiology of Suicide in Greenland [24]	Aggregated data from annual cross-sectional studies of suicide cases. Whole of Greenland from 1972–1995, with age- and gender-specific data only presented for the years 1990–1995. To describe suicide rates in Greenland.	**Cases**: All suicides in Greenland from 1972–1995 of Greenland-born individuals, using the register of causes of death based on death certificates certified by a physician using ICD code diagnoses.	Data collected on gender, age and place of residence were drawn from a computerized register on causes of death for persons born in Greenland. Population suicide rates were age-standardised by direct standardisation. Age groupings used: histograms presented plots for ages 10–15; 15–19; 20–24; 25–29; 30–39; 40–49; 50–59; 60+ (or 15–19; 20–24; 25–29; ≥30)	Measures presented for young men: Visual plots of suicide rates per 100,000 population by sex and age-group. Findings: Visual plotting of suicide rates per 100,000 population by age-group and sex in Greenland in the period 1990–1995 showed that rates in men aged 15–24 were the highest of all age-groups for either gender; approximately 460 per 100,000 population for those aged 15–19, declining with age, but remaining above 400 per 100,000 for those aged 20–24, and approximately 300 per 100,000 for men aged 25–29. Rates for men were higher than for women in all age groups. Risk factors for suicide were not separated out by age-group or gender.	Data for an aggregated period obscured year on year changes. Risk factors for suicide were not separated out by age-group or gender.	From 1990–1995 suicide rates in men aged 15–19 and 20–24 were the two highest ranking age groups for all age-groups in either gender. Rates for men aged 15–24 were approximately 460 per 100,000, and declined with age, remaining higher for men than women in all age groups.
Bjorksten, Bjerregaard et al., (2005) Suicides in the midnight sun—A study of seasonality in suicides in West Greenland [20]	Suicide data from annual cross-sectional studies of suicide cases only for West Greenland, 1968–1995, with rates for 7 specific years reported. To investigate whether there is evidence for seasonality of suicide rates in West Greenland.	**Cases**: Suicides of people of any age living in towns and settlements in West Greenland, as recorded in the register of causes of death in Greenland, and population registers from the National Institute of Public Health in Copenhagen were analysed (n = 833).	Data collected on age, gender, suicide date, country of birth (Greenland/Denmark), residence (town/settlement), latitude, and whether alcohol contributed to their death, based computerised registers on causes of death in Greenland and population registers from the National Institute of Public Health in Copenhagen. Age groupings used: 0–14; 15–24; 25–34; 35–59; ≥60.	Measures presented for young men: suicide rates per 100,000 person-years by age-group; seasonality of suicides by gender (men of all ages) and age group (all men and women ≤ 24 vs. >24). Findings: Of 684 total suicides in men over the period 1968 to 1995, the median age of cases was 25 and the age range was 11 to 84 years. When considering the specific years 1970, 1976, 1982, 1987, 1990, 1993, and 1995 (for which detailed population data were available, allowing calculation of suicide rates per 100,000 person-years), results showed that men aged 15–24 had the highest suicide rates between 1976 and 1995, peaking at 577 per 100,000 person-years in 1990. For the period 1970 to 1995, the longitudinal picture for male suicide rates was one of great change. In 1970, all 122 suicide cases had been men aged 35–59 years of age, with the overall suicide rate for men of all ages at 22 per 100,000 inhabitants per year. From 1976 to 1995, the overall suicide rate in men climbed from 80 per 100,000 in 1976 to a maximum value of 214 per 100,000 in 1990. For young men aged 15–24, suicide rates climbed from 0 in 1970 to 249 per 100,000 in 1976 to a maximum of 577 per 100,000 in 1990. This age group had the highest suicide rates of all male age groups during 1976–1995, plateauing at 238 per 100,000 in 1995. Men of 25–34 years of age also had high suicide rates from 1976 on, peaking at 297 per 100,000 in 1990. In 1995 suicide rates for men aged 15–24 remained the highest at 238 per 100,000, but disparities with those for men aged 25–34 were less at 219 per 100,000 person-years. Analysis of seasonality of suicides was examined for men (all ages), finding a statistically significant midsummer peak, and for men and women combined aged ≤ 24, finding no significant seasonality. Risk factors for suicide were not broken down by age-group or gender.	Estimates of suicide rates by age group for specific years over the period 1968–1995 may not be generalisable to East Greenland or Nuuk, where suicide rates tended to be higher in the 1970s and 1980s. Suicide rates presented as actual numbers of suicides per year for the period 1968–1995, and as crude rates per 100,000 person-years for the specific years 1970, 1976, 1982, 1987, 1990, 1993, and 1995. The finding of no significant seasonality by age may have been due to low numbers	No observations can be made about seasonality of suicide in young men specifically, but there was a significant midsummer peak for men of all ages. For young people, no seasonality was observed in men and women (combined) aged 24 and under for the period 1968–1995. In the whole population, significant seasonality was observed, with a peak in late June and lowest rates in the darkest months (December to February). From 1976–1995 suicide rates were highest in men aged 15–24, peaking at 577 per 100,000 in 1990. In 1995 they remained the highest for all male age groups at 238 per 100,000 but with less of a disparity with rates for men aged 25–34.

Table 1. *Cont.*

Author and Title	Design, Setting and Objectives	Participants	Variables	Results (Those Relevant to Research Question)	Bias and Limitations	Interpretation
Bjerregaard and Lynge (2006) Suicide—A Challenge in Modern Greenland [21]	Aggregated data from annual cross-sectional studies of suicide cases. Whole of Greenland from 1968–1999. To describe suicide rates in young men in Greenland in 1968–1999, compared with those in other demographic groups. Study also presented linked cross-sectional population-based survey data from 2 population surveys in Greenland (in 1993–1994 and 1999–2001) describing past year prevalence of suicidal thoughts (not meeting search criteria for the current study but brief details given here).	**Cases**: All suicides in Greenland from 1968–1999 of Greenland-born individuals (n = 1203), using the register of causes of death based on death certificates certified by a physician using diagnoses based on ICD-8 (1968-1993) or ICD-10 (1994–1999).	Data collected on age, gender, and place of residence. Age groupings used: histogram presented data for a continuous measure of age from 0 to approximately 85, with points specified for 0–4; 10–14; 20–24; 30–24; 40–44; 50–54; 60–64; 70	Measures presented for young men: visual plots of age-specific suicide rates per 100,000 person-years; visual plots of youth suicide rates (aged 15–29) for men per 100,000 births in each cohort. Findings: Overall for the period 1968–1999 and all age groups, suicide rates were 4.3 times higher in men than women. Visual plots of age-specific suicide rates for the aggregated period 1990–1999 showed suicide rates to be considerably higher in men aged 15–24 in Greenland (approximately 470 per 100,000 person-years) than for men of a similar age in Denmark and for women of a similar age in Greenland and in Denmark. This disparity continued up until approximately the age 45, when rates in all four groups converged. Youth suicide rates (aged 15–29) for men in the 1950 birth cohort were plotted as approximately 10 per 100,000 births in that cohort, rising to approximately 60 per 100,000 births in the 1978 birth cohort. Although gender-specific data were not presented, visual plots showed that youth suicide rates (aged 15–29) in East Greenland continued to rise from 1975–1979 (approximately 200 per 100,000 person years) to 1995–1999 (approximately 800 per 100,000 person years), in comparison to West Greenland and the capital Nuuk where they plateaued or fell slightly (to between 100 and 200 per 100,000 person years in 1995–1999). Other risk factors for suicide were not broken down by age-group or gender. Additional analysis of data from population surveys called out in 1993–1994 and 1999–2001 showed that men aged 18–24 were significantly less likely that same-aged women to report lifetime suicidal thoughts (19% *versus* 33%; $p = 0.03$). At age groups above 25 this gender difference was not statistically significant.	Figures in text did not match those in graphical presentations. Data for an aggregated period obscured year on year changes. Suicide rates presented as crude rates per 100,000 person-years for blocks of 5 years, or as crude rates per 100,000 births in a birth cohort.	Aggregated data for 1990–1999 show that suicide rates in young men aged 15–24 were much higher than those for men in other age groups and for men and women of the same age in Denmark, peaking at 450–500 per 100,000 person-years. This pattern differed from that in Denmark, where suicide rates rose across the age groups for men and women, and where a much lower ratio of male:female suicides was reported (1.8 in Denmark *versus* 4.3 in Greenland). There was evidence that male youth suicide rates (in the age group 15–29) were higher in men born in later cohorts (downstream of sociocultural change), as evidenced in data from cohorts born from 1950 to 1978. Regional variations in youth suicide were not broken down by gender, but suggested that young people in remote areas of East Greenland had higher suicide rates, which continued to rise, whilst youth suicide rates had peaked in the capital, Nuuk, in the early 1980s, and plateaued in West Greenland throughout the 1980s and 1990s. Population survey data from 1993–1994 and 1999–2001 suggested that although suicide rates in young men aged 15–24 were considerably higher than those for women of that age in Greenland, women were significantly more likely to report lifetime suicidal thoughts.

Table 1. *Cont.*

Author and Title	Design, Setting and Objectives	Participants	Variables	Results (Those Relevant to Research Question)	Bias and Limitations	Interpretation
Bjerregaard & Larson (2015) Time trend by region of suicides and suicidal thoughts among Greenland Inuit [18]	Aggregated data from annual cross-sectional studies of suicide cases, averaging out rates over periods of between 5 and 30 years Study also presented linked cross-sectional population-based survey data from 2 population surveys in Greenland (in 1993–1994 and 2005–2010) describing past year prevalence of suicidal thoughts (not meeting search criteria for the current study but brief details given here). Whole of Greenland from 1970–2011.To describe time trends in suicide rates (and past year prevalence of suicidal thoughts in 1993–1994 and 2005–2010) in Greenland from 1970 to 2011	**Cases**: All suicides in Greenlandic residents from 1901–2011 (n = 1678), based on routine registry data from the Greenland registry of causes of death. General population sample of Greenlandic residents sampled 1993–1994 and 2005–2010 using the same instrument, and overlapping geographical sampling frames, to collect data on self-reported past year prevalence of suicidal thoughts, with linkage of individuals in cross-sectional surveys to subsequent suicides.	Data collected on age, gender, and region of residence for all suicide cases from 1901–2011. Survey data collected on age, gender, and past year prevalence of suicidal thoughts in two cross-sectional samples. Age groupings used: 10–14; 15–19; 20–24; 25–29; 30–34; 35–44; 45–54; ≥55	Measures presented for young men: suicide rates per 100,000 person-years by age-group (tabulated and visual plots) Findings: Suicide rates for men aged 20–24 were the highest of all male age groups for the period 2000–2011 (at 426 per 100,000 person-years over this whole 12 year period), compared with 297 per 100,000 person years for men aged 15–19 and 251 for men aged 25–29. Suicide rates for men aged 55 and above were the lowest of all male age groups (at 74 per 100,000 person years). Visual plots of suicide rates for men and women by age group for the period 1970 to 2011 also show that rates were highest for men aged 20–24 (at approximately 400 per 100,000 person years). NB: We clarified with the first author that the suicide data presented in his table II are from the registry of causes of death and cover 2000–2011 only, to be comparable to the survey data on suicidal thoughts. He confirmed that the suicide data presented in Figure 1 are also derived from the registry of causes of death but cover the whole time range of the registry, which at that time was 1970–2011. There is hence some overlap between the data from these two sources, and the age pattern was similar in the shorter period and over the whole period. Age patterning of high risk groups for past-year suicidal thoughts mirrored that for suicide rates in men, with the highest rate in men aged 20–24 at 136 per 1000 participants. The proportion of women with past-year suicidal thoughts was greatest for women aged 15–19 (188 per 1000 participants) and fell with increasing age. Age- and gender-specific suicide rates were not presented by region of residence, but the text indicated that male and female patterns were similar to the overall pattern. This pattern, since the 1980s, has been for suicide rates for men, women, and overall to have been significantly higher in East and North Greenland than in West Greenland or the capital, Nuuk. Association between past year suicidal ideation and subsequent suicide was not presented by gender or age group.	Data for an aggregated period obscured year on year changes. Suicide rates presented as crude rates per 100,000 person-years for blocks of between 5 and 30 years. Source and quality of data from 1970–2011 not qualified. Due to small absolute numbers of completed suicides, there was insufficient statistical power to stratify analyses on several levels simultaneously e.g., sex and region. Unclear whether survey instrument validated. [Respondents in the two surveys, and within the same sampling period, could have been sampled twice. No discussion of response rate, recall bias, or social desirability bias.]	Suicide rates for men aged 20–24 were the highest of all age groups for both genders over the 41 year period 1970 to 2011 (at approximately 400 per 100,000 person years), and when considered over the shorter 12 year period of 2000–2011 (at 426 per 100,000 person-years). Age groups at highest risk of suicide match those at highest risk of suicidal thoughts for men.

3.2. Study Quality

The quality of included studies was partly a function of the quality of routine data available, which improved over the years studied. The measures presented also became more meaningful over time; from absolute numbers to rates per 100,000 inhabitants/person-years. Both of these improvements are demonstrated in the chronological presentation of papers in Table 1. For example an early study presented absolute numbers of suicides by age group and gender for 1972 and 1973, but no suicide rates specific to men (or indeed men in specific age groups) [19]. In later studies crude suicide rates were represented as per 100,000 inhabitants [12,22,23] or per 100,000 person-years [18,20,21]. None of the studies used join-point regression models to calculate annual percentage change statistics. Many studies compared suicide rates in different age or gender groups (or between cases and controls) without providing test statistics. Studies investigating risk factors for suicide aggregated genders or age groups, probably for reasons of power, or lacked test statistics.

3.3. Risk of Bias across Studies

The dramatic rise in suicide rates in the early 1970s, increasing from zero in young men aged 15–24 in West Greenland in 1970 to 249 per 100,000 in 1976 [20], may have reflected early under-reporting or misclassification bias. In 1970 all reported suicide cases in West Greenland were for men in the age group 35–59 years of age, but by 1976 suicides were recorded across all age groups above 15 [20]. At the population level, official Greenlandic figures for 1967–1971 indicated an average of 6.4 suicides per year (16.7 per 100,000 Greenlandic-born inhabitants), which had risen to 21.6 per year (53.9 per 100,000) for 1972–1976 [23]. Due to the high relative number of accidents and unidentified causes of death in Greenland, it is possible that a substantial number of suicides are misclassified as accidents [24]. Cultural reasons may underlie the differential recording by age group of suicide *versus* accidental death on the cause of death register, and temporal variations in such practices. The only evidence of selective reporting within studies was for the specific risk factors investigated, which may have reflected inductive bias.

3.4. Results of Individual Studies

3.4.1. Aggregated Data for Specific Periods

Over the aggregated period 1970–2011, the highest suicide rates were in men aged 20–24 (at around 410 per 100,000 person-years), with a clear excess of male suicides [18]. Suicide rates for men for the 41 year period 1970–2011 fell with advancing age group, to their lowest level in men aged 65–69 (at around 50 per 100,000 person-years), then increased slightly for men over 70 [18]. These figures are consistent with aggregated suicide data for a shorter 11 year period (1974–1984), showing that suicide rates in young men aged 20–24 were the highest of all age and gender groups in Greenland (at 387 per 100,000 inhabitants per year) compared with 151 per 100,000 for men aged 15–19, 162 per 100,000 for men aged 25–39, and 73 per 100,000 in men over 40 [12]. The gender gap in suicide rates over this period (1974–1984) was greatest in the age group 20–24 (387 for men *versus* 94 for women per 100,000 inhabitants) [12].

Similarly, aggregated suicide data for the shorter 10 year period 1977–1986 show a consistent pattern of highest suicide rates in Greenlandic men aged approximately 20–23, greatly exceeding those for women of the same age [23]. Visual plots presented for average suicide rates per 100,000 population per year showed that suicide risk increased sharply from the ages of approximately 15–17, peaking at almost 600 per 100,000 per year at approximately age 20–23, falling thereafter but with further peaks aged around 35 (approximately 300 per 100,000) and around 55 (approximately 250 per 100,000). A great disparity between male and female suicide rates was apparent from the age of approximately 17–27 [23].

By the 1990s, aggregated data for the narrower time periods 1990–1999 [21] and 1990–1995 [24] reveal men aged 15–24 to have been the highest risk group. A study presenting visual plots of age-specific suicide rates for the aggregated period 1990–1999 showed suicide rates to be highest in men aged 20–24 in Greenland (approximately 470 per 100,000 person years), closely followed by men aged 15–20 (approximately 430 per 100,000 person years); both considerably higher than those for men or women in other age groups in Greenland, or for men and women of any age in Denmark. This disparity continued up until approximately the age of 45, when rates in men and women in Greenland and Denmark started to converge at around 50 per 100,000 person-years [21].The pattern of suicide risk in Danish men and women over the same period was very different; rising slowly across the age groups but remaining below approximately 70 per 100,000 person-years [21]. In Denmark there was also a much lower ratio of male:female suicides of 1.8 compared to 4.3 for Greenland [21].

Aggregated data for the shortest span of 1990–1995 show that for this period men aged 15–19 had overtaken men aged 20–24 as the group at highest risk, these being the two highest ranking age groups for suicide rates in all age-groups in either gender [24]. Rates for men aged 15–19 during 1990–1995 were approximately 480 per 100,000, declining with age at approximately 440 per 100,999 for men aged 20–24, and 300 per 100,000 for men aged 25–29, and declining thereon. This compared with age-standardised population rates of nearly 110 per 100,000 population over that period. Higher rates for men than women applied in all age groups [24].

3.4.2. Temporal Trends

Only one study described temporal changes in suicide rates for young men, reporting rates for seven specific years over 1970–1995, but solely for West Greenland and using relatively wide age bands [20]. This recorded a zero suicide rate for young men aged 15–24 in 1970, and all reported suicide cases in men that year were within the age group 35–59 years [20]. By 1976 suicide rates in men aged 15–24 had risen dramatically to 249 per 100,000, becoming the group at highest risk, ranking above those for men aged 25–34 at 123 per 100,000 [20]. They remained the highest risk group at all remaining data points (1982, 1987, 1990, 1993, 1995), having peaked at 577 per 100,000 in 1990, and fallen to 238 per 100,000 in 1995 [20]. Throughout this period, men aged 25–34 also had high suicide rates, reaching a maximum of 297 per 100,000 in 1990 (compared with 577 per 100,000 for men aged 15–24). However, by 1995 the disparity between high suicide rates in men aged 15–24 and those aged 25–34 was less marked, at 238 *versus* 219 per 100,000 [20].

3.4.3. Period Effects

Gender-specific suicide data on cohorts born from 1948–1978 provide evidence that suicide rates in the age group 15–29 rose in successive cohorts born from 1952, matching the onset of major sociocultural change. For example suicide rates in men aged 15–29 born in 1952 were approximately 4 per 100,000 births in that cohort, rising to approximately 60 per 100,000 births in the 1978 birth cohort [21]. Whilst this is suggestive of a period effect, we lack comparative data describing other age groups in those cohorts [25].

3.4.4. Regional Variation

One study presented regional variations in youth suicide from 1970–1999, which were not broken down by gender [21]. However, this suggested that young people in remote areas of East Greenland had generally higher suicide rates, rising from approximately 200 per 100,000 person-years in 1970–1974, to approximately 800 per 100,000 person-years in 1995–1999. Meanwhile, youth suicide rates peaked in the capital, Nuuk, in 1980–1984 at above 300 per 100,000 person-years, and fell to between 100 and 200 per 100,000 person-years in 1995–1999. Suicide rates for young people in West Greenland were lowest among regions but rose throughout the 1970s and 1980s, overtaking those in Nuuk in 1985–1989 where they plateaued at approximately 200 per 100,000 person-years throughout the 1980s and 1990s. As the majority of these youth suicides will have been in men, only tentative conclusions can be drawn about these patterns of high suicide rates in rural areas applying to males aged 15–29 [21].

3.4.5. Risk Factors

We were unable to identify specific risk factors for suicide in young men in Greenland because analyses of suicide risk factors aggregated all age groups or both genders, or lacked formal statistical tests. One study investigating seasonality found a significant midsummer peak for men of all ages, but this was not specific to young people [20]. One study, described briefly here, provided limited findings on occupational groups and triggering factors [22]. This reported absolute numbers and proportions of suicides in young men by occupation and triggering factors, for the aggregated period 1982–1986 [22], but did not use rates or statistical tests (perhaps due to relatively low numbers). Only tentative inferences can be made from their finding that 78% of suicides among youth aged 15–30 were in men, and that these men primarily worked in traditional hunting/fishing jobs, unskilled jobs, or were unemployed [22]. Using witness statements from police reports, this study also found that rejection by friends or parents "and other shameful situations" were implicated in 20% of male suicide cases in the age group 15–30 but comparisons with other groups were not presented [22]. One other study found that for all male suicides, 16% of death certificates mentioned alcohol dependence/intoxication, but lacked data on age groups [20], so again little can be inferred about young men.

3.4.6. Suicide Methods

As data on suicide methods in the identified studies did not disaggregate data by age-group and gender we could not identify the methods used by young men in Greenland, or temporal trends in these methods. From 1968–1995, 93% of suicides in the whole population were violent, predominantly shooting or hanging, and this was more common in men than women (96% *versus* 81%; no test statistic provided) [20].

4. Discussion

4.1. Main Findings

Overall our findings support a pattern of consistently higher suicide rates in young Greenlandic men over the period 1976–2011, both when compared with men in other age groups, same-age Greenlandic women, and with men and women of all age groups in Denmark [21]. Aggregated data for periods over 1970–2011 suggests that men aged 20–24 were the highest risk group for suicide in Greenland, followed by men aged 15–19. Evidence from the narrowest period of reporting in 1990–1995 suggests that men aged 15–19 replaced men aged 20–24 as the group at highest risk around this point, at approximately 480 per 100,000 population [24]. This was also the point at which population rates stabilised, at approximately 110 per 100,000 population until the end of 1995 [24]. However, as the only study reporting rates for single years collapsed both age ranges [20] we cannot pinpoint if and when men aged 15–19 overtook men aged 20–24 as the highest ranking age group for suicide.

At the Greenlandic population level, age-standardised population suicide rates over the period 1972–1995 increased markedly from approximately 44 per 100,000 population in 1975 to approximately 110 per 100,000 in 1984–1989, stabilising at 110 per 100,000 from 1990–1995 [24]. This dramatic rise coincides with a time of high suicide rates in Greenlandic men aged 15–24, [12,18,20,21,23,24] suggesting that this age group accounted for the population-level transition. This apparent shift in suicide rates in Greenland from elderly to younger men matched the transition that occurred from 1950–1999 in some high-income countries [1], as suicide rates rose in middle-aged men (aged 35–45) and in young men (aged 15–25), largely supplanting older men as the group at highest risk of suicide [26]. The studies we identified, however, did not present a fine-grained picture of whether and how this age transition occurred. Variation in the years and age ranges reported also hamper direct comparisons of suicide rates across age groups in Greenland to those in other countries.

Geocultural factors are likely to be relevant in explaining suicide epidemiology in Greenland. Primarily covered by an ice sheet, it is sparsely populated by approximately 57,000 people, mainly Inuit people who started to migrate there from Canada in 2500 BC [4]. Its high latitude, ranging from 59° N to 84° N, accounts for extreme seasonal variations in daylight and darkness hours. The country has arctic climatic conditions; seasonal temperatures vary with latitude and distance from the coast, but average yearly temperatures remain below 10 °C [27]. Its remoteness kept it culturally insulated from its neighbours until the 18th century, but in 1953 its colonial status ended and it became an integrated constituency, governed by the Danish state and giving Greenlanders Danish citizenship [28]. This marked the start of a period of rapid sociocultural and socioeconomic changes under the modernising influence of Denmark. These included a change from subsistence hunting and fishing to a wage-earning economy, the migration of non-Inuit people to Greenland, and increased urbanisation, influencing changes in infrastructure and housing [29]. Genetic studies find that over 80% of Greenlanders have some European ancestry due to recent migration (approximating to 25% of their genome), inherited primarily from male Europeans [4]. In 1979, Greenlanders voted for Home Rule, which gave them the right to elect their own parliament, reducing the influence of Denmark. Greenland gained further autonomy in 2008, when a vote for self-governance was passed by a 75% majority and Home Rule was replaced. Greenlanders are now recognised as separate people under international law, and the native Inuit language of west Greenland has official status in place of Danish. The country has autonomous control over areas such as education, health and environment, whilst Denmark still governs justice affairs, national security, civil rights and financial sectors [28].

These geopolitical factors set important context. The dramatic increase in suicide rates in young men in Greenland in the 1970s coincided with a period of rapid modernisation and social change. However, we were unable to identify specific risk factors for suicide in young men as no studies reported these. We can only therefore hypothesise that specific aspects of modernisation increase risk. Superficially, evidence for regional variation in male suicide rates appears to undermine arguments for the effects of modernisation, given that rates are much higher for men of all ages in rural areas of East and North Greenland than in West Greenland, or the capital, Nuuk [18] and for young men in rural East Greenland from 1968–1999 [21]. However, the explanation may lie in disparities in socioeconomic status between men in urban and remote areas. The problem is that the individual-level suicide risk factors identified for young men in other countries (psychiatric disorder, substance misuse, occupational group, ethnicity, rural residence, lower socioeconomic status, single marital status) have not been specifically tested in young men in Greenland, nor have the population-level risk factors (unemployment, social deprivation, media influences) identified in high-income countries [1]. Other explanations for marked rises among young men at the start of this period include improved reporting during the 1970s, misclassification bias, and the effects of rapid social change.

4.2. Findings in the Context of Other Studies

Whilst studies we identified covered the period 1970–2011, insights into suicide patterns in Greenland during earlier periods are provided by studies meeting our exclusion criteria. The Danish physician, Dr. Alfred Bertelsen, recorded 14 suicide cases from 1891–1930 [11]. His findings, published in Danish but reported elsewhere [12], were that eight cases were men, six were women, half the men were under 35 years of age, while all the women were over 35. In cases where he was able to conduct psychological autopsies he ascertained that all had been diagnosed with mental disorders [11]. His estimates of average annual suicide rates were 4 per 100,000 for the period 1891–1903 and 3 per 100,000 for the period 1901–1930, noting that the high comparative rate for accidental deaths suggested misclassification of suicides [12]. A 1955 anthropological study concluded that, at that time, suicide deaths largely occurred in elderly people who were no longer economically active, did not wish to burden their families, and attempted suicide only after consultation with family members [30]. Our findings for the period following this, coinciding with a period of rapid sociocultural and socioeconomic change, suggest a marked shift in risk towards much younger age groups.

Studies specifically describing suicide epidemiology in young men are lacking worldwide, and wide variation in the format of mortality statistics creates problems in comparing temporal patterns by age and gender group [1]. However, there is evidence that indigenous group status predicts high suicide rates in young men in regions neighbouring Greenland; specifically indigenous Sami in Arctic Norway [31], Native American men aged 15–24 in the US [32], and Inuit men aged 15–24 in Canada [33]. A study of men aged 15–34 years in small Alaskan communities from 1980–2007 found that community-level characteristics such as remoteness, fewer non-Natives, and cultural divides had higher suicide risks, whilst those with higher incomes, more married couples, and traditional elders had lower risks [34]. This is consistent with evidence that rural or remote residence is associated with risk of suicide in young men in Denmark, with explanations relating to the migration of healthy workers to cities, and the increasing socioeconomic disparity between men in rural and urban areas [35]. Analysis of data from Danish longitudinal psychiatric registers shows that Greenlandic men aged 15–24 have significantly higher first psychiatric admission rates than Danish men of the same age [36]. Psychiatric disorder is therefore implicated as one explanation for high suicide rates in young men in Greenland compared with their Danish counterparts.

4.3. Strengths and Limitations

4.3.1. Strengths

We used a clear research question, and a comprehensive search strategy, with methods to identify unindexed papers. We registered our study protocol with PROSPERO, followed STROBE and PRISMA guidelines when conducting our review, and independently screened and critiqued papers.

4.3.2. Limitations at Study Level

These have largely been covered under study quality. A major limitation for many studies was that of statistical power, which may explain why included studies provided very limited data on risk factors by gender and age-group, and why many studies aggregated successive years of data. Presenting average age and gender-specific suicide rates for periods of up to 41 years obscured any year-on-year changes and made it hard to identify short-term variations during a period of rapid social change [18]. Even where gender- and age-specific suicide rates were presented for specific years over the period 1970–1995, these lacked statistical tests for time trends, and only related to West Greenland [20] so were unlikely to be generalisable to the rest of Greenland [21]. Without join-point regression models to calculate annual percentage change statistics we lacked an understanding of the magnitude and direction of short-term and long-term trends in age-specific suicide rates.

Misclassification bias was another general potential problem. All studies derived suicide cases from population death registers, or from the death certificates and police reports on which those registers were based. Only one was questionable in terms of whether it used comprehensive methods of identifying all cases [19]. In one included study, 94% of deaths from 1968–1999 (as recorded in death certificates) had been certified by a physician and issued with an underlying International Classification of Diseases (ICD) code for cause of death [21]. This highlighted that where no diagnosis is made or where classified as other injuries, it is possible that there was under-recording of suicide, which may have been differential by age group for cultural reasons. Given high rates of accidental death in young men, and wide international variation in the quality of suicide data, the great potential for underestimating, through misclassification, suicide deaths in young men has been acknowledged [1]. It is possible that high rates of suicide in young men in included studies might be under-estimates. Conversely, it is possible that biases might lead to the deaths of young men being more likely to be classified as suicide than deaths in other demographic groups. Validation studies, however, suggest that registration within official Greenlandic statistics is generally reliable [37].

4.3.3. Limitations at Review Level

Due to our eligibility criteria we may have overlooked studies published in Danish or Greenlandic, and our search for unpublished studies may have been incomplete, thereby reflecting publication bias. By confining our study to studies reporting suicide rates we lacked the context of age and gender patterns in suicidal ideation and suicide attempts over a similar period, particularly where they provided gender [38] or cross-national [39] comparisons.

4.4. Policy Implications

Although the studies we identified suggest that from the mid-1970s to 2011 young men were the demographic group of most concern in Greenland, and that their high rates may have stabilised in 1995 [24], data on suicide rates subsequent to 2011 are lacking. Apart from suicide data published by Statistics Greenland reporting 24 suicides among men aged 15–24 in 2012–2013 [5] we lack clear risk estimates enabling temporal trends to be charted. To ensure that current Greenlandic suicide prevention efforts target the groups at highest risk, there is a clear need for updated mortality indicators to be made available so that up-to-date surveillance is directly linked to suicide prevention activity.

Young people were a major focus of Greenland's 2004 plans for suicide prevention [9], which proposed interventions to improve resilience throughout the education system, and national helplines and radio programmes to overcome barriers such as access to psychological support in remote areas [21,40]. They also proposed strengthening the research base [21], a need reinforced by the findings of this review. The multi-level approaches suggested have been shown to have synergistic benefits [41], but Greenlandic suicide prevention interventions have never been evaluated [42]. Future versions should consider including means restriction, generally understood to be the most effective suicide prevention intervention [43]. The high proportion of violent suicides in Greenland [20], and the historical predominance of hunting/fishing jobs in men who die by suicide [22], suggest a role for gun control interventions such as limiting access to firearms outside working hours. Other issues to address include the impact of labour market changes [10], given qualitative accounts of hopelessness among young men about their opportunities [44]. Evidence from other Arctic Indigenous communities suggests that reform of alcohol policies may not reduce suicides as much as hoped [34]. International collaborations such as the RISING SUN project are likely to be important to Greenland in terms of research infrastructure, and intervention implementation. Implementation of the Canadian Inuit Suicide Prevention Strategy will also be of interest to Greenland, where any successes in Inuit-specific approaches might be emulated [2].

4.5. Future Research

Beyond identifying very high suicide rates in young men in Greenland, the studies we identified in our review covered only a very narrow period, providing little detail on how and when age group transitions occurred, or specific suicide risk factors. There is a clear need for studies describing fine-grained temporal changes by age and gender, using appropriate statistical tests, and charting patterns beyond 2011. There is also a need for studies investigating specific risk factors for suicide in young men in Greenland, as those for all age groups and genders are of limited utility in understanding young men. This relies on improved recording of the clinical and socio-demographic characteristics of all suicide cases, with testing for age and gender variation. However, the problem of a lack of power will continue to be an issue in answering specific research questions at the country level. This has been an issue in studies investigating whether there is a differential effect of seasonality in young people in Greenland, where even combining genders was not felt to have overcome the issue of power [20,45]. Seasonality is of interest because of the potential impact of ambient light on sleep patterns, as well as associations with social behavior and working patterns.

Population registers present a valuable opportunity to conduct population-based analysis including routine clinical data, and to describe how risk factors for suicide are distributed demographically. One such study found that young Greenlandic men are more likely than the general Greenlandic population to be diagnosed with a psychiatric disorder [36]. Using suicide data available for the years following 2011, and the potential for linkage with Danish health registers, there is scope to investigate the contribution of psychiatric disorder to suicide risk in different age groups. Primary data collection will be needed to investigate whether seasonal affective disorder, acculturative stress through rapid modernisation, or imitative suicide [46] are contributory factors, as these have not yet been investigated in young men in Greenland. The association between ambient light and suicide rates, the effects of colonial relationships [47] and of climate change, and the influence of the media reporting of suicide on young people in Greenland are also important areas for investigation.

5. Conclusions

Our systematic review identified evidence to support a dramatic rise in suicide mortality among Greenlandic men aged 15–24 from the mid-1970s, who then represented the highest risk group from 1976–2011 compared with men and women of all age groups in Denmark and Greenland. However, as no articles investigated risk factors for suicide in young men we lack clear explanations for these disparities. Our findings identify a gap in the evidence describing how and when age group transitions occurred in Greenland, and whether this was related to the rapid social change in Greenland at that time. Our findings also identified the need for studies identifying specific suicide risk factors in young men to inform future suicide prevention strategies.

Author Contributions: Conceptualization, H.S. and A.P.; methodology, H.S. and A.P.; database search, H.S.; study screening & critical appraisal, H.S., R.F., and A.P.; writing—original draft preparation, H.S. and A.P.; writing—review and editing, H.S., R.F., and A.P.

Funding: This research received no external funding. Pitman is part-supported by the UCLH NIHR Biomedical Research Centre (BRC).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pitman, A.; Krysinska, K.; Osborn, D.; King, M. Suicide in young men. *Lancet* **2012**, *379*, 2383–2392. [CrossRef]
2. Collins, P.Y.; Delgado, R.A.; Pringle, B.A.; Roca, C.; Phillips, A. Suicide prevention in Arctic indigenous communities. *Lancet Psychiatry* **2017**, *4*, 92–94. [CrossRef]
3. Young, T.K.; Revich, B.; Soininen, L. Suicide in circumpolar regions: An introduction and overview. *Int. J. Circumpolar Health* **2015**, *74*, 27349. [CrossRef] [PubMed]
4. Moltke, I.; Fumagalli, M.; Korneliussen, T.S.; Crawford, J.E.; Bjerregaard, P.; Jørgensen, M.E.; Grarup, N.; Gulløv, H.C.; Linneberg, A.; Pedersen, O.; et al. Uncovering the genetic history of the present-day Greenlandic population. *Am. J. Hum. Genet.* **2015**, *96*, 54–69. [CrossRef] [PubMed]
5. StatbankGreenland. National Board of Health Mortality Statistics 1990–2013. Available online: http://bank.stat.gl/ (accessed on 24 April 2018).
6. WHO. Country Reports and Charts. Available online: http://www.who.int/topics/suicide/en/ (accessed on 27 April 2018).
7. WHO. Global Health Observatory (GHO) Data: Suicide Rates per (100,000 Population). Available online: http://www.who.int/gho/mental_health/suicide_rates_crude/en/ (accessed on 17 April 2018).
8. White, A.; Holmes, M. Patterns of mortality across 44 countries among men and women aged 15–44 years. *J. Mens Health Gender* **2006**, *3*, 139–151. [CrossRef]
9. PAARISA. *Proposal for a National Strategy for the Prevention of Suicides in Greenland*; PAARISA: Nuuk, Greenland, 2004.
10. Bjerregaard, P.; Larsen, C.V. Three lifestyle-related issues of major significance for public health among the Inuit in contemporary Greenland: A review of adverse childhood conditions, obesity, and smoking in a period of social transition. *Public Health Rev.* **2018**, *39*, 5. [CrossRef] [PubMed]
11. Bertelsen, A. Gronlandsk medicinsk statistik og nosografi (medical statistics and nosography of Greenland). *Meddl. Gronland* **1935**, *1*, 58–59.
12. Lynge, I. Suicide in Greenland. *Arct. Med. Res.* **1985**, *40*, 53–60.
13. Kapur, N.; Cooper, J.; O'Connor, R.C.; Hawton, K. Non-suicidal self-injury v. Attempted suicide: New diagnosis or false dichotomy? *Br. J. Psychiatry* **2013**, *202*, 326–328. [CrossRef] [PubMed]
14. Cooper, J.; Kapur, N.; Webb, R.; Lawlor, M.; Guthrie, E.; Mackway-Jones, K.; Appleby, L. Suicide after deliberate self-harm: A 4-year cohort study. *Am. J. Psychiatry* **2005**, *162*, 297–303. [CrossRef] [PubMed]

15. Liberati, A.; Altman, D.G.; Tetzlaff, J.; Mulrow, C.; Gøtzsche, P.C.; Ioannidis, J.P.A.; Clarke, M.; Devereaux, P.J.; Kleijnen, J.; Moher, D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. *BMJ* **2009**, *339*. [CrossRef] [PubMed]
16. Elm, E.v.; Altman, D.G.; Egger, M.; Pocock, S.J.P.; Gøtzsche, P.C.; Vandenbroucke, J.P. Strengthening the reporting of observational studies in epidemiology (strobe) statement: Guidelines for reporting observational studies. *BMJ* **2007**, *335*, 806–808. [CrossRef] [PubMed]
17. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med.* **2009**, *6*, e1000097. [CrossRef] [PubMed]
18. Bjerregaard, P.; Larsen, C. Time trend by region of suicides and suicidal thoughts among Greenland Inuit. *Int. J. Circumpolar Health* **2015**, *74*, 26053. [CrossRef] [PubMed]
19. Grove, O.; Lynge, J. Suicide and attempted suicide in Greenland. A controlled study in Nuuk (godthaab). *Acta Psychiatr. Scand.* **1979**, *60*, 375–391. [CrossRef] [PubMed]
20. Bjorksten, K.S.; Bjerregaard, P.; Kripke, D.F. Suicides in the midnight sun–a study of seasonality in suicides in West Greenland. *Psychiatry Res.* **2005**, *133*, 205–213. [CrossRef] [PubMed]
21. Bjerregaard, P.; Lynge, I. Suicide—A challenge in modern Greenland. *Arch. Suicide Res.* **2006**, *10*, 209–220. [CrossRef] [PubMed]
22. Thorslund, J. Suicide among Inuit youth in Greenland 1977–1986. *Arct. Med. Res.* **1991**, *50*, 299–302.
23. Thorslund, J. Inuit suicides in Greenland. *Arct. Med. Res.* **1990**, *49*, 25–33.
24. Leineweber, M.; Arensman, E. Culture change and mental health: The epidemiology of suicide in Greenland. *Arch. Suicide Res.* **2003**, *7*, 41–50. [CrossRef]
25. Blanchard, R.D.; Bunker, J.B.; Wachs, M. Distinguishing aging, period and cohort effects in longitudinal studies of elderly populations. *Socio-Econ. Plan. Sci.* **1977**, *11*, 137–146. [CrossRef]
26. World Health Organization (WHO). *Multisite Intervention Study on Suicidal Behaviours-Supre-Miss: Protocol of Supre-Miss*; World Health Organization: Geneva, Switzerland, 2002.
27. Box, J.E. Survey of Greenland instrumental temperature records: 1873–2001. *Int. J. Climatol.* **2002**, *22*, 1829–1847. [CrossRef]
28. Government of Greenland. Politics in Greenland. Available online: http://naalakkersuisut.gl/en/About-government-of-greenland/About-Greenland/Politics-in-Greenland (accessed on 30 September 2018).
29. Bjerregaard, P. Rapid sociocultural change and health in the Arctic. *Int. J. Circumpolar Health* **2001**, *60*, 102–111. [PubMed]
30. Leighton, A.H.; Hughes, C.C. Notes on Eskimo patterns of suicide. *Southwest. J. Anthropol.* **1955**, *11*, 327–338. [CrossRef]
31. Silviken, A.; Haldorsen, T.; Kvernmo, S. Suicide among indigenous Sami in Arctic Norway, 1970–1998. *Eur. J. Epidemiol.* **2006**, *21*, 707–713. [CrossRef] [PubMed]
32. EchoHawk, M. Suicide prevention efforts in one area of Indian health service, USA. *Arch. Suicide Res.* **2006**, *10*, 169–176. [CrossRef] [PubMed]
33. Boothroyd, L.J.; Kirmayer, L.J.; Spreng, S.; Malus, M.; Hodgins, S. Completed suicides among the Inuit of northern Quebec, 1982–1996: A case-control study. *CMAJ Can. Med. Assoc. J.* **2001**, *165*, 749–755.
34. Berman, M. Suicide among young Alaska native men: Community risk factors and alcohol control. *Am. J. Public Health* **2014**, *104*, S329–S335. [CrossRef] [PubMed]
35. Qin, P. Suicide risk in relation to level of urbanicity—A population-based linkage study. *Int. J. Epidemiol.* **2005**, *34*, 846–852. [CrossRef] [PubMed]
36. Lynge, I.; Mortensen, P.B.; Munk-Jorgensen, P. Mental disorders in the Greenlandic population. A register study. *Int. J. Circumpolar Health* **1999**, *58*, 188–197. [PubMed]
37. Thorslund, J.; Misfeldt, J. On suicide statistics. *Arct. Med. Res.* **1989**, *48*, 124–130.
38. Bjerregaard, P.; Curtis, T. Cultural change and mental health in Greenland: The association of childhood conditions, language, and urbanization with mental health and suicidal thoughts among the Inuit of Greenland. *Soc. Sci. Med.* **2002**, *54*, 33–48. [CrossRef]
39. Broderstad, A.R.; Eliassen, B.M.; Melhus, M. Prevalence of self-reported suicidal thoughts in SLiCA. The survey of living condition in the Arctic (SLiCA). *Glob. Health Action* **2011**, *4*, 10226. [CrossRef] [PubMed]
40. Le Fevre, A.C. The challenge of reducing youth suicide in Greenland—Interventions, strategies and roads to be explored. *Int. J. Circumpolar Health* **2004**, *63* (Suppl. 2), 38–40. [CrossRef] [PubMed]

41. Van der Feltz-Cornelis, C.M.; Sarchiapone, M.; Postuvan, V.; Volker, D.; Roskar, S.; Grum, A.T.; Carli, V.; McDaid, D.; O'connor, R.; Maxwell, M. Best practice elements of multilevel suicide prevention strategies. *Crisis* **2011**, *32*, 319–333. [CrossRef] [PubMed]
42. Redvers, J.; Bjerregaard, P.; Eriksen, H.; Fanian, S.; Healey, G.; Hiratsuka, V.; Jong, M.; Larsen, C.V.L.; Linton, J.; Pollock, N. A scoping review of indigenous suicide prevention in circumpolar regions. *Int. J. Circumpolar Health* **2015**, *74*, 27509. [CrossRef] [PubMed]
43. Zalsman, G.; Hawton, K.; Wasserman, D.; van Heeringen, K.; Arensman, E.; Sarchiapone, M.; Carli, V.; Höschl, C.; Barzilay, R.; Balazs, J. Suicide prevention strategies revisited: 10-year systematic review. *Lancet Psychiatry* **2016**, *3*, 646–659. [CrossRef]
44. Soule, S. *An Evaluation of the Implementation of Greenland's National Strategy for Suicide Prevention with Recommendations for the Future*; PAARISA: Nuuk, Greenland, 2008.
45. Björkstén, K.S.; Bjerregaard, P. Season of birth is different in Inuit suicide victims born into traditional than into modern lifestyle: A register study from Greenland. *BMC Psychiatry* **2015**, *15*, 147. [CrossRef] [PubMed]
46. Leineweber, M.; Bjerregaard, P.; Baerveldt, C.; Voestermans, P. Suicide in a society in transition. *Int. J. Circumpolar Health* **2001**, *60*, 280–287. [PubMed]
47. Bolliger, L.; Gulis, G. The tragedy of becoming tired of living: Youth and young adults' suicide in Greenland and Denmark. *Int. J. Soc. Psychiatry* **2018**, *64*, 389–395. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

An Epidemiological Analysis of Head Injuries in Taiwan

I-Lin Hsu [1], Chung-Yi Li [2,3], Da-Chen Chu [4] and Li-Chien Chien [5,*]

1 Department of Surgery, College of Medicine, National Cheng Kung University Hospital, National Cheng Kung University, Tainan 70403, Taiwan; yilinhsu@gmail.com
2 Department of Public Health, College of Medicine, National Cheng Kung University, Tainan 70403, Taiwan; cyli99@mail.ncku.edu.tw
3 Department of Public Health, College of Public Health, China Medical University, Taichung 40402, Taiwan
4 Department of Neurosurgery, Taipei City Hospital, Taipei 10065, Taiwan; DAD57@tpech.gov.tw
5 Department of Emergency Medicine, Taipei City Hospital, Taipei 10065, Taiwan
* Correspondence: maxchien@mail2000.com.tw; Tel.: +886-988360837; Fax: +886-6-276-6676

Received: 30 September 2018; Accepted: 1 November 2018; Published: 4 November 2018

Abstract: Traumatic head injuries occur frequently in Taiwan, having catastrophic consequences for the victims, their families, and society as a whole. However, little is known about the risk factors at the population level in Taiwan. The primary aim of this study was to obtain more information on these variables and their relationships. Another aim was to analyze the effects of independent variables such as sex, age, residency, pre-existing conditions, mechanisms of injury, associated injuries, and severity on the probability of in-hospital death. Using the 2007–2008 total admissions claim dataset from Taiwan's National Health Insurance system, total admissions due to acute head injury were selected for further analysis. The obtained data included patient demographics and trauma hospitalization rate. A total of 99,391 patients were admitted with head injury, 48,792 of which had moderate-to-severe head injury. There were 4935 cases recorded as in-hospital mortality and the standardized in-hospital mortality rate was 10.7 deaths per 100,000 person-years. The mortality rate increased with age. After adjustments, male sex, age older than 54 years, living in a rural area, lower monthly income, a Charlson comorbidity index greater than one, being a pedestrian hit by a motor vehicle, fall from a height, and having significant chest, abdominal, or lower extremity injury increased the risk of death during admission. This population-based analysis provides information about the incidence rate and death rate for admissions in Taiwan due to acute head injury and the factors that affect in-hospital mortality. Our results that highlight the risk factors for adverse outcome can help us prevent or improve rural area trauma care of head injury patients in the future.

Keywords: head injury; incidence; national health insurance; population-based study; mortality; traumatic brain injury

1. Introduction

Trauma is a devastating cause of morbidity and mortality in Taiwan and other countries around the world. Head injury is still one of the leading causes of death in modern society [1–3]. Approximately 1.5 million patients are suffering from traumatic head injuries in the United States [4] and traumatic head injuries caused more than 53,000 deaths every year during 1997 to 2007, at a rate of 18.4 deaths per 100,000 people [5]. Critical head injuries are preventable deaths. Head injuries lead to tremendous loss of capacity and resources in long-term care. Some studies have noted the severity of social burden caused by head injuries and claimed that more than half of trauma-related deaths are caused by head injury [6,7]. Evidence has accumulated from studies in other regions indicating that head injury is a global health threat [8–11]. Some epidemiological reports have revealed various incidence rates in different countries [7,12–18]. Comparison amongst studies is difficult because the definitions of head injury, socio-economic status, and the inclusion criteria are different. A descriptive, population-based study of patients with severe associated head injuries has not been carried out in Taiwan. Therefore, our aim was to estimate the trauma hospitalization rate in Taiwan due to head injury and describe where people are being treated for their injuries. We also aimed to investigate the prevalence of concomitant injuries among hospitalized patients with acute head injury in Taiwan, and how hospital outcomes are affected by variables including (1) patient characteristics including age, sex, residence, and presence of pre-existing conditions (PECs) and (2) characteristics of the injury including body region of principal injury, trauma severity, mechanism of the injury. There are two major events affecting trauma case mortality rate: the introduction of the motorcycle helmet law in 1997 and the new trauma care hospital classification system in 2009. The helmet law effectively decreased the mortality and morbidity from motorcycle-related head injuries after its inception [19]. Even in rural areas, people now wear helmets when riding their motorcycle leading to a decrease in accident mortality rate. We wanted to find more predictors associated with mortality of head injury allowing policy makers to introduce new policies or improve medical care, thus decreasing the loss.

2. Methods

In Taiwan, the National Health Insurance (NHI) system covers nearly 99% of the population and this study retrieved the entire claim admissions dataset for those who were coded with discharged diagnostic numbers defined by the International Classification of Diseases, 9th revision, Clinical Modification (ICD-9-CM). Therefore, this study included all NHI trauma patients who were hospitalized and discharged between 1 January 2007 and 31 December 2008 and is population-based. We captured all the trauma hospitalizations of patients defined as any admission claim coded with at least one ICD-9-CM diagnosis between 800.00 and 959.99, except 905.00 to 909.99, which indicate "late effects of injuries, poisonings, toxic effects and other external causes: 930.00 to 939.99, which include "effects of a foreign body entering through an orifice", and 958, "complications". In Taiwan, computed tomography (CT) for patients with traumatic head injury is generally covered by NHI and moderate-to-severe head injury was mainly diagnosed by radiological findings on CT, as in other studies [13]. This research is a based on the national data bank with informed consent. This research did not involve human participants and/or animals.

2.1. Variables

According to capacity and services, hospitals in Taiwan were classified into three levels: medical centers (MC), regional hospitals (RH), and local hospitals (LH). Taiwan's definition for medical centers and regional hospitals includes capacity and services. Empirically, the patients were stratified into nine age groups: 0–14, 15–24, 25–34, 35–44, 45–54, 55–64, 65–74, 75–84, and >84 years of age. The patient area of residency was classified into two groups: urban (including suburban) and rural areas. The average population density (persons/km^2) is higher in urban areas (2635 persons/km^2) than in rural areas (230 persons/km^2). According to the level of monthly income, injured patients were classified into four groups: monthly income less than USD $660, between USD $660 and $1320, more than USD $1320, and dependents, which are those who were injured but not employed, including those insured (NHI) by a spouse, other family member, or social welfare. During the study period, the average Taiwan monthly living cost was USD $595/person. Pre-existing conditions were defined as the medical co-morbidities of the injured patients who were diagnosed and recorded with ICD-9-CM codes in the same admission data file [18]. We used Charlson's comorbidity index to quantify the PECs [20]. According to Charlson's study in 1987, 17 diseases are included in the formula for calculating the Charlson comorbidity index [21]. The patients were classified into four levels according to their Charlson comorbidity index: 0, 1, 2, and more than 2.

For assessing injury severity, a computerized mapping method that used the ICD codes to obtain injury severity scores, such as the AIS, was employed [22,23]. A computerized mapping system, ICDMAP, for converting injury-related International Classification of Diseases, ninth revision (ICD-9-CM) rubrics into AIS scores, was proposed by MacKenzie et al. in 1989, and their results have been verified [24]. This ICD mapping system has been refined over the years and has been used in several large or population-based studies to classify severity using ICD diagnostic codes [25,26]. Our study applied ICDMAP to the dataset and derived ICD/AIS scores for each injury diagnosis and an ICD/ISS (injury severity score) for each admission.

As the patient group coded with an AIS head score of 1 or 2, as generated by ICDMAP software, was classified as mild head injury, we defined moderate-to-severe head injury as significant head injury with an AIS head score more than 2, which included moderate (AIS head = 3) and severe head injury (AIS head 4, 5, and 6), as shown in Table 1.

Table 1. Abbreviated Injury Score (AIS) head score and head injury severity classification.

AIS Head Score	Head Injury Severity Classification
AIS score 1–2	Mild
AIS score 3	Moderate
AIS score 4–6	Severe

ICD E-codes in the claim data were also used to classify the mechanism of injury [27]. A patient is classified as the driver or passenger in a motor vehicle traffic accident if the number after the decimal point is zero or one, respectively (e.g., E811.0, E811.1, E812.0, or E812.1). A patient is classified as a motorcycle rider or passenger injured in a traffic accident if the number after the decimal is two or three, respectively (e.g., E811.2, E811.3, E812.2, or E812.3). A patient is classified as having an injury of an unspecified nature in a motor vehicle traffic accident if the number after the decimal point is nine (e.g., E811.9, E812.9, etc.), as shown in Table 2.

Table 2. ICD-E codes and injury mechanisms.

ICD-E Code	Injury Mechanism
The number after the decimal point is zero or one, e.g., E811.0	Driver or passenger in a motor vehicle traffic accident
The number after the decimal point is two or three, e.g., E811.2	Motorcycle rider or passenger injured in a traffic accident
The number after the decimal point is nine, e.g., E811.9	An injury of an unspecified nature in a motor vehicle traffic accident

ICD-E code: International Classification of Disease-external cause of injury code.

2.2. Statistical Analysis

The percentage of sex, four categories of monthly income, four categories of PECs, and mortality rates were compared to understand the differences between all patients with head injuries and the patients with moderate-to-severe head injuries. Means, medians, and standard deviations of continuous variables, such as age, monthly income, Charlson comorbidity index, and ICD/ISS were analyzed using analysis of variance (ANOVA). The relative odds of death in the admission course were estimated by using binary logistic regression with the backward conditional stepwise method. Statistical significance was set at $p < 0.05$. All the data analyses were generated by using SPSS software, version 20.0 (IBM Corporation, New York, NY, USA). Institutional board review was waived because this study involved secondary data analysis of a well-encrypted database.

3. Results

3.1. Incidence of Hospitalization and In-Hospital Fatality Rate

During 2007–2008, 476,241 cases were retrieved as trauma admissions. There were 50,599 cases of mild head injury in which the AIS in the head region was 1 or 2. Among these cases, 151 deaths occurred and the in-hospital mortality rate was 0.3%. There were 9666 cases of moderate head injury (AIS head of 3) and 126 deaths were recognized; the in-hospital mortality rate was 1.3%. There were 39,126 cases of severe head injury (AIS head of 4, 5, or 6) and 4809 deaths were recognized; the in-hospital mortality rate was 12.3%. The raw hospitalization rate for all head injuries in Taiwan was 215.3 per 100,000, and for moderate-to-severe head injury, 105.9 per 100,000. After calculations using the U.S. 2000 standard population, the standardized incidence of all head injury was 220.6 per 100,000 in Taiwan and the standardized incidence of moderate-to-severe head injury was 110.5 per 100,000. Table 3 shows the incidence of trauma hospitalization with head injury in the nine age strata. The lowest trauma hospitalization rate for patients with moderate-to-severe head injury in Taiwan was in the 0–14 year age group (29.8/100,000), followed by a higher incidence in the 15–24 year age group (97.9/199,999), and gradually increased up to 582.4/100,000 in the age group older than 84 years. The graph pattern of hospitalization rate in the nine age strata and the in-hospital fatality rate for patients with moderate-to-severe head injury are shown in Figure 1. The in-hospital mortality rate for patients with moderate-to-severe head injury was 10.7% and the standardized hospitalization death rate was 11.8/100,000. The hospitalization death rate for patients with moderate-to-severe head injury increased gradually from 1.2/100,000 in the group younger than 15 years old to the highest rate, 108.5/100,000, in the group older than 84 years, except for slightly lower rates in the 25–34 year age group than in the 15–24 year age group. Men had higher rates than women in all age groups.

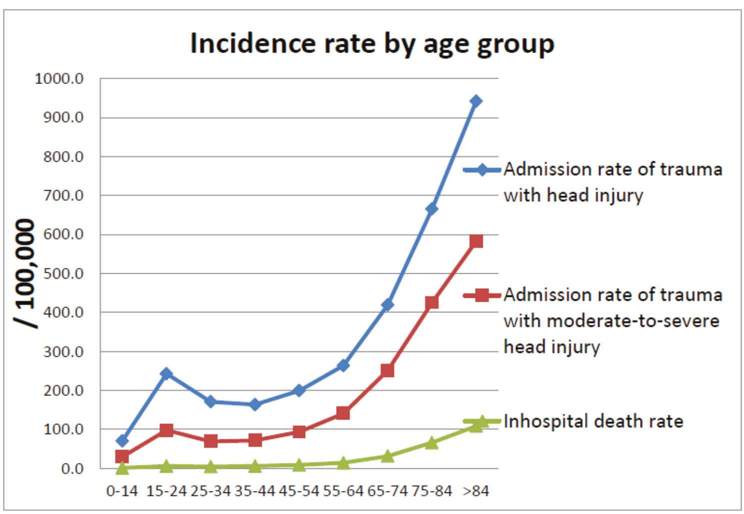


Figure 1. Incidence rate of head injury admission and death by age groups.

Table 3. Age distribution of trauma hospitalizations with head injury in Taiwan in 2007–2008.

Age	Total	Hospitalizations with Head Injury	% of Total Admission with Head Injury	Moderate-to-Severe Head Injury	Population in Taiwan, 2008	Admission Rate of Trauma with Moderate-to-Severe Head Injury (per 100,000)	In-hospital Deaths with Moderate-to-Severe Head Injury	In-hospital Death Rate in Age Group (per 100,000)
0–14	27,764	5497	5.5%	2324	3,905,203	29.8	90	1.2
15–24	61,981	15,778	15.9%	6363	3,248,257	97.9	383	5.9
25–34	61,409	13,368	13.4%	5404	3,911,237	69.1	342	4.4
35–44	59,961	12,099	12.2%	5302	3,702,101	71.6	452	6.1
45–54	72,028	14,405	14.5%	6751	3,611,743	93.5	633	8.8
55–64	59,074	11,901	12.0%	6358	2,256,270	140.9	642	14.2
65–74	55,950	11,506	11.6%	6886	1,372,060	250.9	859	31.3
75–84	56,108	10,990	11.1%	7026	825,992	425.3	1091	66.0
>84	21,966	3847	3.9%	2378	204,168	582.4	443	108.5
Total	476,241	99,391	100.0%	48,792	23,037,031	105.9	4935	10.7

In Taiwan, 6,935,205 people resided in rural areas in 2008. During the study period, 41,570 rural people were admitted with head injury and 18,108 of these rural people had moderate-to severe head injury. Thus, focusing on moderate-to-severe head injury, the patients who resided in rural areas had a higher hospitalization rate than those residing in urban areas (130.6 vs. 95.3 per 100,000, respectively; $p < 0.01$). Among the patients who had in-hospital mortality, 2038 were registered in rural areas. The rural population yielded a higher raw hospitalization death rate for patients with moderate-to-severe head injury than the urban population (14.2 vs. 9.2 per 100,000, respectively; $p = 0.005$). Monthly distribution of case numbers during 2007–2008 is shown in Figure 2, and in general, the occurrence did not significantly change with month.

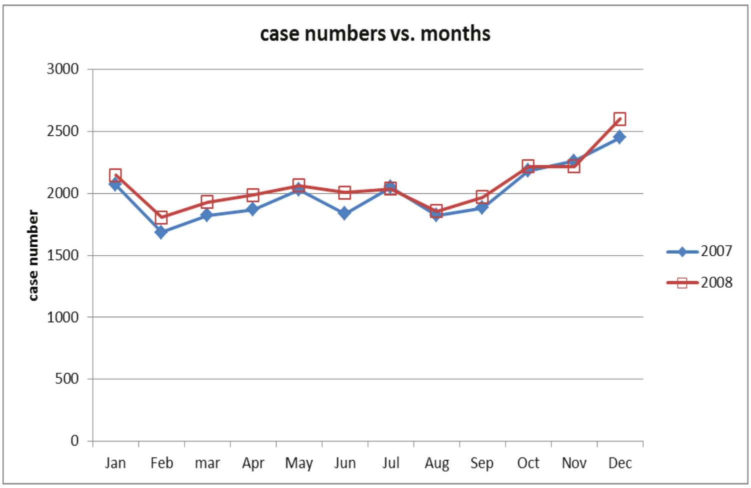


Figure 2. Monthly distribution of case numbers during 2007–2008.

Table 4 presents the demographics of two groups: all patients with head injury and those with moderate-to-severe head injury. Table 2 summarizes the comparisons of age, sex, and other demographics, the associated injury characteristics, and distribution of the locations of treatment between these two groups. The patients with moderate-to-severe injury were older (50.6 ± 23.1 vs. 46.8 ± 22.9 years), a higher proportion were men, and rural people had a higher Charlson comorbidity index and a higher ICDISS score (18.2 ± 6.4 vs. 12.0 ± 8.1 years), and had a longer median stay in the hospital (7, 4–13 vs. 5, 2–6 (days, 25th–75th centile). In the moderate-to-severe head injury group, more patients were treated at a medical center and fewer at a local hospital. Figure 3 shows the different distribution of the mechanisms of injury among these nine age groups. Among those in the younger population, being injured as a motorcycle rider or passenger was more frequently the external cause, and a fall on the same level was more frequently the cause in older patients. The proportion of case numbers for motorcycle injury in people aged 15–24 years was significantly higher than that in people of other ages (49.4% vs. 26.4%, $p < 0.001$).

Table 4. Patients demographics in mild and moderate-to-severe head injury cases.

Patient Characteristics	Total Head Injury Cases	Moderate-to-Severe Head Injury Cases	*p*-Value
Variable	***N* = 99,391**	***N* = 48,792 (49.1%)**	
Age (years)			
Mean (± SD)	46.8 ± 22.9	50.6 ± 23.1	<0.001
>64 years old	26,343 (26.5%)	16,290 (33.4%)	<0.001
Sex			
Male	61,746 (62.1%)	32,072 (65.7%)	<0.001
Residency			
Rural	41,570 (41.8%)	18,108 (37.1%)	<0.001
Income			
Income group (dependent)	30,673 (30.9%)	14,793 (30.7%)	<0.001
Income group (<USD $660)	28,071 (28.2%)	14,257 (29.6%)	
Income group (USD $660–1320)	35,719 (35.9%)	16,792 (34.8%)	
Income group (>USD $1320)	4918 (4.9%)	2404 (5.0%)	
Charlson Index groups			
Charlson index = 0	81,865 (82.4%)	38,159 (78.2%)	<0.001
Charlson index = 1	12,551 (12.6%)	7233 (14.8%)	
Charlson index = 2	3317 (3.3%)	2231 (4.6%)	
Charlson index > 2	1658 (1.7%)	1169 (2.4%)	
Mechanism of Injury			
Car accidents	2785 (2.8%)	917 (1.9%)	<0.001
Motorcycle riders or passengers	29,930 (30.1%)	12,586 (25.8%)	
Bicycle	2219 (2.4%)	981 (1.9%)	
Pedestrians (hit by motor vehicle)	2751 (2.8%)	1488 (3.0%)	
Traffic accident, mechanism unknown	8722 (8.8%)	4456 (9.1%)	
Same level fall	7072 (14.0%)	8694 (17.8%)	
Fall from a height	15,705 (4.9%)	2716 (5.6%)	
Missing	20,767 (20.9%)	12,598 (25.8%)	
Associated injury (ICD/ISS ≥ 3)			
Face injury	28	12	0.279
Neck injury	21	13	0.573
Chest injury	3397 (3.4%)	2298 (4.7%)	<0.001
Abdominal and pelvic injury	1017 (1.0%)	531 (1.1%)	0.047
Spinal injury (Thoraco-lumbar spine)	1591 (1.6%)	658 (1.3%)	<0.001
Upper extremity injury	444 (0.4%)	171 (0.4%)	<0.001
Lower extremity injury	3099 (3.1%)	1674 (3.4%)	<0.001
ICDISS			
Mean	12.0 ± 8.1	18.2 ± 6.4	<0.001
Level of treating hospitals			
Medical Center (MC)	20,546 (20.7%)	16,611 (34.0%)	<0.001
Regional Hospitals	46,258 (46.5%)	23,354 (47.9%)	
Local Hospitals	32,587 (32.8%)	8827 (18.1%)	
Ownership of treating hospitals			
Public Hospitals	22,439 (22.6%)	10,795 (22.1%)	0.001
Treatment Outcome			
Discharge Status			
Death	5086(5.1%)	4935 (10.1%)	<0.001
Length of Stay (day)			
Mean	7.8	10.7	<0.001
Median	5	7	
25th–75th centile	2–6	4–13	

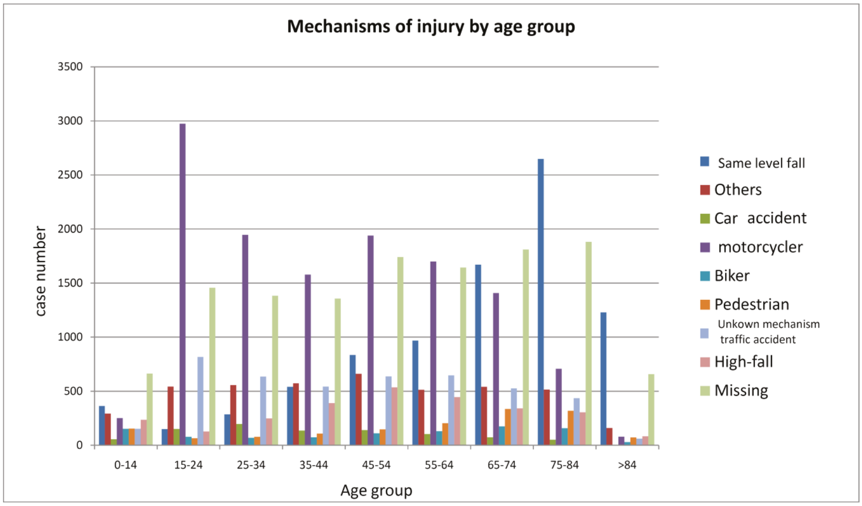


Figure 3. Mechanisms of head injury by age groups.

3.2. Results of Regression Analyses

We included characteristics of the patients and their injuries, including age, sex, rural residency, level of monthly income, Charlson comorbidity index, mechanism of injury, and ICD/ISS in a logistic regression to estimate the adjusted relative odds of in-hospital fatality for the patients with moderate-to-severe head injury. The results are shown in Table 5. Different age strata experienced different effects; being younger than 15 years decreases the risk of death, but being older than 54 years is associated with a higher risk of death. The older the patient, the higher the hospitalization mortality rate. Rural residency increases the risk of death (odds ratio (OR): 1.19, 95% confidence interval (CI): 1.11–1.27). Compared with the lowest level of monthly income (less than USD $660), the group with a monthly income between USD $660 and 1320 had a lower probability of death (OR: 0.78, 95% CI: 0.72–0.85), but dependents had a higher probability of in-hospital death (OR: 1.49, 95% CI: 1.38–1.61) Compared with the patients treated at medical centers, being treated at a regional hospital was associated with a higher probability of in-hospital fatality (OR: 1.11, 95% CI: 1.03–1.19). Compared with the patients treated at private hospitals, being treated at a public hospital was associated with slightly lower risk of death (OR: 0.92, 95% CI: 0.86–1.00).

As an indicator of PECs, higher Charlson comorbidity index scores were associated with a higher risk of death. The relative odds of dying in the group with a Charlson comorbidity index of two was 1.52 times higher than those for the group with a Charlson comorbidity index of zero (95% CI: 1.34–1.72) and the relative odds of dying in the group with a Charlson comorbidity index more than two was 2.5 times higher (95% CI: 2.16–2.90). Several mechanisms of injury did not significantly affect the probability of death. Compared with falls on the same level, only pedestrians hit by a motor vehicle and falls from a height had a higher probability of in-hospital death. Furthermore, being associated with another significant injury significantly increased the risk of mortality; the mortality rate in the group associated with a significant abdominal, chest, or lower extremity injury was 3.41, 3.41, and 1.37 times higher, respectively, than the group without a significant injury.

Table 5. Regression results. Adjusted odds of in-hospital death for the patients with moderate-to-severe head injury.

Vairables	Relative Odds	Lower 95% CI	Higher 95% CI	*p*-Value
Men (Women as reference)	1.18	1.10	1.26	<0.001
Age (15–54 years old (y/o) as reference)				
Child (<15 y/o)	0.53	0.42	0.66	<0.001
55–64 y/o	1.43	1.29	1.58	<0.001
65–74 y/o	1.77	1.62	1.95	<0.001
>74 y/o	2.38	2.19	2.59	<0.001
Residency (urban as reference)				
Rural resident	1.19	1.11	1.27	<0.001
Income group (<USD $660 USD as reference)				
Income group (dependent)	1.49	1.38	1.61	<0.001
Income group (USD $660–1320)	0.78	0.72	0.85	<0.001
Income group (>USD $1320)	0.97	0.82	1.14	0.717
Level of treating places (Medical center as reference)				
Regional hospitals	1.11	1.03	1.19	0.004
Local hospitals	0.52	0.47	0.58	<0.001
Being treated at a public hospital	0.92	0.86	1.00	0.037
PECs (Charlson index = 0 as reference)				
Charlson index = 1	1.08	0.99	1.18	0.066
Charlson index = 2	1.52	1.34	1.72	<0.001
Charlson index > 2	2.50	2.16	2.90	<0.001
Mechanism (Fall on the same level as reference)				
Other mechanism	1.08	0.95	1.23	0.223
Car driver or passenger	1.09	0.86	1.39	0.483
Motorcycle rider or passenger	0.93	0.84	1.03	0.169
Pedal cyclist	0.93	0.73	1.18	0.540
Pedestrian hit by auto vehicle	1.60	1.36	1.89	<0.001
Traffic accident without clear mechanism	1.11	0.98	1.26	0.108
Fall from a height	1.58	1.38	1.81	<0.001
E-code missing	1.03	0.94	1.14	0.532
Anatomic injury (ICD/AIS < 3 as reference)				
Severe chest injury	3.41	3.07	3.79	<0.001
Severe Abdominal injury	3.41	2.76	4.21	<0.001
Severe lower limb	1.37	1.18	1.58	<0.001
Constant	0.05			<0.001

4. Discussion

Head injury continues to be a major health problem around the world in both developed and developing countries [1–4]. The aim of this study was to examine the incidence and in-hospital head injury-related mortality in Taiwan. Compared with other regions or countries, the incidence rate in our study, 220.6/100,000, was higher than in the U.S. (103/100,000), Finland (101/100,000), and India (160/100,000) but lower than in Europe (235/100,000) [2,9,16,28]. However, the definitions and inclusion criteria are different, which complicates comparisons between studies and regions.

Among those aged 15 to 74 years, the most common mechanism of injury is motorcycle-related traffic accident because motorcycles are the most popular transportation vehicles in Taiwan. Motorcycle riders and passengers should be the target population for injury prevention, even in the age group between 65 to 74 years. For those older than 65 years, falls caused more than two-thirds of the injuries and hospitalization rates in this age group. This deserves more attention because trauma does not only affect the younger population; therefore, a focus on injury prevention programs should be aggressively advocated for the elderly. For the people younger than 15 years old in Taiwan in the study by Tsai et al. in 2004, the incidence rate was higher in the age groups of 4 to 9 and 10 to 14 years. The main cause of pediatric head injury was traffic injury followed by falls. Of all pediatric traffic injuries, motorcycle-related injury had the highest incidence, followed by pedestrian and bicycle-related injury [29].

In the United States, head injuries cause about 2.5 million emergency department (ED) visits, 280,000 hospitalizations, and 50,000 deaths annually [4]. In Taiwan, our study showed a hospitalization death rate of 11.8/100,000. Compared with the head injury-related death rates in the U.S. (18.1/100,000) [18], Finland (18.3/100,000) [28], India (20.0/100,000), and Europe (15.4/100,000) [22], the standardized death rate in Taiwan, at 11.8/100,000, was lower than in these other studies. However, our dataset only included patients who had been admitted to hospital, and head injury-related accidents caused 4935 in-hospital deaths among the total 7640 hospitalization mortality cases in 2007–2008 in Taiwan. The dataset did not include deaths at the scene of the accident, during transportation, or in the emergency department. In Taiwan, there were 14,207 accidental deaths in 2007–2008 and our in-hospital dataset only included half of the deaths. It is reasonable to assume that the annual head injury-related mortality rate is underestimated in this study. With the aging population, the devastating effects of head injury will become increasingly serious.

Our regression model showed that only pedestrians hit by a motor vehicle and those who had a fall from a height were associated with a higher probability of mortality; other causes of head injury were not. This result might be caused by the lack of helmet protection and pedestrians hit by a motor vehicle is a well-known dangerous mechanism of injury [30]. In Taiwan, NHI provides a high accessibility to hospital resources to all citizens, but our results show that rural residents and the subgroup with lower monthly income were associated with lower survival probability after moderate-to-severe head trauma. Some researchers have found that the medical institutions providing higher levels of trauma care are often located in urban areas [31]. Their results showed that the odds ratio for death was higher in the hospitals located in rural areas than in hospitals located in more urbanized areas. One study found huge rural-urban disparities in mortality from unintentional injuries [32]. Lower monthly incoming and living in a rural area are associated with many factors leading to higher mortality rate, such as low education and health literacy, risky environments, and low economic support for medical care. We need to focus on these vulnerable populations, reinforce cause-specific prevention programs to reduce the mortality rate, relocate more resources to fairly provide optimal care, and periodically monitor the effectiveness of these programs.

This study had some limitations. First, this is a retrospective study including all the NHI claim data about injury-related hospitalizations during the study period. However, the accuracy of NHI claim data is assured given the severe penalty for fraud and erroneous claiming. Second, this dataset does not include the Glasgow Coma Scale (GCS), which may be misleading for assessing the severity of head injury. The NHI covers brain CT for head injury in an emergency setting, and the severity score ICD/ISS in this study was basically determined by brain CT. The third limitation is that nearly 10% of all admissions were recorded as traffic accidents without further information about the mechanisms, and one-quarter of admissions were recorded without any information about the mechanism of injury. This is reasonable because there was a higher proportion of trauma patients who were found injured without any witnesses or other information source compared to in the United States [14]. Because this study is based on the NHI research database, some detailed information that was not included may lead to ignoring some clinical important factors such as alcohol or drug consumption, related to head injury mortality.

5. Conclusions

In conclusion, this epidemiological study that focused on traumatic head injury hospitalizations in Taiwan is the first population-based study providing abundant objective information about the severity of the damage caused by head injury in our society. Although the government and medical institutes have attempted to mitigate the damage caused by head injury for decades, head injury still requires public attention to alleviate the impact on human health and society. Our results that highlight the risk factors for adverse outcome can help us to prevent or improve rural area trauma care for head injury patients in the future.

Author Contributions: Conceptualization, L.-C.C. and I-L.H.; Methodology, C.-Y.L.; Software, L.-C.C. and I-L.H.; Validation, L.-C.C., I-L.H., D.-C.C., and C.-Y.L.; Formal Analysis, L.-C.C. and I-L.H.; Investigation, L.-C.C. and D.-C.C.; Resources, L.-C.C. and I-L.H.; Data Curation, L.-C.C.; Writing-Original Draft Preparation, I-L.H.; Writing-Review & Editing, L.-C.C. and C.-Y.L.; Visualization, I-L.H.; Supervision, D.-C.C. and C.-Y.L.; Project Administration, L.-C.C. and I-L.H.; Funding Acquisition, L.-C.C.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Department of Statistics, Ministry of Health and Welfare, Taiwan. 2011 Statistics of Causes of Death. Available online: http://www.mohw.gov.tw/EN/Ministry/Statistic.aspx?f_list_no=474&fod_list_no=3485 (accessed on 20 January 2015).
2. Langlois, J.A.; Rutland-Brown, W.; Thomas, K.E. *Traumatic Brain Injury in the United States: Emergency Department Visits, Hospitalizations, and Deaths*; Centers for Disease Control and Prevention: Atlanta, GA, USA, 2006.
3. Selassie, A.W.; Zaloshnja, E.; Langlois, J.A.; Miller, T.; Jones, P.; Steiner, C. Incidence of long-term disability following traumatic brain injury hospitalization in the United States. *J. Head Trauma Rehabil.* **2008**, *23*, 123–131. [CrossRef] [PubMed]
4. Langlois, J.A.; Kegler, S.R.; Butler, J.A.; Gotsch, K.E.; Johnson, R.L.; Reichard, A.A.; Webb, K.W.; Coronado, V.G.; Selassie, A.W.; Thurman, D.J. Traumatic brain injury related hospital discharges: Results from a 14-state surveillance system, 1997. *MMWR Surveill. Summ.* **2003**, *52*, 1–20. [PubMed]
5. Coronado, V.G.; Xu, L.; Basavaraju, S.V.; McGuire, L.C.; Wald, M.M.; Faul, M.D.; Guzman, B.R.; Hemphill, J.D.; Centers for Disease Control and Prevention (CDC). Surveillance for traumatic brain injury-related deaths—United States, 1997–2007. *MMWR Surveill. Summ.* **2011**, *60*, 1–32. [PubMed]
6. Chiu, W.T.; Hung, C.C.; Shih, C.J. Epidemiology of head injury in rural Taiwan—A four year survey. *J. Clin. Neurosci.* **1995**, *2*, 210–215. [CrossRef]
7. Langlois, J.A.; Rutland-Brown, W.; Wald, M.M. The epidemiology and impact of traumatic brain injury: A brief overview. *J. Head Trauma Rehabil.* **2006**, *21*, 375–378. [CrossRef] [PubMed]
8. Andersson, E.H.; Björklund, R.; Emanuelson, I.; Stålhammar, D. Epidemiology of traumatic brain injury: A population based study in western Sweden. *Acta Neurol. Scand.* **2003**, *107*, 256–259. [CrossRef] [PubMed]
9. Baldo, V.; Marcolongo, A.; Floreani, A.; Majori, S.; Cristofolettil, M.; Dal Zotto, A.; Vazzoler, G.; Trivello, R. Epidemiological aspect of traumatic brain injury in Northeast Italy. *Eur. J. Epidemiol.* **2003**, *18*, 1059–1063. [CrossRef] [PubMed]
10. Masson, F.; Vecsey, J.; Salmi, L.R.; Dartigues, J.F.; Erny, P.; Maurette, P. Disability and handicap 5 years after a head injury: A population-based study. *J. Clin. Epidemiol.* **1997**, *50*, 595–601. [CrossRef]
11. Tagliaferri, F.; Compagnone, C.; Korsic, M.; Servadei, F.; Kraus, J. A systematic review of brain injury epidemiology in Europe. *Acta Neurochir. (Wien)* **2006**, *148*, 255–268. [CrossRef] [PubMed]
12. Aenderl, I.; Gashaw, T.; Siebeck, M.; Mutschler, W. Head injury—A neglected public health problem: A four-month prospective study at Jimma University Specialized Hospital, Ethiopia. *Ethiop J. Health Sci.* **2014**, *24*, 27–34. [CrossRef] [PubMed]
13. Andelic, N.I.; Anke, A.; Skandsen, T.; Sigurdardottir, S.; Sandhaug, M.; Ader, T.; Roe, C. Incidence of hospital-admitted severe traumatic brain injury and in-hospital fatality in Norway: A national cohort study. *Neuroepidemiology* **2012**, *38*, 259–267. [CrossRef] [PubMed]
14. Andriessen, T.M.; Horn, J.; Franschman, G.; van der Naalt, J.; Haitsma, I.; Jacobs, B.; Steyeberg, E.W.; Vos, P.E. Epidemiology, severity classification, and outcome of moderate and severe traumatic brain injury: A prospective multicenter study. *J. Neurotrauma* **2003**, *107*, 256–259. [CrossRef] [PubMed]
15. Annoni, J.M.; Beer, S.; Kesselring, J. Severe traumatic brain injury—Epidemiology and outcome after 3 years. *Disabil. Rehabil.* **1992**, *14*, 23–26. [CrossRef] [PubMed]
16. Bruns, J., Jr.; Hauser, W.A. The epidemiology of traumatic brain injury: A review. *Epilepsia* **2003**, *44* (Suppl. 10), 2–10. [CrossRef] [PubMed]
17. Lee, S.T.; Lui, T.N.; Chang, C.N.; Wang, D.J.; Heimburger, R.F.; Fai, H.D. Features of head injury in a developing country—Taiwan (1977–1987). *J. Trauma* **1990**, *30*, 194–199. [CrossRef] [PubMed]

18. Masson, F.; Thicoipe, M.; Aye, P.; Mokni, T.; Senjean, P.; Schmitt, V.; Dessalles, P.H.; Cazauqade, M.; Labadens, P.; Aquitaine Group for Severe Brain Injuries Study. Epidemiology of severe brain injuries: A prospective population-based study. *J. Trauma* **2001**, *51*, 481–489. [CrossRef] [PubMed]
19. Chiu, W.T.; Kuo, C.Y.; Hung, C.C.; Chen, M. The effect of the Taiwan motorcycle helmet use law on head injuries. *Am. J. Public Health* **2000**, *90*, 793–796. [PubMed]
20. Morris, J.A., Jr.; MacKenzie, E.J.; Edelstein, S.L. The effect of preexisting conditions on mortality in trauma patients. *JAMA* **1990**, *263*, 1942–1946. [CrossRef] [PubMed]
21. Charlson, M.E.; Pompei, P.; Ales, K.L.; MacKenzie, C.R. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *J. Chronic Dis.* **1987**, *40*, 373–383. [CrossRef]
22. Baker, S.P.; O'Neill, R.; Haddon, W.; Long, W.B. The injury severity score: A method for describing patients with multiple injuries and evaluating emergency care. *J. Trauma.* **1974**, *14*, 187–196. [CrossRef] [PubMed]
23. Linn, S. The injury severity score—importance and uses. *Ann. Epidemiol.* **1995**, *5*, 440–446. [CrossRef]
24. MacKenzie, E.J.; Steinwachs, D.M.; Shankar, B. Classifying trauma severity based on hospital discharge diagnoses. Validation of an ICD-9CM to AIS-85 conversion table. *Med. Care* **1989**, *27*, 412–422. [CrossRef] [PubMed]
25. Durbin, D.R.; Localio, A.R.; MacKenzie, E.J. Validation of the ICD/AIS MAP for pediatric use. *Inj. Prev.* **2001**, *7*, 96–99. [CrossRef] [PubMed]
26. Mullins, R.J.; Veum-Stone, J.; Helfand, M.; Zimmer-Gembeck, M.; Hedges, J.R.; Southard, P.A.; Trunkey, D.D. Outcome of hospitalized injured patients after institution of a trauma system in an urban area. *JAMA* **1994**, *271*, 1919–1924. [CrossRef] [PubMed]
27. Annest, J.L.; Fingerhut, L.A.; Gallagher, S.S.; Grossman, D.C.; Hedegaard, H.; Johnson, R.L.; Kohn, M.; Pickett, D.; Thomas, K.E.; et al. Strategies to improve external cause-of-injury coding in state-based hospital discharge and emergency department data systems: Recommendations of the CDC workgroup for improvement of external cause-of-injury coding. *MMWR Recomm. Rep.* **2008**, *57(RR-1)*, 1–15.
28. Koskinen, S.; Alaranta, H. Traumatic brain injury in Finland 1991–2005: A nationwide register study of hospitalized and fatal TBI. *Brain Inj.* **2008**, *22*, 205–214. [CrossRef] [PubMed]
29. Tsai, W.H.; Chiu, W.T.; Chiou, H.Y.; Choy, C.S.; Hung, C.C. Pediatric traumatic brain injuries in Taiwan: An 8-year study. *J. Clin. Neurosci.* **2004**, *11*, 126–129. [CrossRef]
30. American College of Surgeons. In *Advanced Trauma Life Support*; American College of Surgeons: Chicago, IL, USA, 2018.
31. Mullins, R.J.; Diggs, B.S.; Hedges, J.R.; Newgard, C.D.; Arthur, M.; Adams, A.L.; Veum-Stone, J.; Lenfesty, B.; Trunkey, D.D. Regional differences in outcomes for hospitalized injured patients. *J. Trauma* **2006**, *60*, 691–698. [CrossRef] [PubMed]
32. Singh, G.K.; Azuine, R.E.; Siahpush, M.; Kogan, M.D. All-cause and cause-specific mortality among US youth: Socioeconomic and rural-urban disparities and international patterns. *J. Urban Health* **2013**, *90*, 388–405. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

Generalized Self-Efficacy, Dispositional Optimism, and Illness Acceptance in Women with Polycystic Ovary Syndrome

Ewa Rzońca [1], Grażyna Iwanowicz-Palus [1], Agnieszka Bień [1,*], Artur Wdowiak [2], Ryszard Szymański [3] and Gustaw Chołubek [2]

1 Faculty of Health Sciences, Medical University of Lublin, 4-6 Staszica St., 20-081 Lublin, Poland; eva.rzonca@gmail.com (E.R.); spupalus@gmail.com (G.I.-P.)
2 Diagnostic Techniques Unit, Medical University of Lublin, 4-6 Staszica St., 20-081 Lublin, Poland; wdowiakartur@gmail.com (A.W.); gustaw.cholubek@umlub.pl (G.C.)
3 Gynecological-obstetrics Ward, Independent Public Complex of Health Care Facilities in Nowa Dęba, 1a M.C. Skłodowska St., 39-460 Nowa Dęba, Poland; rsz18@poczta.fm
* Correspondence: agnesmbien@gmail.com; Tel.: +48-81-448-6840; Fax: +48-81-448-6841

Received: 12 October 2018; Accepted: 4 November 2018; Published: 7 November 2018

Abstract: Polycystic ovary syndrome (PCOS) is one of the most common chronic endocrinopathies affecting between 5 and 10% of reproductive age women. A diagnosis of PCOS very often causes a deterioration in the woman's self-esteem and self-image, and consequently her quality of life (QoL). The purpose of the study was to investigate generalized self-efficacy, dispositional optimism and acceptance of illness in women with PCOS and to explore factors that affect these variables. The study was performed between January and November 2016 among women with PCOS using health care services. The study used a diagnostic survey with questionnaires. The research instruments included the Generalized Self-Efficacy Scale (GSES), the Life Orientation Test-Revised (LOT-R), the Acceptance of Illness Scale (AIS), and a standardized interview questionnaire. Among the PCOS patients studied, the mean score for generalized self-efficacy was 28.74 (±5.16), dispositional optimism—13.56 (±4.28), and acceptance of illness—27.90 (±7.74). The respondents' generalized self-efficacy was determined by their residence, education, socio-economic standing, BMI, and time from diagnosis ($p < 0.05$), while socio-economic standing was a determinant of dispositional optimism ($p < 0.05$). Determinants of illness acceptance in women with PCOS included their residence, socio-economic standing, and time from diagnosis ($p < 0.05$). Increased generalized self-efficacy and dispositional optimism contributed to more illness acceptance in PCOS patients. The present study, compared with others on the subject, will enable specialists providing care to women with PCOS to gain a deeper and more comprehensive understanding of the situation and condition of their patients. It will also allow for a better response to the needs of PCOS patients, and provide them with individualized, holistic specialist care, diagnostics, and treatment.

Keywords: polycystic ovary syndrome; self-efficacy; optimism; acceptance of illness

1. Introduction

Every person reacts differently to the diagnosis of a chronic illness, in accordance with their individual characteristics. Any illness affects the psychological state and functioning of the person diagnosed with it, and chronic illnesses are very often associated with a number of adaptations and changes in a person's life and activities [1–3]. Significant factors affecting the acceptance of illness include a person's belief in their capacity to adapt, their self-esteem, self-efficacy, optimism, and socio-demographic factors [1–4].

Polycystic ovary syndrome (PCOS) is one of the most common chronic endocrinopathies affecting between 5 and 10% of reproductive age women. The syndrome presents a variety of symptoms, including hyperandrogenism, menstrual disorders, obesity, metabolic syndrome, type 2 diabetes mellitus, fertility disorders, or emotional and psychological disorders, while its etiology has not yet been fully understood [5–7]. Polycystic ovary syndrome is most commonly diagnosed based on the Rotterdam criteria, i.e., when two out of the following three findings are present: oligomenorrhea or oligo-ovulation, clinical and/or biochemical hyperandrogenism, and/or polycystic ovaries on ultrasound [7]. A diagnosis of PCOS very often causes a deterioration in the woman's self-esteem and self-image, and consequently her quality of life (QoL) [8,9]. Polycystic ovary syndrome presents a broad variety of symptoms, including metabolic, endocrine, and emotional or psychological ones, which is extremely challenging to women diagnosed with the disorder [5,7,10]. Numerous studies on PCOS focus on its biological and physiological aspects, though a growing number of publications discuss the adverse impact of PCOS on women's lives and how women function. These publications emphasize the fact that PCOS patients are at a higher risk of depression, anxiety, and low QoL [5–13]. Thus, women with PCOS, faced with the challenge of such a multidimensional disorder, must cope with their new situation. This may be affected by multiple factors, including the patient's individual predispositions [6,8–13].

Individual capabilities and predispositions, including optimism or self-efficacy, have an impact on a person's acceptance of their situation, including the situation of being ill, and this motivated the authors to tackle this subject in the present paper. The purpose of the study was to investigate generalized self-efficacy, dispositional optimism and acceptance of illness in women with PCOS and to explore factors that affect these variables.

2. Materials and Methods

2.1. Participants

The cross-sectional study was performed between January and November 2016 in PCOS patients using health care services (primary care, specialist outpatient care, and inpatient/hospital care) in four regions of Poland: Lubelskie, Podkarpackie, Pomorskie, and Wielkopolskie provinces. Inclusion criteria for cases were as follows: Age over 18 years, PCOS diagnosis made based on the Rotterdam criteria, and use of health care services in Poland. Exclusion criteria were: Cancer and psychological disorders. The study was performed in accordance with the Helsinki Declaration, and approved by the Lublin Medical University Bioethics Committee (approval no. KE-0254/189/2015). It was also approved by the respective health care units. Respondents were informed that participation was voluntary, and that study results were anonymous and to be used exclusively for research purposes. Out of the 300 survey questionnaires distributed to respondents, 250 correctly completed questionnaires were analyzed. Fifty questionnaires were excluded from further analysis due to incomplete data. The data effectiveness rate was 83.33%.

2.2. Assessments

The study used a diagnostic survey with questionnaires. Research instruments included the Generalized Self-Efficacy Scale (GSES), the Life Orientation Test-Revised (LOT-R), the Acceptance of Illness Scale (AIS), and an own sociodemographic questionnaire to collect data on respondents' characteristics (age, residence, marital status, education, professional activity, socio-economic standing—self-reported, having children, BMI—collected from respondents' medical records, time from PCOS diagnosis).

The GSES was developed by Schwarzer and Jerusalem, and adapted for Polish settings by Juczyński. It measures an individual's sense of generalized self-efficacy in difficult situations. The scale comprises 10 statements, which are scored by the respondent as follows: 1—not true at all, 2—hardly true, 3—moderately true, 4—exactly true. The total score ranges between 10 and 40 points, representing a person's level of self-efficacy. Higher scores indicate a stronger sense of self-efficacy. Scores should be interpreted using sten ranges, with scores of sten 1–4 considered low, and sten 7–10 considered high. The scale has good psychometric properties, with a Cronbach's α of 0.85 [14].

The LOT-R was authored by Scheier, Carver, and Bridges, and adapted for Polish settings by Poprawa and Juczyński. It measures dispositional optimism based on 10 statements rated 0–4, where 0 stands for "I disagree a lot", while 4 stands for "I agree a lot". Total score ranges between 0 and 24 points, with higher scores denoting more optimism. The raw score is converted into sten results, with scores of sten 1–4 considered low, i.e., indicating pessimistic tendencies, while sten 7–10 scores are considered high, i.e., indicating an optimistic attitude. The reliability of the LOT-R as measured by Cronbach's α is 0.76 [15].

The AIS was developed by Felton et al., and adapted for Polish settings by Juczyński. It allows for measuring illness acceptance in adults, using 8 statements concerning the negative consequences of poor health, rated by the respondents using a 5-item scale, with 1 standing for "strongly agree", and 5 for "strongly disagree". The total, between 8 and 40 points, is a measure of illness acceptance. Low scores indicate a lack of acceptance, poor adaptation to the illness, and significant emotional problems associated with the illness, while high scores indicate a respondent's acceptance of their condition. Internal consistency as measured by Cronbach's α is 0.85 for the original version of the AIS, and very similar for the Polish version, with an α of 0.82 [16].

2.3. Statistical Analyses

Statistical analysis of data from the questionnaires was performed using STATISTICA version 12 (StatSoft, Kraków, Poland). Qualitative data are described by numbers (n) and percentages (%), and quantitative data are described by means (M), standard deviations (SD) and medians (Me). To analyze the variables affecting GSES, LOT-R, and AIS scores, standard multiple regression analysis was performed. With this method, results are interpreted by comparing β values, indicating the direction and strength of relationships between predictors. The fit of the regression model to the data is indicated by the corrected R^2 value. Reference categories were provided for nominal and ordinal variables: For age it was under 25 years old, for residence—urban: Province capital, for education—college/university, and for time from PCOS diagnosis—less than 1 year. As to BMI, it was included in the analysis as a scale variable. If a variable comprised two categories, the analysis involved comparing the two categories, e.g., in the case of marital status—single, professional activity—not working, socio-economic status—unsatisfactory, and for having children—none. Correlations between selected variables were tested using Pearson's r; correlation strength was evaluated using Guilford's classification. The study used a significance threshold of $p < 0.05$.

3. Results

The study included 250 women with PCOS. Most respondents were women aged 26–35 years (43.20%), lived in urban: Province capital (38.40%), married (64.40%), college/university-educated (59.60%), professionally active (77.20%), who viewed their socio-economic standing as satisfactory (62.40%), had no children (53.20%), mean BMI 26.41, and had been diagnosed with PCOS 1 to 5 years before (50.40%) (Table 1).

Table 2 shows the mean scores for generalized self-efficacy (28.74 $\pm$ 5.16), dispositional optimism (13.56 $\pm$ 4.28), and acceptance of illness (27.90 $\pm$ 7.74).

The regression model for variables generalized self-efficacy (GSES) is shown in Table 3. The regression model that was developed accounts for 18.8% of variance for the GSES variable ($F(15.234) = 4.834$; $p < 0.001$). Statistically significant predictors for the GSES regression model included residence—urban: Other ($\beta = -0.140$; $p = 0.033$), education—primary/vocational ($\beta = -0.137$; $p = 0.033$), socio-economic standing ($\beta = 0.275$; $p = 0.001$), BMI ($\beta = -0.152$; $p = 0.012$) and time from PCOS diagnosis—1–5 years ($\beta = -0.233$; $p = 0.010$) and 6–10 years ($\beta = -0.197$; $p = 0.023$). Living in a city other than a province capital, having finished one's education at the primary/vocational level, having a higher BMI, and a 1 to 10 years from PCOS diagnosis (1–5 and 6–10 years) were all associated with a lower sense of self-efficacy in the women studied. Better socio-economic status predicted a stronger sense of self-efficacy.

Table 1. Participants' characteristics.

Participants' Characteristics		*n*	%
Age	Under 25 years old	47	18.80
	26–35	108	43.20
	Over 35 years old	95	38.00
Residence	Urban: Province capital	96	38.40
	Urban: Other	74	29.60
	Rural	80	32.00
Marital status	Single	89	35.60
	Married	161	64.40
Education	Primary/vocational	22	8.80
	High school	79	31.60
	College/university	149	59.60
Professional activity	Working professionally	193	77.20
	Not working	57	22.80
Socio-economic standing	Unsatisfactory	94	37.60
	Satisfactory	156	62.40
Having children	No children	133	53.20
	Children	117	44.80
Body Mass Index (BMI)	Mean BMI	26.41	
Time from PCOS diagnosis	Up to 1 year	38	15.20
	1–5 years	126	50.40
	6–10 years	52	20.80
	more than 10 years	34	13.60

Table 2. Mean scores for generalized self-efficacy (GSES), dispositional optimism (LOT-R), and illness acceptance (AIS) in the polycystic ovary syndrome (PCOS) patients studied.

	M	Me	SD
GSES	28.74	30	5.16
LOT-R	13.56	13	4.28
AIS	27.90	28	7.74

M—mean; SD—standard deviation; Me—median.

Table 3. Regression model for variables GSES.

Variables	GSES $R^2 = 0.188$ $F(15.234) = 4.834$ $p < 0.001$			
	B	β	t	p
Age: 26–35	0.023	0.002	0.024	0.981
Age: Over 35 y/o	−0.090	−0.008	−0.076	0.939
Residence: Urban—other	−1.576	−0.140	−2.150	0.033
Residence: Rural	−1.216	−0.110	−1.665	0.097
Marital status	0.315	0.029	0.438	0.662
Education: Primary/vocational	−2.487	−0.137	−2.148	0.033
Education: High school	−0.001	0.000	−0.002	0.999
Professional activity	−0.676	−0.055	−0.904	0.367
Socio-economic standing	2.921	0.275	4.561	0.001
Having children	0.103	0.010	0.127	0.899
BMI	−0.169	−0.152	−2.541	0.012
Time from PCOS diagnosis: 1–5 years	−2.395	−0.233	−2.599	0.010
Time from PCOS diagnosis: 6–10 years	−2.504	−0.197	−2.282	0.023
Time from PCOS diagnosis: More than 10 years	0.319	0.021	0.246	0.806

Table 4 shows regression model for variables LOT-R. The model was found to fit the data well ($F(15.234) = 2.427$; $p = 0.003$), while it only accounted for 7.9% of variance in the LOT-R variable. For the LOT-R variable, the only statistically significant variable was socio-economic standing ($\beta = 0.259$; $p = 0.001$)—the better the socio-economic status, the higher the optimism score.

Table 4. Regression model for variables LOT-R.

Variables	LOT-R $R^2 = 0.079$ $F(15.234) = 2.427$ $p = 0.003$			
	B	β	t	p
Age: 26–35	0.009	0.001	0.010	0.992
Age: Over 35 y/o	0.016	0.002	0.015	0.988
Residence: Urban—other	−0.851	−0.091	−1.313	0.190
Residence: Rural	−0.572	−0.062	−0.886	0.377
Marital status	−0.192	−0.021	−0.303	0.762
Education: Primary/vocational	−1.506	−0.100	−1.472	0.142
Education: High school	−0.167	−0.018	−0.271	0.787
Professional activity	−0.855	−0.084	−1.293	0.197
Socio-economic standing	2.284	0.259	4.034	0.001
Having children	0.489	0.057	0.691	0.490
BMI	−0.040	−0.044	−0.688	0.492
Time from PCOS diagnosis: 1–5 years	1.253	0.147	1.538	0.125
Time from PCOS diagnosis: 6–10 years	1.324	0.126	1.366	0.173
Time from PCOS diagnosis: More than 10 years	2.025	0.162	1.763	0.079

The regression model for AIS is shown in Table 5. This regression model accounts for 20.7% of variance for the AIS variable ($F(115.234) = 5.340$; $p < 0.001$). For the AIS variable model, statistically significant predictors included residence—urban: Other ($\beta = -0.134$; $p = 0.037$), socio-economic standing ($\beta = 0.335$; $p = 0.001$), and time from PCOS diagnosis—1–5 years ($\beta = -0.263$; $p = 0.003$). Living in a city other than a province capital and time from PCOS diagnosis between 1 and 5 years were associated with poorer acceptance of illness by the PCOS patients studied, while better socio-economic standing predicted more PCOS acceptance.

Table 5. Regression model for variables AIS.

Variables	AIS $R^2 = 0.207$ $F(15.234) = 5.340$ $p < 0.001$			
	B	β	t	p
Age: 26–35	1.292	0.083	0.885	0.377
Age: Over 35 y/o	−1.914	−0.120	−1.096	0.274
Residence: Urban—other	−2.272	−0.134	−2.092	0.037
Residence: Rural	−1.918	−0.116	−1.772	0.078
Marital status	1.722	0.107	1.617	0.107
Education: Primary/vocational	−3.145	−0.115	−1.834	0.068
Education: High school	1.347	0.081	1.302	0.194
Professional activity	−0.168	−0.009	−0.152	0.880
Socio-economic standing	5.335	0.335	5.623	0.001
Having children	0.758	0.049	0.628	0.530
BMI	−0.077	−0.046	−0.777	0.438
Time from PCOS diagnosis: 1–5 years	−4.056	−0.263	−2.971	0.003
Time from PCOS diagnosis: 6–10 years	−2.169	−0.114	−1.334	0.183
Time from PCOS diagnosis: More than 10 years	−0.016	0.000	−0.009	0.993

Statistical analyses showed significant positive correlations between generalized self-efficacy on the one hand, and dispositional optimism and illness acceptance on the other, as well as between dispositional optimism and illness acceptance ($p < 0.001$). The strength of correlations was rated between 0.434 and 0.610 (Table 6).

Table 6. Correlations between GSES, LOT-R, and AIS in the PCOS patients studied.

	GSES	LOT-R	AIS
GSES	– –	$r = 0.474$ $p < 0.001$	$r = 0.610$ $p < 0.001$
LOT-R	$r = 0.474$ $p < 0.001$	– –	$r = 0.434$ $p < 0.001$
AIS	$r = 0.610$ $p < 0.001$	$r = 0.434$ $p < 0.001$	– –

GSES—generalized self-efficacy; LOT-R—dispositional optimism; AIS—illness acceptance.

4. Discussion

Illness, especially a chronic one such as diabetes, cardiovascular disease, or thyroid disease, is a major challenge for the person diagnosed with it. It has a considerable impact on the patient's life and functioning, which is associated with multiple and diverse factors, including concern about one's health and life, possible complications, impact of the disease on family or professional life, self-efficacy, optimism, and a number of other social, demographic, and economic considerations [2,10,17–19]. Polycystic ovary syndrome matches the above description of the impact illness has on the patient's life, as the literature emphasizes that a diagnosis of PCOS affects the whole of a woman's life. Symptoms of PCOS first manifest during puberty or in early adulthood, but its complications or sequelae have a negative impact on women's functioning and QoL at subsequent stages as well, including menopause or even old age [5,7,10,20–23]. Therefore, the authors of the present paper attempted to evaluate generalized self-efficacy, dispositional optimism, and illness acceptance in women with PCOS, as well as the determinants of these characteristics.

Self-efficacy is strongly associated with an ability to control one's actions so as to achieve the desired results, even in unfavorable circumstances. It reflects a person's beliefs about their capability to change their behavior when diagnosed with an illness or to prevent potential health problems [24–26]. The author of the Polish adaptation of GSES reported mean scores for Polish populations, including a population of women who had undergone a mastectomy—30.07 [14]. In Polish hysterectomy patients, a self-efficacy score of 32.46 was reported [1], in University of the Third Age students—a mean of 30.12 [25], and in patients with heart failure—32 [27]. These data indicate a strong sense of self-efficacy in the various groups [1,14,25,27]. As to the present study, it demonstrated a moderate level of self-efficacy in PCOS patients (28.74), similar to that found in diabetics (28.34) or in women with migraines (28.57) [14].

Kozica et al. (2013) studied self-efficacy and self-care in women with and without PCOS, demonstrating that significant predictors of self-efficacy in their respondents included health vigilance, overall health history, and diagnosed infertility. The authors state that women with PCOS should be aware that their illness is a chronic condition that requires them to carefully manage their health by introducing lifestyle modifications and undergoing regular screening [24]. Rogala et al. (2015) investigated the relationship between mental adaptation to cancer and self-efficacy in women who had undergone hysterectomies due to cancer. Their analysis included socio-demographic factors, such as education, residence, relationship status, and economic standing, which could affect the respondents' self-efficacy. The only statistically significant variable was the patients' relationship status, as women in a relationship had a higher level of self-efficacy than single ones [1]. In a study by Zielińska-Więczkowska (2017), concerning the relationships between satisfaction with life and selected personal resources in University of the Third Age students, self-efficacy was associated with education, financial standing, and perceived health. Lower levels of education were associated with poorer self-efficacy, while a better financial standing or better perceived health was related to higher GSES scores [25]. In the present study, lower levels of self-efficacy were found in PCOS patients who lived in cities other than province capitals, had had a primary/vocational education, had higher BMI values, and had been living with PCOS for a longer time (1–10 years), while higher levels were found in women with a better socio-economic standing.

The present study also analyzed another important individual characteristic or tendency, namely optimism, defined as an overall expectation of positive rather than negative life experiences in the future. This trait is conducive to motivation, perseverance, and determination in pursuing and achieving goals [2,28–30]. The literature on the subject provides a highly inconsistent view on dispositional optimism and its predictors in various patient groups studied in the world, depending on their health status [2,25,28–31]. In their study on the relationship between dispositional optimism and acceptance of illness in a group of patients with Graves' disease, Basińska et al. (2008) found comorbidity to be the only variable significantly affecting the patients' optimism levels. Those diagnosed with other diseases beside Graves' had lower levels of dispositional optimism than those with no comorbidities [2]. A study by Chung et al. (2016) showed that, in stroke patients, there was a negative correlation between optimism and depression symptoms—lower optimism levels were associated with more severe symptoms of depression [28]. Bleil et al. (2012), in turn, found that a more pessimistic attitude may contribute to the failure of fertility treatment [29]. Moyer et al. (2009) analyzed optimism/pessimism and health-related quality of life (HRQoL) in pregnant women from China, Ghana, and the US. Optimism levels were found to be associated with the women's country of origin, professional activity, education, and treatment for emotional disorders [30]. Zielińska-Więczkowska (2017) demonstrated that financial standing had a significant impact on dispositional optimism in the University of the Third Age students she studied [25]. The present results corroborate those reported by this author. Satisfaction with one's socio-economic standing was associated with a higher level of dispositional optimism, which overall was moderate in the PCOS patients studied.

Acceptance of illness is a very important stage in the life of a person diagnosed with a disease, especially a chronic one. It allows the patient to adapt to their new situation, and higher levels of illness acceptance contribute to better coping with the illness, and ultimately, to a better QoL [2,32,33]. In the study on Graves' patients by Basińska et al. (2008), the reported acceptance of illness score was 28.48 [2]. In turn, Bień et al. (2016) reported a score of 30.66 in a group of patients with gestational diabetes mellitus (GDM) [32]. A comparison between the cited scores and the mean score of 27.90 found in the present study demonstrates that PCOS patients have a lower level of illness acceptance than other populations. This finding may be explained by the complexity and variety of PCOS symptoms, which significantly affect the life of patients, as well as their self-perception in relation to the expectations of the contemporary world [6,8,9,24,34].

The multiple factors affecting acceptance of illness also have an impact on a person's adaptation to the illness, which occurs to various degrees, as shown in the literature on the subject, as well as in another part of the present study [2,32,33,35]. In a study by Jankowska-Polańska et al. (2016), determinants of illness acceptance in patients with chronic obstructive pulmonary disease (COPD) included age, education, and duration of illness. Better illness acceptance was found in younger COPD patients, in those with college/university education, and with a shorter duration of illness [35]. Bień et al. (2016) demonstrated that acceptance of illness in women with GDM improved with very good financial standing, dietary treatment of diabetes, and very good overall perceived health [32]. In turn, Kostyła et al. (2013), in a study on patients with psoriasis, found acceptance of illness to increase as psycho-pathological symptom intensity decreased [33]. In the present study, PCOS patients who lived in cities other than province capitals and had been diagnosed with PCOS between 1 and 5 years before were found to have lower levels of illness acceptance. Better socio-economic standing was associated with greater acceptance of PCOS in the group studied.

The findings of the present study demonstrate a single common determinant of all the aspects analyzed, including self-efficacy, dispositional optimism, and acceptance of illness: The PCOS patients' socio-economic standing. The literature on the subject emphasizes the importance of financial situation for health, as it affects access to health care in general and to specific methods of treatment, the ability to maintain a healthy lifestyle, as well as how the patients manage on a daily basis and their family life [19,25,32,36].

The final stage of the present study involved an analysis of correlations between self-efficacy, dispositional optimism, and acceptance of illness in women with PCOS. Popa-Velea & Purcarea (2014), investigating the psychological factors affecting HRQoL in COPD patients, demonstrated that self-efficacy and optimism were positively correlated with the patients' HRQoL. Moreover, patients with low levels of self-efficacy and optimism may experience more problems and discomfort than those with higher levels of these traits in the case of a comparable decrease in lung function [37]. Zielińska-Więczkowska (2017) found satisfaction with life to be positively correlated both with generalized self-efficacy and with dispositional optimism [25]. Basińska et al. (2008) demonstrated an association between dispositional optimism and acceptance of illness in their group of patients with Graves' disease [2]. As to the present study, a stronger sense of self-efficacy was associated with both greater optimism and more acceptance of PCOS, and similarly, higher levels of optimism were associated with greater illness acceptance. This warrants the conclusion that acceptance of illness is the product of an individual's personal resources that allow them to come to terms with their illness and to adapt to it, so as to live and function better with the illness, as evidenced both by literature reports and by the present study [2,25,37].

So far, the authors have found no studies on the level and determinants of generalized self-efficacy, dispositional optimism, and illness acceptance among women with PCOS in the available literature. Therefore, the study is an original one. The present study has certain limitations. The study concept did not include an analysis of the impact of PCOS symptoms (including hirsutism or menstrual disorders) or psychological condition on the patient's life, though these analyses were included in a number of studies worldwide, providing information and direction to the authors of the present paper. Further studies are still warranted on PCOS, its clinical circumstances, etiology, impact on patients' lives, and long-term consequences.

5. Conclusions

Women with PCOS have moderate levels of generalized self-efficacy, dispositional optimism, and illness acceptance. The patients' generalized self-efficacy is determined by their residence, education, socio-economic standing, BMI, and time from diagnosis, while socio-economic standing was the only determinant of dispositional optimism. Determinants of illness acceptance in women with PCOS included their residence, socio-economic standing, and time from diagnosis. Increased generalized self-efficacy and dispositional optimism contributed to greater illness acceptance in PCOS patients.

The present study, compared with others on the subject, will enable specialists providing care to women with PCOS to gain a deeper and more comprehensive understanding of the situation and condition of their patients. It will also allow for a better response to the needs of PCOS patients, and provide them with individualized, holistic specialist care, diagnostics, and treatment.

Author Contributions: E.R. designed the study, collected data, analyzed the statistics, interpreted the data, searched the literature, wrote the paper, and performed the manuscript review; G.I.-P. designed the study, interpreted the data, and carried out a critical review of the manuscript; A.B. analyzed the statistics, overviewed and analyzed the literature, and wrote the paper; A.W. and R.S. collected and analyzed the research material; G.C.: interpreted the data, and searched the literature. All authors have read and approved the final manuscript.

Funding: This research received no external funding.

Acknowledgments: We would like to express our deepest gratitude to the heads of health care services (primary care, specialist outpatient care, and inpatient/hospital care) that hosted the study, and to all the women who agreed to participate.

Conflicts of Interest: The authors declare no conflict of interest. Authors declare to have no financial or personal relationships with other people or organizations that could inappropriately influence (bias) the results presented in this manuscript.

References

1. Rogala, D.; Mazur, A.; Maślińska, M.; Koper, K.; Wysocka, J. Self-efficacy and strategies of adaptation to disease in patients with cancer of reproductive organs. *Curr. Gynecol. Oncol.* **2015**, *13*, 154–164. [CrossRef]

2. Basińska, A.M.; Zalewska-Rydzkowska, D.; Wolańska, P.; Junik, R. Dispositional optimism and acceptance of illness among a group of individuals with Graves-Basedow's disease. *Pol. J. Endocrinol.* **2008**, *59*, 23–28. (In Polish)
3. Basinska, A.; Andruszkiewicz, A. Health Locus of Control in Patients with Graves-Basedow Disease and Hashimoto Disease and Their Acceptance of Illness. *Int. J. Endocrinol. Metab.* **2012**, *10*, 537–542. [CrossRef] [PubMed]
4. Lipińska-Grobelny, A. Effects of gender role on personal resources and coping with stress. *Int. J. Occup. Med. Environ. Health* **2011**, *24*, 18–28. [CrossRef] [PubMed]
5. De Leo, V.; Musacchio, M.C.; Cappelli, V.; Massaro, M.G.; Morgante, G.; Petraglia, F. Genetic hormonal and metabolic aspects of PCOS: An update. *Reprod. Biol. Endocrinol.* **2016**, *14*. [CrossRef] [PubMed]
6. Elsenbruch, S.; Benson, S.; Hahn, S.; Tan, S.; Mann, K.; Pleger, K.; Kimmig, R.; Janssen, O.E. Determinants of emotional distress in women with polycystic ovary syndrome. *Hum. Reprod.* **2006**, *21*, 1092–1099. [CrossRef] [PubMed]
7. El Hayek, S.; Bitar, L.; Hamdar, L.H.; Mirza, F.G.; Daoud, G. Poly Cystic Ovarian Syndrome: An Updated Overview. *Front. Physiol.* **2016**, *5*. [CrossRef] [PubMed]
8. Bazarganipour, F.; Ziaei, S.; Montazeri, A.; Foroozanfard, F.; Kazemnejad, A.; Faghihzadeh, S. Health-Related Quality of Life in Patients with Polycystic Ovary Syndrome (PCOS): A Model-Based Study of Predictive Factors. *J. Sex. Med.* **2014**, *11*, 1023–1032. [CrossRef] [PubMed]
9. Bazarganipour, F.; Ziaei, S.; Montazeri, A.; Foroozanfard, F.; Kazemnejad, A.; Faghihzadeh, S. Health-related quality of life and its relationship with clinical symptoms among Iranian patients with polycystic ovarian syndrome. *Iran. J. Reprod. Med.* **2013**, *11*, 371–378. [PubMed]
10. Teede, H.; Deeks, A.; Moran, L. Polycystic ovary syndrome: A complex condition with psychological, reproductive and metabolic manifestations that impacts on health across the lifespan. *BMC Med.* **2010**, *8*, 41. [CrossRef] [PubMed]
11. Veltman-Verhulst, S.M.; Boivin, J.; Eijkemans, M.J.; Fauser, B.J. Emotional distress is a common risk in women with polycystic ovary syndrome: A systematic review and meta-analysis of 28 studies. *Hum. Reprod. Update* **2012**, *18*, 638–651. [CrossRef] [PubMed]
12. Kolahi, L.; Asemi, N.; Mirzaei, M.; Adibi, N.; Beiraghdar, M.; Mehr, A.M. The relationship between quality of life and coping strategies in polycystic ovary syndrome patients. *Adv. Biomed. Res.* **2015**, *4*, 168. [PubMed]
13. Cooney, L.G.; Lee, I.; Sammel, M.D.; Dokras, A. High prevalence of moderate and severe depressive and anxiety symptoms in polycystic ovary syndrome: A systematic review and meta-analysis. *Hum. Reprod.* **2017**, *32*, 1075–1091. [CrossRef] [PubMed]
14. Schwarzer, R.; Jerusalem, M.; Juczyński, Z. Skala Uogólnionej Własnej Skuteczności—GSES. In *Measurement Instruments in Health Promotion and Psychooncology (Narzędzia Pomiaru w Promocji i Psychologii Zdrowia)*; Juczyński, Z., Ed.; Pracowania Testów Psychologicznych Polskiego Towarzystwa Psychologicznego: Warszawa, Poland, 2012; pp. 89–94.
15. Scheier, M.F.; Carver, C.S.; Bridges, M.W. (adaptation: Poprawa, R.; Juczyński, Z.); Test Orientacji Życiowej—LOT-R. In *Measurement Instruments in Health Promotion and Psychooncology (Narzędzia Pomiaru w Promocji i Psychologii Zdrowia)*; Juczyński, Z., Ed.; Pracowania Testów Psychologicznych Polskiego Towarzystwa Psychologicznego: Warszawa, Poland, 2012; pp. 61–66.
16. Felton, B.J.; Revenson, T.A.; Hinrichsen, G.A. (adaptation: Juczyński, Z.); Skala Akceptacji Choroby—AIS. In *Measurement Instruments in Health Promotion and Psychooncology (Narzędzia Pomiaru w Promocji i Psychologii Zdrowia)*; Juczyński, Z., Ed.; Pracowania Testów Psychologicznych Polskiego Towarzystwa Psychologicznego: Warszawa, Poland, 2012; pp. 162–166.
17. Hamilton, J.G.; Lobel, M. Psychosocial Factors Associated with Risk Perceptions for Chronic Diseases in Younger and Middle-Aged Women. *Women Health* **2015**, *55*, 921–942. [CrossRef] [PubMed]
18. Papaspurou, M.; Laschou, V.C.; Partsiopoulou, P.; Fradelos, E.C.; Kleisiaris, C.F.; Kalota, M.A.; Neroliatsiou, A.; Papathanasiou, I.V. Fears and Health Needs of Patients with Diabetes: A Qualitative Research in Rural Population. *Med. Arch.* **2015**, *69*, 190–195. [CrossRef] [PubMed]
19. Viego, V.; Temporelli, K. Socioeconomic status and self-reported chronic diseases among Argentina's adult population: Results based on multivariate probability models. *J. Public Health Res.* **2017**, *6*, 56–63. [CrossRef] [PubMed]

20. Moran, L.J.; Deeks, A.A.; Gibson-Helm, M.E.; Teede, H.J. Psychological parameters in the reproductive phenotypes of polycystic ovary syndrome. *Hum. Reprod.* **2012**, *27*, 2082–2088. [CrossRef] [PubMed]
21. Amini, L.; Valian, K.; Sdeghi Avvalshahr, H.; Montaeri, A. Self-Confidence in Women with and without Polycystic Ovary Syndrome. *J. Fam. Reprod. Health* **2014**, *8*, 113–116.
22. Bates, G.W.; Lergo, R.S. Longterm management of Polycystic Ovarian Syndrome (PCOS). *Mol. Cell. Endocrinol.* **2013**, *373*, 91–97. [CrossRef] [PubMed]
23. Avery, J.C.; Braunack-Mayer, A.J. The information needs of women diagnosed with Polycystic Ovarian Syndrome—Implications for treatment and health outcomes. *BMC Womens Health* **2007**, *20*, 9. [CrossRef] [PubMed]
24. Kozica, S.L.; Gibson-Helm, M.E.; Teede, H.J.; Moran, L.J. Assessing self-efficacy and self-help methods in women with and without Polycystic Ovary Syndrome. *Behav. Med.* **2013**, *39*, 90–96. [CrossRef] [PubMed]
25. Zielińska-Więczkowska, H. Correlations between satisfaction with life and selected personal resources among students of Universities of the Third Age. *Clin. Interv. Aging* **2017**, *31*, 1391–1399. [CrossRef] [PubMed]
26. Carpenter, D.M.; Blalock, S.J.; Sayner, R.; Muir, K.W.; Robin, A.L.; Hartnett, M.E.; Giangiacomo, A.L.; Tudor, G.E.; Sleath, B.L. Communication Predicts Medication Self-Efficacy in Glaucoma Patients. *Optom. Vis. Sci.* **2016**, *93*, 731–737. [CrossRef] [PubMed]
27. Rydlewska, A.; Krzysztofik, J.; Libergal, J.; Rybak, A.; Banasiak, W.; Ponikowski, P.; Jankowska, E.A. Health locus of control and the sense of self-efficacy in patients with systolic heart failure: A pilot study. *Patient Prefer. Adherence* **2013**, *19*, 337–343. [CrossRef] [PubMed]
28. Chung, M.L.; Bakas, T.; Plue, L.D.; Williams, L.S. Effects of Self-esteem, Optimism, and Perceived Control on Depressive Symptoms in Stroke Survivor-Spouse Dyads. *J. Cardiovasc. Nurs.* **2016**, *31*, E8–E16. [CrossRef] [PubMed]
29. Bleil, M.E.; Pasch, L.A.; Gregorich, S.E.; Millstein, S.G.; Katz, P.P.; Adler, N.E. Infertility Outcomes Program Project Group.: Fertility treatment response: Is it better to be more optimistic or less pessimistic? *Psychosom. Med.* **2012**, *74*, 193–199. [CrossRef] [PubMed]
30. Moyer, C.A.; Yang, H.; Kwawukume, Y.; Gupta, A.; Zhu, Y.; Koranteng, I.; Elsayed, Y.; Wei, Y.; Greene, J.; Calhoun, C.; et al. Optimism/pessimism and health-related quality of life during pregnancy across three continents: A matched cohort study in China, Ghana, and the United States. *BMC Pregnancy Childbirth* **2009**, *1*. [CrossRef] [PubMed]
31. Moyer, C.A.; Elsayed, Y.; Zhu, Y.; Wei, Y.; Engmann, C.M.; Yang, H. Is generalized maternal optimism or pessimism during pregnancy associated with unplanned cesarean section deliveries in China? *J. Pregnancy* **2010**, 754938. [CrossRef] [PubMed]
32. Bień, A.; Rzońca, E.; Kańczugowska, A.; Iwanowicz-Palus, G. Factors Affecting the Quality of Life and the Illness Acceptance of Pregnant Women with Diabetes. *Int. J. Environ. Res. Public Health* **2015**, *22*. [CrossRef] [PubMed]
33. Kostyła, M.; Tabała, K.; Kocur, J. Illness acceptance degree versus intensity of psychopathological symptoms in patients with psoriasis. *Postepy Dermatol. Alergol.* **2013**, *3*, 134–139. [CrossRef] [PubMed]
34. McCook, J.G.; Reame, N.E.; Thatcher, S.S. Health-related quality of life issues in women with polycystic ovary syndrome. *J. Obstet. Gynecol. Neonatal. Nurs.* **2005**, *34*, 12–20. [CrossRef] [PubMed]
35. Jankowska-Polańska, B.; Kasprzyk, M.; Chudiak, A.; Uchmanowicz, I. Relation between illness acceptance and quality of life in patients with chronic obstructive pulmonary disease (COPD). *Pneumonol. Alergol. Pol.* **2016**, *84*, 3–10. [CrossRef] [PubMed]
36. Fort, M.P.; Alvarado-Molina, N.; Peña, L.; Mendoza, M.C.; Murrillo, S.; Martínez, H. Barriers and facilitating factors for disease self-management: A qualitative analysis of perceptions of patients receiving care for type 2 diabetes and/or hypertension in San José, Costa Rica and Tuxtla Gutiérrez, Mexico. *BMC Fam. Pract.* **2013**, *14*. [CrossRef] [PubMed]
37. Popa-Velea, O.; Purcarea, V.L. Psychological factors mediating health-related quality of life in COPD. *J. Med. Life* **2014**, *7*, 100–103. [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

Aircraft-Assisted Pilot Suicides in the General Aviation Increased for One-Year Period after 11 September 2001 Attack in the United States

Alpo Vuorio [1,2,*,†], Tanja Laukkala [3,†], Ilkka Junttila [4], Robert Bor [5,6], Bruce Budowle [7], Eero Pukkala [8], Pooshan Navathe [9] and Antti Sajantila [2]

1 Mehiläinen Airport Health Centre, 01530 Vantaa, Finland
2 Department of Forensic Medicine, University of Helsinki, 00014 Helsinki, Finland; antti.sajantila@helsinki.fi
3 Mehiläinen Kielotie Health Centre, 01300 Vantaa, Finland; tanja.laukkala@duodecim.fi
4 Faculty of Medicine and Life Sciences, University of Tampere, Finland and Fimlab Laboratories, 33014 Tampere, Finland; ilkka.junttila@uta.fi
5 Royal Free Hospital, Pond Street, London NW3 2QG, UK; robertbor@hotmail.com
6 Centre for Aviation Psychology, London NW3 1ND, UK
7 Center for Human Identification, University of North Texas Health Science Center, 3500 Camp, Bowie Blvd, Fort Worth, TX 76107, USA; bruce.budowle@unthsc.edu
8 Faculty of Social Sciences, University of Tampere, 33014 Tampere, Finland; Eero.Pukkala@cancer.fi
9 The Maitland Hospital, Maitland 2320, Australia; Pooshan.Navathe@hnehealth.nsw.gov.au
* Correspondence: alpo.vuorio@gmail.com; Tel.: +010-414-2702
† These authors contributed equally to this work.

Received: 13 September 2018; Accepted: 6 November 2018; Published: 12 November 2018

Abstract: Pilot aircraft-assisted suicides (AAS) are rare, and there is limited understanding of copycat phenomenon among aviators. The aim of this study was to evaluate the possible effect the 11 September 2001, terrorist attacks had on pilot AASs in the U.S. Fatal aviation accidents in the National Transportation Safety Board (NTSB) database were searched using the following search words: "suicide", "murder-suicide" and "homicide-suicide". The timeline between 11 September 1996, and 11 September 2004, was analyzed. Only those accidents in which NTSB judged that the cause of the accident was suicide were included in the final analysis. The relative risk (RR) of the pilot AASs in all fatal accidents in the U.S. was calculated in order to compare the one, two, and three-year periods after the September 11 terrorist attacks with five years preceding the event. The RR of a fatal general aviation aircraft accident being due to pilot suicide was 3.68-fold (95% confidence interval 1.04–12.98) during the first year after 11 September 2001, but there was not a statistically significant increase in the later years. This study showed an association, albeit not determinate causal effect, of a very specific series of simultaneous terrorist murder-suicides with subsequent pilot AASs.

Keywords: September 11 terrorist attacks; pilot aircraft-assisted suicide; copycat effect

1. Background

Copycat phenomenon (suicidal behavior provoked by media exposure) in pilot aircraft-assisted suicides (AASs) has been studied since the 1970s with inconclusive results [1,2], while longitudinal case series on pilot AASs have shown the difficulty of predicting rare events [3,4].We have previously analyzed pilot AASs after the Germanwings pilot murder-suicide incident [5], with no increase in pilot AASs was observed in the U.S. or Germany during a two-year period after the incident as compared to the previous five-year period.

The negative impact of exposure to atrocities on mental health is well documented, from trauma and-stress related disorders, to the level of copycat suicides [6–9]. The exposure to traumatic events through media alone is sufficient to cause psychological distress or even stress-related disorders [10]. Copycat behavior can be defined as suicidal behavior provoked by media exposure, also described as the "Werther effect" (from Goethe's novel Die Leiden des Jungen Werthers). In addition, the term "Papageno effect" has been proposed to describe protective media measures [7,11]. The name originally refers to a bird catcher character from Mozart's Magic Flute, who became suicidal but recovered after his friends intervened [7,11]. World Health Organization (WHO) has guidance on responsible media reporting after suicide, and recent National Institute for Health and Care Excellence (NICE) guidelines on suicide prevention have also focused on the impact of media [12,13].

Pilots are generally regarded as being psychologically resilient and healthier than the general population due to initial health assessment of specialized aeromedical examiners, which excludes individuals with severe risk factors for suicidal behavior, such as repeated suicide attempts or current major depression. Mandatory health assessments are regularly repeated, as requested in aviation authorities' guidance on aviation medical examiners (AME)s [14,15]. At an individual pilot level, mental health fitness assessments may have an impact on aviation safety [16]. However, the use of several aircrafts in the September 11 attacks focused attention internationally on the actions of aviators and also on the potential vulnerability of aviation safety with excessive and repeated media coverage. In relation to military pilots, the U.S. Air Force has published fairly constant suicide rates for that time period in an analysis of the U.S. Air Force suicide prevention program [17].

A study by Claassen et al. [18] has shown no increase in suicide rates in the general population in areas surrounding the three airline crash sites in New York, the Pentagon in Washington D.C, and Somerset County in Pennsylvania after 11 September 2001. Claassen et al. [18] concluded that geographical proximity is less important compared to other event characteristics and more research is needed regarding relevant social factors. This is in line with the study of Mezuk and coworkers [19], who reported no increase in New York suicide rates after the attack. Monthly homicide statistics in the New York City area did not reveal an increase in homicide or suicide rates after 11 September 2001 [20,21].

Although Pridemore and colleagues [21] showed that the rate of homicide or suicide deaths in New York was not increased following the attacks of 11 September 2001, Jordan et al. [22] showed that the standardized mortality ratio (SMR) for suicide was elevated (SMR = 1.82) among rescue recovery workers, while SMR for all-cause mortality was not elevated (SMR = 0.69). In another study of mortality among World Trade Center rescue workers, Stein et al. [23] reported a lowered all-cause and cause-specific mortality among rescue workers during a time period of 2002–2011 and did not find any association relating to the duration of rescue work and mortality. These long-term follow-up study results are not conclusive, inviting more analyses among different professional groups related to similar types of incidents. Starkman [24] has reported a one-year increase in suicide attempts after 11 September 2001 in Michigan, U.S, and the effect was greatest in the months following the attacks.

One aspect to explore is the research on suicidal ideation. In a survey of 871 adults who experienced loss during the September 11 attacks, individuals with complicated grief had significantly high rates of suicidal ideation even after adjusting to comorbid depression [25]. Additionally, suicidal ideation was increased in another primary care cohort after 11 September 2001 (N = 444) in Manhattan [6].

The aim of this study is to evaluate possible changes in pilot aircraft-assisted suicides after the 11 September 2001, terrorist murder–suicides in the U.S.

2. Methods

The U.S. National Transportation Safety Board (NTSB) database was searched on 26 June 2018, using the following search words: "suicide", "murder–suicide", and "homicide–suicide" [26]. Fatal aviation accidents in the U.S, with full formal accident investigation reports finalized at the time of the search and the cause of accidents assessed as pilot suicide in the NTSB accident investigation, were included as index cases to this study. These NTSB accident investigations were further analyzed in-depth to assess the descriptive factors in these suicide processes. Only those accidents in which NTSB judged that the cause of the accident was suicide were included in the statistical analysis.

Frequency of the aircraft-assisted pilot suicides out of all fatal accidents in the U.S was calculated for the reference period (5 years before the 11 September 2001 terrorist murder-suicides event) and separately for the first, second, and third year after the September 11 terrorist attacks. The relative risk (RR) of the likelihood of an aircraft accident being due to pilot suicide for the three post-attack years was calculated.

3. Results

The number of fatal aviation accidents and relative risks of the likelihood of pilot AAS being due to suicide three years after 11 September 2001, calculated for each year separately, is presented in Table 1. All pilots who died were male. The RR was 3.68 in the first year and decreased close to 1.0 in the third year. Only the observation for the first year was statistically significant.

Table 1. The number and the risk ratios of pilot aircraft-assisted suicides five years before and three years after 11 September 2001, in the U.S.

Follow-Up Period	Dates	No. of Suicides	No. of Fatal Aviation Accidents	FRQ	Risk Ratio	95% Confidence Interval
Reference	11 September 1996–10 September 2001	6	1861	0.32%	1.00	Reference
1st year	12 September 2001–11 September 2002	4	337	1.19%	3.68	1.04–12.98
2nd year	12 September 2002–11 September 2003	3	375	0.80%	2.48	0.62–9.88
3rd year	12 September 2003–11 September 2004	1	354	0.28%	0.88	0.11–7.26

FRQ = frequency.

Altogether 23 fatal aviation accident reports were obtained with the specified search words. All index cases were identified with the search word "suicide", search words "homicide-suicide", or "murder-suicide" did not yield any additional index case incidents. Fourteen of these incidents were caused by a pilot's or co-pilot's suicidal act according to NTSB accident investigations, and these incidents are described in detail in Table 2. All the incidents were operationally related to general aviation, although three pilots had also Class I (commercial pilot) medical certification even though they were not flying commercial aircraft at the time of the suicide crash. It should be noted that three of the deceased pilots were also flight instructors. Altogether fourteen pilots or student pilots died in these incidents. Their ages ranged from 15 to 69 years. Eight of these aircraft-assisted suicides occurred after 11 September 2001; four within the first year, three during the second year, and one in the third year, while six aircraft-assisted suicides took place during the five-year-period before it (see timeline in Figure 1). NTSB causes of death for excluded cases in this search were: four undetermined causes of fatal aviation accidents, two related to psychological problems (drugs or alcohol were mentioned at least as a contributing factor), in two reports weather conditions were mentioned, and in one case death occurred due to a passenger's suicidal act.

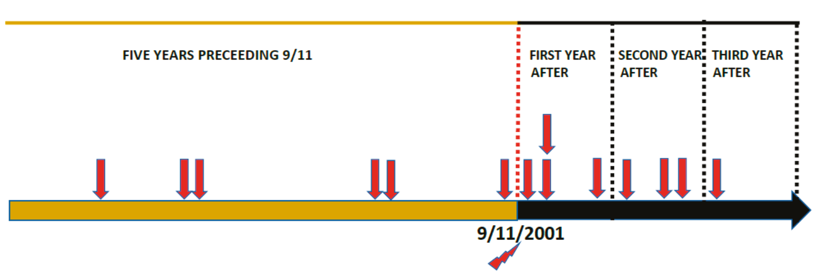


Figure 1. Timeline of aircraft-assisted pilot suicides five years before and three years after 11 September 2001. Red arrows indicate aircraft assisted suicides.

Table 2. Aircraft-assisted pilot suicides in the National Transportation Safety Board database five years before and three years after the 11 September 2001.

Event Date and State	Medical Certificate * (Assessment)	PM Toxicology	Health before the Flight	Cause of Accident by NTSB **	Other Information on Events before the Incident Flight
16 September 2003, Georgia	Class 2 (14 May 2002)	Ethanol	Normal	Intentional suicidal flight into the ground while impaired by alcohol.	Pilot threatened to commit suicide by flying into a mountain. No suicide note.
21 July 2003, Minnesota	Class 3 (9 April 2003)	Citalopram, fluoxetine, diphenhydramine	Depression	An act of suicide.	Suicide attempt with medication within a week with psychiatric hospitalization, "he wanted to get the courage to kill himself", left the hospital a day before the accident flight. No suicide note.
25 February 2003, Florida	Class 1 (12 March 2002)	Toxicology negative	No data	The intentional suicidal act.	No data. Autopsy: Manner of death suicide. No suicide note.
17 November 2002, Texas	Class 2 expired (3 September 1987)	Toxicology not done	Depression	An intentional suicide by exiting from the airplane, contributory was psychological condition.	Friends and co-workers: ongoing treatment for depression and a recent intent to take his own life by using an aircraft. A suicide note.
12 August 2002, Nebraska	Class 2 (7 July 1998)	Toxicology done, negative	No data	The pilot's intentional flight into terrain in an act of suicide.	On-going criminal investigation and was told to have threatened to kill himself by intentionally crashing an airplane before he would go to jail. No suicide note.
5 January 2002, Florida	Class 3 (17 November 2001)	Toxicology done, negative	No data	The pilot's unauthorized use of an aircraft for the purpose of committing suicide.	The airplane impacted the office building at the 28th-floor level. A suicide note.
5 January 2002, Colorado	Class 2 (28 September 2000)	Venlafaxine and its metabolite	Depression	Intentional suicidal flight into terrain. Contributing was depressive state, and inappropriate medication.	Received psychotherapy for severe depression. Told that if he killed himself, he would crash a plane with only himself in it. No suicide note. Manner of death: suicide.
4 January 2002, California	Class 3 (30 March 2001)	Done, no data given	No data	Intentional flight into terrain in an act of suicide.	Subject of a criminal investigation, the pilot's computer was seized from his home in the accident day. Additional emotional earlier distress. Coroner: suicide. No suicide note.
25 August 2001, New Hampshire	Class 3 (15 March 2000)	Done, no data given	No data	Suicide, the pilot intentionally crashed his airplane into his house.	A day before the pilot was issued a restraining order at his home and was escorted off his property. Cause of death: suicide. No suicide note.

Table 2. *Cont.*

Event Date and State	Medical Certificate * (Assessment)	PM Toxicology	Health before the Flight	Cause of Accident by NTSB **	Other Information on Events before the Incident Flight
2 October 2000, South Dakota	Class 1 (17 December 1999)	Toxicology negative	No data	Suicide.	Manner of death: suicide. No suicide note.
3 July 2000, Alaska	Class 1 (2 March 1999)	Ethanol, diazepam, cocaine, and their metabolites	Substance abuse history	Suicide.	A history of substance abuse, previously convicted of arson, sought by police. Friend: On the drive to the airport restless, agitated, depressed. Suicide note. Manner of death suicide.
11 October 1998, Oklahoma	Class 3 (7 August 1997)	Diazepam and its metabolites	No data	Suicide.	The accident site was located adjacent to a church where a friend, reported to have declined pilot's marriage proposal the night before, was attending services. No suicide note.
6 September 1998, Florida	Class 2 (31 January 1998)	Measured ethanol level referred to ethanol consumption before the incident	No data	The pilot's use of the aircraft to commit suicide.	Suicide note, "I do not want to live". Manner of death: suicide.
24 November 1997, California	Class 3 (4 November 1996)	Ethanol, postmortem production	Coroner: previous heart attack	Act of suicide by intentionally diving the aircraft into the ocean.	Girlfriend contacted the operator asking if the pilot had taken off and expressed concern about suicidality. A recent mother's death, worried about health. A will dated day before the incident, no suicide note. Mode of death: suicide.

* Class 1 = Airline Transport Pilot; 2 = Commercial Pilot; 3 = Private or Recreational Pilot; ** NTSB = National Transportation Safety Board.

4. Discussion

The one-year RR of 3.68 of aircraft-assisted pilot suicides in the U.S. after 11 September 2001, is in line with the results of rescue worker's suicide mortality reported by Jordan et al. [22], and some earlier copycat reports by Fink et al. [8]; see also Sisask and Värnik [7]. While fatal aviation accidents are rare, large or long-term register studies such as the study of Blettner et al. [27] and Politano and Walton [3] enable assessments to detect plausible effects of such events. General population suicide rates have been relatively stable during the time period of this study, but have increased in the U.S. during the last few years [28,29].

In studies from other countries, in England and Wales no increase of suicides was observed after 11 September 2001 [30] and there was no detectable change in the suicide rates in the Germany study [31]. An analysis of suicide rates from the Netherlands by de Lange and Neeleman [32] showed evidence of an increase in the suicide rates in the weeks after 11 September. Regarding trauma and stress related disorders, a Danish study by Hansen et al. [33] reported an increase in trauma and stress -related disorders after 11 September 2001, in line with several U.S. studies [34,35].

In one suicide flight involving a young pilot after 11 September 2001, the plane crashed intentionally into a high office building. According to the report there was a suicide note but it was not included in the accident investigation data. The helicopter pilots who intercepted this airplane attempted to signal the young pilot to land. According to the helicopter pilots, the student pilot saw their hand gestures and gestured back, but they could not determine the meaning of the gestures. Investigation of this incident led to a Federal Aviation Administration security notice release.

In the incidents investigated in this study, previous depression with suicide attempts was described in one case, ongoing treatment for depression in two cases, and a history of substance abuse and arson conviction in one case. Toxicology revealed that in six of these incidents there were medications or substance misuse not compatible with flying according to FAA. Some pilots with no information on previous psychological or psychiatric issues had recent stressful personal situations, such as legal or interpersonal difficulties.

The choice of crashing an aircraft to commit suicide among some pilots may in part be understood through the desire to combine passion for flying with death, reflecting an intimate bond between the pilot and his means of dying. Occasionally there may also be a hope that the incident would be construed as a medical fatality if the evidence were to be destroyed in the crash (for instance if the aircraft disappears).

One of the Class 1 pilots had a history of substance dependence. He was convicted of driving when intoxicated on several occasions and of arson, and had acute stressors. He left a suicide note. Another Class 1 pilot with no known immediate stressful events or health-related issues asked air traffic control to call for airport rescue and firefighting, and "also if you could tell my family and friends that I love them very much". Thus, he could be reached immediately before the lethal act. The third Class 1 pilot with military helicopter experience told the dispatcher he was going to meet someone. At no time did he display any signs of stress or unusual behavior. There were no records of previous incidents, and he had flight experience with various helicopters. This description gives an impression of a person who made a decision and enacted it; he was not in contact with any air traffic control facility at the time of the incident.

Previous suicide attempts and major depression are risk factors for completed suicide, but data on depression among aviators is limited [13,36]. In civil aviation, ongoing symptomatic major depression is not compatible with flying duties. After full remission and sufficient follow-up, International Civil Aviation Organization (ICAO) and aviation authorities currently accept psychological and some pharmacological treatments for aviators to prevent recurrence [14,15].

In aviation, full remission of depressive or stress-related symptoms, good compliance, and a reasonable follow-up time post-recovery are a prerequisite to the consideration of an aviator's return to flying duties [13–15,36–38]. Accepted antidepressants used to prevent symptoms are assessed through specific programs that focus on comprehensively assessing fitness to fly, symptoms, treatment adherence, and especially aero-medically relevant side-effects (e.g., fatigue).

In relation to mental health and suicide, it should be noted that among pilots regular health examinations contribute to the recognition of risk factors of suicidal behavior. The actual risk factors for pilots may somewhat differ from the general population, in relation to the role of any defined psychiatric disorders [13,39,40].

The limitations of this study include the fact that our data are based on information in the accident investigations, and we do not have access to additional medical information for determining aviator suicides nor to any information on aviators' attempted suicides prior to the incidence. The data are reliant on NTSB reports, and these reports are incomplete in some cases and certain details remain unknown. The news information after 11 September and timing of news delivery was not analyzed, but there are several studies from the U.S. on psychological effects after 11 September 2001 [6,25,41]. We do not think that any flying pilot in the U.S. could have avoided the news, due to immediate alerts and airspace closures in the U.S. The role of media research and balancing Werther and Papageno effects in general population suicide prevention is more thoroughly assessed elsewhere [42–45]. However, we consider our pilot AAS estimates to be conservative since those fatal aviation incidents where autopsy indicated suicide, but NTSB accident investigations did not agree, were excluded.

5. Conclusions

This study showed an association of 11 September 2001, with pilot AASs. The copycat effect was present for one year after 11 September 2001. The causal factors behind this statistical association remain unclear in context to the theoretical approaches on suicidality. The use of aircraft as the means to commit the terrorist attacks in the U.S. on 11 September 2001, may have had a negative effect on a few acutely vulnerable pilots. This vulnerability warrants further investigation, particularly with reference to the copycat phenomenon and it needs to be taken into account in aviation medical safety risk assessments.

Author Contributions: A.V. contributed to the study design, data procurement, results interpretation and drafted the article. T.L. drafted first version of the ms. A.V., T.L., I.J., E.P. were in charge of data analysis. A.V., T.L., I.J. contributed to results interpretation. All authors A.V., B.B., A.S., T.L., I.J., E.P., P.N., R.B. critically revised the article and gave their final approval for publication.

Funding: The Academy of Finland, Finnish Medical Foundation, Sigrid Juselius Foundation, and the Tampere Tuberculosis Foundation.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

AAS	Aircraft Assisted Suicide
AME	Aviation Medical Examiner
CAA	Civil Aviation Authority
CI	Confidence Interval
FAA	Federal Aviation Administration
ICAO	International Civil Aviation Organization
NTSB	National Transportation Safety Board
PM	Post-Mortem
RR	Relative Risk

References

1. Phillips, D.P. Airplane accident fatalities increase just after newspaper stories about murder and suicide. *Science* **1978**, *201*, 748–750. [CrossRef] [PubMed]
2. Altheide, D.L. Airplane Accidents, Murder, and the Mass Media: Comment on Phillips. *Soc. Forces* **1981**, *60*, 593–596. [CrossRef]

3. Politano, P.M.; Walton, R.O. Analysis of NTSB aircraft-assisted pilot suicides: 1982–2014. *Suicide Life Threat. Behav.* **2016**, *46*, 234–238. [CrossRef] [PubMed]
4. Vuorio, A.; Laukkala, T.; Navathe, P.; Budowle, B.; Eyre, A.; Sajantila, A. Aircraft-assisted pilot suicides: Lessons to be learned. *Aerosp. Med. Hum. Perform.* **2014**, *85*, 841–846. [CrossRef] [PubMed]
5. Laukkala, T.; Vuorio, A.; Bor, R.; Budowle, B.; Navathe, P.; Pukkala, E.; Sajantila, A. Copycats in Pilot Aircraft-Assisted Suicides after the Germanwings Incident. *Int. J. Environ. Res. Public Health* **2018**, *15*, 491. [CrossRef] [PubMed]
6. Neria, Y.; Wickramaratne, P.; Olfson, M.; Gameroff, M.J.; Pilowsky, D.J.; Lantigua, R.; Shea, S.; Weissman, M.M. Mental and physical health consequences of the September 11, 2001 (9/11) attacks in primary Care: A longitudinal study: 9/11 health consequences in primary care. *J. Trauma Stress* **2013**, *26*, 45–55. [CrossRef] [PubMed]
7. Sisask, M.; Värnik, A. Media roles in suicide prevention: A systematic review. *Int. J. Environ. Res. Public Health* **2012**, *9*, 123–138. [CrossRef] [PubMed]
8. Fink, D.S.; Santaella-Tenorio, J.; Keyes, K.M. Increase in suicides during the months after the death of Robin Williams in the U.S. *PLoS ONE* **2018**, *13*, e0191405. [CrossRef] [PubMed]
9. Jang, S.A.; Sung, J.M.; Park, J.Y.; Jeon, W.T. Copycat suicide induced by entertainment celebrity suicides in South Korea. *Psychiatry Investig.* **2016**, *13*, 74–81. [CrossRef] [PubMed]
10. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders fifth edition DSM-5*; American Psychiatric Publishing: Washington, DC, USA, 2013.
11. Niederkrotenthaler, T.; Voracek, M.; Herberth, A.; Till, B.; Strauss, M.; Etzersdorfer, E.; Eisenwort, B.; Sonneck, G. Role of media reports in completed and prevented suicide: Werther v. Papageno effects. *Br. J. Psychiatry* **2010**, *197*, 234–243. [CrossRef] [PubMed]
12. World Health Organisation. *WHO: Preventing Suicide: A Resource for Media Professionals*; WHO reference number: WHO/MSD/MER/17.5; WHO: Geneva, Switzerland, 2017.
13. National Institute for Health and Care Excellence. *Preventing Suicide in Community and Custodial Settings*; NICE guideline [NG105] September; National Institute for Health and Care Excellence: London, UK, 2018.
14. International Civil Aviation Organization (ICAO). Manual of Civil Aviation Medicine. 2012. Available online: http://www.caa.co.za/Documents/Aviation%20Medicine/Manual%20of%20Civil%20Aviation%20Medicine.pdf (accessed on 24 June 2018).
15. Federal Aviation Authority (FAA). Guide for Aviation Medical Examiners. Decision Consideration—Aerospace Medical Dispositions. Item 47. Psychiatric Conditions—Use of Antidepressants; 2018. Available online: www.faa.gov/about/office_org/headquarters_offices/avs/offices/aam/ame/guide/app_process/exam_tech/item47/amd/antidepressants (accessed on 24 June 2018).
16. Bor, R.; Eriksen, C.; Oakes, M.; Scragg, P. *Pilot Mental Health Assessment and Support: A Practitioner's Guide*; Routledge: London, UK, 2017.
17. Knox, K.L.; Pflanz, S.; Talcott, G.W.; Campise, R.L.; Lavigne, J.E.; Bajorska, A.; Tu, X.; Caine, E.D. The US Air Force Suicide Prevention Program: Implications for Public Health Policy. *Am. J. Public Health* **2010**, *100*, 2457–2463. [CrossRef] [PubMed]
18. Claassen, C.A.; Carmody, T.; Stewart, S.M.; Bossarte, R.M.; Larkin, G.L.; Woodward, W.A.; Trivedi, M.H. Effect of 11 September 2001 terrorist attacks in the USA on suicide in areas surrounding the crash sites. *Br. J. Psychiatry* **2010**, *196*, 359–364. [CrossRef] [PubMed]
19. Mezuk, B.; Larkin, G.L.; Prescott, M.R.; Tracy, M.; Vlahov, D.; Tardiff, K.; Galea, S. The influence of a major disaster on suicide risk in the population. *J. Traumatic Stress* **2009**, *22*, 481. [CrossRef] [PubMed]
20. Pridemore, W.A.; Chamlin, M.B.; Trahan, A. A test of competing hypotheses about homicide following terrorist attacks: An interrupted time series analysis of September 11 and Oklahoma City. *J. Quant. Criminol.* **2008**, *24*, 381–396. [CrossRef]
21. Pridemore, W.A.; Trahan, A.; Chamlin, M.B. No Evidence of Suicide Increase Following Terrorist Attacks in the United States: An Interrupted Time-Series Analysis of September 11 and Oklahoma City. *Suicide Life Threat. Behav.* **2009**, *39*, 659–670. [CrossRef] [PubMed]
22. Jordan, H.T.; Stein, C.R.; Li, J.; Cone, J.E.; Stayner, L.; Hadler, J.L.; Brackbill, R.M.; Farfel, M.R. Mortality among rescue and recovery workers and community members exposed to the September 11, 2001 World Trade Center terrorist attacks, 2003–2014. *Environ. Res.* **2018**, *163*, 270–279. [CrossRef] [PubMed]

23. Stein, C.R.; Wallenstein, S.; Shapiro, M.; Hashim, D.; Molinen, J.M.; Udasin, I.; Crane, M.A.; Luft, B.J.; Lucchini, R.G.; Holden, W.L. Mortality among World Trade Center rescue and recovery workers, 2002–2011. *Am. J. Ind. Med.* **2016**, *59*, 87–95. [CrossRef] [PubMed]
24. Starkman, M.N. The terrorist attacks of September 11, 2001, as psychological toxin: Increase in suicide attempts. *J. Nerv. Ment. Dis.* **2006**, *194*, 547–550. [CrossRef] [PubMed]
25. Neria, Y.; Gross, R.; Litz, B.; Maguen, S.; Insel, B.; Seirmarco, G.; Rosenfeld, H.; Suh, E.J.; Kishon, R.; Cook, J.; Marshall, R.D. Prevalence and psychological correlates of complicated grief among bereaved adults 2.5–3.5 years after September 11th attacks. *J. Trauma Stress* **2007**, *20*, 251–262. [CrossRef] [PubMed]
26. National Transportation Safety Board (NTSB). Accident Database and Synopsis. Available online: http://www.ntsb.gov/aviationquery/index.aspx (accessed on 26 June 2018).
27. Blettner, M.; Zeer, H.; Auvinen, A.; Ballard, T.J.; Caldora, M.; Haldorsen, T.; Hammar, N.; Hammer, G.P.; Irvine, D.; Langner, I.; et al. Mortality from cancer and other causes among male airline cockpit crew in Europe. *Int. J. Cancer* **2003**, *106*, 946–952. [CrossRef] [PubMed]
28. McKeown, R.E.; Cuffe, S.P.; Schulz, R.M. US Suicide Rates by Age Group, 1970–2002: An Examination of Recent Trends. *Am. J. Public Health* **2006**, *96*, 1744–1751. [CrossRef] [PubMed]
29. Stone, D.M.; Simon, T.R.; Fowler, K.A.; Kegler, S.R.; Yuan, K.; Holland, K.M.; Ivey-Stephenson, A.Z.; Crosby, A.E. Vital Signs: Trends in State Suicide Rates—United States, 1999–2016 and Circumstances Contributing to Suicide—27 States, 2015. *MMWR* **2018**, *67*, 617–624. [CrossRef] [PubMed]
30. Salib, E. Effect of 11 September 2001 on suicide and homicide in England and Wales. *Br. J. Psychiatry* **2003**, *183*, 207–212. [CrossRef] [PubMed]
31. Medenwald, D. The terror attacks of 9/11 and suicides in Germany a time series analysis. *Medicine* **2016**, *95*, 1–5. [CrossRef] [PubMed]
32. De Lange, A.W.; Neeleman, J. The Effect of the September 11 terrorist attacks on suicide and deliberate self-harm: A time trend study. *Suicide Life Threat. Behav.* **2004**, *34*, 439–447. [CrossRef] [PubMed]
33. Hansen, B.T.; Østergaard, S.D.; Sønderskov, K.M.; Dinesen, P.T. Increased incidence rate of trauma- and stressor-related disorders in Denmark after the September 11, 2001, terrorist attacks in the United States. *Am. J. Epidemiol.* **2016**, *184*, 494–500. [CrossRef] [PubMed]
34. Updegraff, J.A.; Silver, R.C.; Holman, E.A. Searching for and finding meaning in collective Trauma: Results from a national longitudinal study of the 9/11 terrorist attacks. *J. Pers. Soc. Psychol.* **2008**, *95*, 709–722. [CrossRef] [PubMed]
35. Silver, R.C.; Holman, E.A.; McIntosh, D.N.; Poulin, M.; Gil-Rivas, V. Nationwide longitudinal study of psychological responses to September 11. *JAMA* **2002**, *288*, 1235–1244. [CrossRef] [PubMed]
36. Pasha, T.; Stokes, P.R.A. Reflecting on the Germanwings disaster: A systematic review of depression and suicide in commercial airline pilots. *Front. Psychiatry* **2018**, *9*, 86. [CrossRef] [PubMed]
37. Laukkala, T.; Bor, R.; Budowle, B.; Navathe, P.; Sajantila, A.; Sainio, M.; Vuorio, A. Pilot Posttraumatic Stress Disorder and Fatal Aviation Accidents: A Descriptive Study. *Aviat. Psychol. Appl. Hum. Factors* **2018**, *8*, 93–99. [CrossRef]
38. Vuorio, A.; Budowle, B.; Sajantila, A.; Laukkala, T.; Junttila, I.; Kravik, S.E.; Griffiths, R. Duty of Notification and Aviation Safety—A Study of Fatal Aviation Accidents in the United States in 2015. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1258. [CrossRef] [PubMed]
39. Vuorio, A.; Laukkala, T.; Navathe, P. Major depression and fitness to fly by different aviation authorities. *Aviat. Space Environ. Med.* **2012**, *83*, 909–911. [CrossRef] [PubMed]
40. Aerospace Medical Association; Ad Hoc Working Group on Pilot Mental Health. Pilot Mental Health Expert Working Group Recommendations Revised 2015. *Aerosp. Med. Hum. Perform.* **2016**, *87*, 505–507. [CrossRef] [PubMed]
41. Silver, R.C.; Holman, E.A.; Andersen, J.P.; Poulin, M.; McIntosh, D.N.; Gil-Rivas, V. Mental- and physical-health effects of acute exposure to media images of the September 11, 2001, attacks and the Iraq War. *Psychol. Sci.* **2013**, *24*, 1623–1634. [CrossRef] [PubMed]
42. Reinemann, C.; Scherr, S. Der Werther-Defekt: Plädoyer für einen neuen Blick auf den Zusammenhang von suizidalem Verhalten und Medien [The Werther defect: Plea for a new view on the association between suicidal behavior and the media]. *Publizistik* **2011**, *56*, 89–94. [CrossRef]

43. Scherr, S. Medien und Suizide: Überblick über die kommunikationswissenschaftliche Forschung zum Werther-Effekt [Media and suicide: Review of research on the Werther effect in communication science]. *Suizidprophylaxe* **2013**, *40*, 96–107.
44. Scherr, S. *Depression—Medien—Suizid: Zur empirischen Relevanz von Depressionen und Medien für die Suizidalität [Depression—Media—Suicide: On the Empirical Relevance of Depression and Media for Suicidality]*; Springer VS: Wiesbaden, Germany, 2016.
45. Scherr, S.; Steinleitner, A. Zwischen dem Werther- und Papageno-Effekt. *Nervenarzt* **2015**, *86*, 557–565. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

Body Mass Index and *Helicobacter pylori* among Obese and Non-Obese Patients in Najran, Saudi Arabia: A Case-Control Study

Ali M. Al-Zubaidi [1,*], Abdo H. Alzobydi [2], Saeed A. Alsareii [3], AbdulazizTurky Al-Shahrani [4], Naweed Alzaman [5] and Saba Kassim [6]

1 Department of Medicine, Endoscopy Unit, King Khalid Hospital, Najran 1120, Saudi Arabia
2 Department of Surgery, Bariatric Surgery, King Khalid Hospital-Najran 1120, Saudi Arabia; alzobydi@yahoo.com
3 Department of Surgery, Najran University Medical College and Hospital, Najran 1988, Saudi Arabia; alsareii997@hotmail.com
4 Department of Internal Medicine; Najran University Medical College and Hospital, Nagran 1988, Saudi Arabia; dr.abdulaziz2012@hotmail.com
5 Department of Internal Medicine, Taibah University Medical College, Al-Madinah Al-Munawwarah 42353, Saudi Arabia; naweedalzamanmd@gmail.com
6 Department of Preventive Dental Sciences, Taibah University Dental College and Hospital, Al-Madinah Al-Munawwarah 42353, Saudi Arabia; saba262003@gmail.com
* Correspondence: dr_ali26@yahoo.com; Tel.: +96-56-239-7881

Received: 12 October 2018; Accepted: 16 November 2018; Published: 19 November 2018

Abstract: *Objective*: We examine obese and non-obese patients with respect to Helicobacter pylori (*H. pylori*) positive-infection (HPPI) and associated factors, specifically body mass index (BMI). *Methods*: This study took place in the Department of Endoscopy of a central hospital in the Najran region of Saudi Arabia (SA). A total of 340 obese Saudi patients (BMI $\geq$ 30 kg/m^2) who had undergone diagnostic upper endoscopy before sleeve gastrectomy, were compared with 340 age and gender-matched control patients (BMI < 30 kg/m^2) who had undergone diagnostic upper endoscopy for other reasons. Data collected included diagnosis of HPPI. Descriptive and multivariable binary logistic regression was conducted. *Results*: Mean patient age was 31.22 $\pm$ 8.10 years, and 65% were males. The total prevalence of HPPI was 58% (95% CI = 54–61%) with obese patients presenting significantly more HPPI than non-obese patients (66% vs. 50%, OR = 1.98, 95% CI = 1.45–2.70, $p < 0.0005$). Age and gender did not associate significantly with HPPI ($p = 0.659$, 0.200, respectively) and increases in BMI associated significantly with increases in HPPI ($p < 0.0005$). BMI remained a significant factor in HPPI when modelled with both age and gender (OR = 1.022, 95% CI = 1.01–1.03, $p < 0.0005$). *Conclusions*: Within the limitations of this study, the significance of HPPI in obese Saudi patients residing in the Najran region in SA was demonstrated alongside the significance role of BMI in HPPI.

Keywords: obesity; endoscopy; *Helicobacter pylori*; BMI; Saudi patients

1. Introduction

Helicobacter pylori (*H. pylori*) is one of the most common infections affecting the epithelial lining of the stomach [1]. It is correlated with antral gastritis, peptic ulcers and promotes gastric malignancies [2]. There were approximately 4.4 billion individuals with *H. pylori* infection worldwide in 2015 [1]. *H. pylori* infection is prevalent in developing countries [1] and risk factors for *H. pylori* including age and socioeconomic status [2].

There is ongoing debate over the relationship between obesity and *H. pylori* infection, within the acknowledgment that the etiology of obesity is far more complex [3–5]. A recent ecological review of several cross sectional studies found an inverse correlation between *H. pylori* prevalence and rate of overweight/obesity in countries of the developed world i.e., increase in *H. pylori* positive infection associated with reduction in obesity [6]. This was further corroborated with Intervention studies reporting that eradication of *H. pylori* was associated with significant weight gain as compared to subjects with untreated *H. pylori* [7,8]. In contrast, observational or clinical studies from developing countries reported a linear relationship between *H. pylori* positive infection and obesity [9–11].

Saudi Arabia has the highest obesity and overweight prevalence rates in the Middle East which is linked to multiple factors including adapting a westernized life style [12]. The prevalence of *H. pylori* within the Saudi Arabian population and related factors remains unknown and most available data is reported from medical care settings with a range of prevalence between 33–85% [9,13–17]. A study randomly selecting university students reported a prevalence of 35% for *H. pylori* [18]. Specifically, high *H. pylori* prevalence was reported among morbid obese patients who underwent upper endoscopy prior to bariatric surgery was 88% [9]. Yet the matching of this latter group with normal body weight patients to substantiate such findings within the context of Saudi Arabia has not been attempted specifically in the light of controversy of the association of *H. pylori* with obesity [6–8].

Morbid obesity has been widely treated with various type of bariatric surgery. Bariatric surgery includes laparo-scopic adjustable gastric bands (LAGB), lap band, the Roux-Y gastric bypass (RYGB) and sleeve gastrectomy or gastric sleeve [19]. Routine upper endoscopy with *H. pylori* screening and biopsies to rule out pathological abnormalities (e.g., gastritis, duodenitis, esophagitis, ulceration, hiatus hernia, ..., etc.) is able to detect abnormalities in up to 91% of bariatric candidates [20] with a higher rate in patients with concomitant *H. pylori* infection [21–24].

The American Association of Clinical Endocrinologists/The Obesity Society/American Society for Metabolic and Bariatric Surgery guidelines [25] have not provided any clear indication about preoperative *H. pylori* screening and management but recommended *H. pylori* screening in patients belonging to high-prevalence areas and upper endoscopy in selected cases. On the other hand, the European guidelines [26] have recommended upper gastrointestinal endoscopy before bariatric surgery in both symptomatic and asymptomatic patients and to treat *H. pylori* infection and other abnormalities, which may cause postoperative complications.

In light of the limited literature to assess the prevalence of HPPI in patients undergoing bariatric surgery from the Middle East and North Africa (MENA) and specifically Saudi Arabia therefore, we aimed in this hospital-based study to determine the estimate of *H. pylori* positive infection (HPPI) in group of morbidly obese subjects undergoing sleeve gastrectomy and compare it to a match control group (age and gender) who had an upper endoscopy in the same period and same setting with different indications but had normal body weight. Based on the conflicting results of the current literature, we hypothesized, after controlling for age and gender, that we will be able to find out whether there is correlation between HPPI and obesity.

2. Material and Methods

2.1. Study Design and Setting

A case-control study design was adopted in the department of endoscopy at a central hospital (King Khalid) which received all medical referrals in Najran region southwest of Saudi Arabia.

2.2. Subjects' Recruitment and Inclusion and Exclusion Criteria

A consecutive test group consisted of 340 obese patients; defined with a BMI of $\geq$30 kg/m^2 with a serious comorbidity (e.g., diabetes mellitus, ischemic heart disease, obstructive sleep apnea) or with a BMI of $\geq$40 [27] who had undergone diagnostic upper endoscopy before sleeve gastrectomy were enrolled into the study over the period between January 2013–December 2014. Over the same period, a similar number of 340 patients; with BMI of <30 kg/m^2 who had undergone diagnostic upper endoscopy in the same hospital and department due to other reasons and complain (e.g., epigastric pain and dyspepsia) were selected as controls with approximately 1:1 ratio test to control. Both groups self-reported neither tested nor treated for *H. Pylori* in the past 5 years.

We excluded patients with advanced liver disease, malignancy, renal failure, those taking bismuth subcitrate, history of taking eradication therapy of HPPI within six months, with acute infection and, gastro intestinal bleeding or patients with history of proton pump inhibitor (PPI) or antibiotics intake 2 weeks before the endoscopic procedure as these are known to may lead to false negative results for HPPI test.

2.3. Variables and Clinical Procedures

The data for both groups included the demographics (age, gender), medication use, laboratory results (e.g., CBC, coagulation profile, RFT, LFT, Cortisol, TFT, blood sugar, lipid profile), endoscopic and histologic finding (e.g., Gastritis, PUD, GERD, Hiatus Hernia). All patients underwent the endoscopic procedure with standard technique (using Olympus GF260, 9.6 mm diameter scope) to assess for HPPI test. Patient were lying in the left lateral position with a mouth piece placed and their throat sprayed with adequate xylocaine. Two antral biopsies were taken from antrum and lower body for each patient, as per the routine practice of the endoscopy unit, and submitted for histologic evaluation. No complication related to endoscopy was reported in any subject. The specimens were stained by hematoxylin and eosin stain or methylene blue. The laboratory investigator was blinded to the sample status to avoid bias.

Body mass index (BMI) is a simple index of weight-for-height that is commonly used to classify overweight and obesity in adults. It is defined as a person's weight in kilograms divided by the square of his height in meters (kg/m^2). BMI greater than or equal to 25 is considered to be overweight, and BMI greater than or equal to 30 is considered to be obesity [28]. As for body mass index a digital scale was used to measure the weight of patient to the nearest 0.1 kg and height was considered in a standing position without shoes and in light clothing.

2.4. Ethical Considerations

The study's proposal was reviewed and guaranteed clearance from the Ethical Committee of King Khalid Hospital, Najran (H-11-N081, Ali Mothanna Al-Zubaidi, 2012). The study was conducted in accordance with the Declaration of Helsinki wherein respondents were briefed about the study and the anticipated benefits of participating in the study. An information sheet explaining the aims and objectives of the study was made available for each respondent. Voluntary participation was emphasized and withdrawal from the study at any time without giving reason(s) would not disadvantage the respondent from receiving any treatment. An informed consent form was obtained before participating in the study and confidentiality of the information obtained was assured.

2.5. Statistical Analysis

Descriptive analysis was conducted and continuous variables were reported using Mean $\pm$ SD and for categorical variables we used frequencies with percentages. Unpaired *T*-Test and Chi squared tests were performed to explore the relationship of continuous and categorical variables with the dependent variable ('HPPI' (*H. pylori* positive infection) and 'HPNI' (*H. pylori* negative infection)). Binary logistic regression analysis was used to model the relationship of BMI after controlling for age and gender with the dependent variable. A p value of <0.05 was considered significant. The statistical analyses aforementioned were performed withe SPSS Software version 16.0 (SPSS Inc., Chicago, IL, USA).

3. Results

3.1. Sample Characteristics and Distribution of HPPI between and within Obese and Non-Obese

The demographic characteristics for both groups are shown in Table 1. There were no significant differences in the age of both obese and none obese patients (p = 0.307) with a male predominance in both groups (66.5% and 64.4%). The total mean of BMI for both groups was 35.57 ± 13.97 and there were statistical significant differences in the mean BMI between the obese and none obese (BMI = 23.13 ± 3.80 for none obese vs. 48.04 ± 8.10 for obese, p < 0.0005). As demonstrated in Table 1, the total prevalence of *H. pylori* positive-infection (HPPI) was 58% (95% CI = 54–61%). A statistical significant difference in HPPI between obese and non-obese was observed (225 (66.2%) for obese vs. 169 (49.7%) non-obese, OR (95% CI) = 1.98 (1.45–2.70), p < 0.0005). Within the gender groups, females were found to slightly predominate in HPPI. Males and females with respect to HPPI among obese patients was non-significant (146 (64.6%) vs. 79 (69.3%), OR = 1.23 (0.76–2.00), p = 0.388) and likewise among non-obese (104 (47.5%) vs. 65 (53.7%), OR (95% CI) = 1.28 (0.82–2.00), p = 0.271). However, between groups, Table 1 shows that both males and females were statistical significantly different in HPPI (p < 0.0005, 0.014, respectively). Unpaired *T*-Test and Chi squared tests revealed that age and gender did not associate with HPPI (p = 0.659, 0.200, respectively) and increases in BMI associated significantly with increases in HPPI (p < 0.0005).

Table 1. Characteristics of obese (N = 340) and none-obese (N = 340) and distributions of HPPI between groups and gender.

Variables	Mean ± SD or *N* (%)		OR (95% CI)	*p*-Value
	Obese (Cases)	Non-Obese (Controlled)		
Age	31.54 ± 8.27	30.90 ± 7.93	-	0.307
Gender				
Male	226 (66.5)	219 (64.4)		
Female	114 (33.5)	121 (35.6)	0.93 (0.67–1.25)	0.572
BMI	48.04 ± 8.10	23.13 ± 3.80	-	0.0005
HPPI	225 (66.2)	169 (49.7)	1.98 (1.45–2.70)	0.0005
HPPI in male	146 (58.4)	104 (41.6)	2.02 (1.38–2.95)	0.0005
HPPI in female	79 (54.9)	65 (45.1)	1.95 (1.14–3.32)	0.014

SD = Standard deviation; F (%) = Frequency with percentage; HPPI = *H. pylori* positive infection; OR (95% CI) = Odd ratio with 95% confidence interval.

3.2. Multivariable Analyses of Adjusted Factors of HPPI

Multivariable binary logistic regression revealed that BMI remained significant factor in HPPI when modelled with both age and gender (OR = 1.022, 95% CI = 1.01–1.03, p < 0.0005) (Table 2).

Table 2. Adjusted model for the association of BMI with HPPI.

Explanatory Variables	Wald	AOR (95% CI) *	*p*-Value
Age	0.371	0.99 (0.98–1.03)	0.542
Gender			
Male		1	
Female	2.11	1.28 (0.92–1.77)	0.146
BMI	13.67	1.02 (1.01–1.03)	0.0005

* AOR (95% CI) = adjusted odd ratio and 95% confidence interval.

4. Discussion

To the best of our knowledge, this case-control study is the first conducted in this region of Najran, SA that considered the none-modifiable factors (age and gender) in modelling relationship between obesity and *H. pylori* positive infection. The key finding was that increased prevalence of HPPI was associated with increased BMI levels, both with and without adjustment for none modifiable factors (age and gender). The findings of this study added to the current debate about the relationship of obesity with HPPI. A recent review showed a significantly inverse correlation between *H. pylori* prevalence and rate of overweight/obesity in countries of the developed world. Thus, the obesity endemic observed in the Western world was attributed to the gradual decrease of the HPPI [6]. These later results were also supported with studies in certain areas in Asian region e.g., Taiwan, [29].

However, our study findings were in consensus with previous studies reported within SA [9] and elsewhere [11,30]. In Turkey a prevalence of 57.2% of HPPI in Turkish obese subjects compared to 27.0% in normal body weight was found and this further supported our finding [31].

The estimate of HPPI among obese patients who underwent sleeve gastrectomy in this study sample was 66% which was in mid-range (50–88) of previous studies conducted in in SA [13,32,33] and higher than a study among young medical students that showed a low prevalence of 35% [18]. Different study design, participants' recruitments and clinical investigation may reflect such variations in the estimate of HPPI among obese subjects. In addition, the role of the difference in geographic, economic or environmental factors should be considered. Notably, our study was comparable to one US study which found a higher prevalence rate of 61% in morbidly obese patients compared to 48% in the control group [34] though this later used *H. pylori* serologies, while our study used histopathology diagnosis. With respect to the role of age and gender in HPPI, they were found to be non-significantly associated with HPPI and these findings were in consensus with other studies reported in the region [15,33] and elsewhere [35].

As for the strengths of this study, firstly our study findings have substantiated the relationship of obesity with *H. pylori* among obese patients who were matched with none obese patients consecutively recruited in the same setting with a standardized investigation procedures i.e., internal validity plausible. Secondly, this was a hospital-based case-control as such more prone to bias, specifically when selecting participants. However, we were in line with what has been reported previously [30] patients underwent a gastroscopy and received a histological examination of gastric mucosal biopsies as a standard diagnostic examination before bariatric surgery. Thus, the majority of the examined patients were asymptomatic which excludes a strong selection bias toward HPPI. Our study results were of validity as the inclusion and exclusion criteria was adequate to rule out any confounding effects; e.g., use of antibiotics [34]. The enrolment of patients from the same region and in one Centre (Najran Hospital) is another strength of our study, the difference in the estimate of *H. pylori* between the two groups was less likely to represent a differences in geographic prevalence of *H. pylori* as reported elsewhere [34]. Finally, the feasibility of the study with respect of time and resources was additional strength. As for the limitations, the sampling method adopted precludes the generalizability of the study findings to individuals outside of this setting. The patients self-reported of neither tested nor treated for *H. pylori* was not objectively validated, therefore, recall bias was possible. The impact of the latter on the validity of the results might be inevitable. Importantly, in SA the role of race in HPPI should be investigated as elsewhere [29]. The use of the concept of nationality is not sufficient as many Saudi national of different ethnicity/origin backgrounds that presumably influence HPPI. The contribution of other environmental and host factors [36] which has not been controlled for in this study should be considered in future research.

5. Conclusions

Within the limitations of this study, the significance of HPPI in obese Saudi patients residing in the Najran region in SA was demonstrated alongside the significance role of BMI in HPPI. Further research should consider a comprehensive and context specific factors (e.g., environmental and host) contribution to clarify the association of obesity with HPPI.

Author Contributions: A.M.A-Z. contributed to study conception and design, conducted the endoscopic procedures, acquired data, reviewed the literature, interpreted the results, wrote the manuscript draft and approved the final draft of the manuscript; A.H.A. Conducted the bariatric sleevectomy, reviewed the manuscript and data analysis; A.A-s. managed data collection; S.A.A. conducted the bariatric sleevectomy, reviewed the manuscript; S.K. reviewed the literature, analyzed the data, interpreted the results and wrote the manuscript draft and approved the final draft of the manuscript, N.A. drafted the original manuscript, participated in literature review and critically commented on the draft.

Funding: This research received no external funding.

Acknowledgments: We would like to thank Madia Divikaran, Muhammed Farouk, Esam Alharoun and endoscopy nurses Rinu Kurain, Mini Mathew Sharly, minimol and Abdullah Alshahi for their endless efforts and support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hooi, J.K.Y.; Lai, W.Y.; Ng, W.K.; Suen, M.M.Y.; Underwood, F.E.; Tanyingoh, D.; Malfertheiner, P.; Graham, D.Y.; Wong, V.W.S.; Wu, J.C.Y.; et al. Global Prevalence of Helicobacter pylori Infection: Systematic Review and Meta-analysis. *Gastroenterology* **2017**, *153*, 420–429. [CrossRef] [PubMed]
2. Pellicano, R.; Ribaldone, D.G.; Fagoonee, S.; Astegiano, M.; Saracco, G.M.; Megraud, F. A 2016 panorama of Helicobacter pylori infection: Key messages for clinicians. *Panminerva Med.* **2016**, *58*, 304–317. [PubMed]
3. Boone-Heinonen, J.; Gordon-Larsen, P.; Kiefe, C.I.; Shikany, J.M.; Lewis, C.E.; Popkin, B.M. Fast food restaurants and food stores: Longitudinal associations with diet in young to middle-aged adults: The CARDIA study. *Arch. Intern. Med.* **2011**, *171*, 1162–1170. [CrossRef] [PubMed]
4. Singh, M.; Kirchengast, S. Obesity prevalence and nutritional habits among Indian women: A comparison between Punjabi women living in India and Punjabi migrants in Vienna, Austria. *Anthropol. Anz.* **2011**, *68*, 239–251. [CrossRef] [PubMed]
5. Zillikens, M.C.; van Meurs, J.B.; Rivadeneira, F.; Amin, N.; Hofman, A.; Oostra, B.A.; Sijbrands, E.J.; Witteman, J.C.; Pols, H.A.; van Duijn, C.M.; et al. SIRT1 genetic variation is related to BMI and risk of obesity. *Diabetes* **2009**, *58*, 2828–2834. [CrossRef] [PubMed]
6. Lender, N.; Talley, N.J.; Enck, P.; Haag, S.; Zipfel, S.; Morrison, M.; Holtmann, G.J. Review article: Associations between Helicobacter pylori and obesity—An ecological study. *Aliment. Pharmacol. Ther.* **2014**, *40*, 24–31. [CrossRef] [PubMed]
7. Kamada, T.; Sugiu, K.; Haruma, K. Development of obesity and hyperlipidemia after eradication of Helicobacter pylori. *Nihon Rinsho* **2005**, *63* (Suppl. 1), 536–538. [PubMed]
8. Lane, J.A.; Murray, L.J.; Harvey, I.M.; Donovan, J.L.; Nair, P.; Harvey, R.F. Randomised clinical trial: Helicobacter pylori eradication is associated with a significantly increased body mass index in a placebo-controlled study. *Aliment. Pharmacol. Ther.* **2011**, *33*, 922–929. [CrossRef] [PubMed]
9. Al-Akwaa, A.M. Prevalence of Helicobacter pylori infection in a group of morbidly obese Saudi patients undergoing bariatric surgery: A preliminary report. *Saudi J. Gastroenterol.* **2010**, *16*, 264–267. [CrossRef] [PubMed]
10. Papasavas, P.K.; Gagne, D.J.; Donnelly, P.E.; Salgado, J.; Urbandt, J.E.; Burton, K.K.; Caushaj, P.F. Prevalence of Helicobacter pylori infection and value of preoperative testing and treatment in patients undergoing laparoscopic Roux-en-Y gastric bypass. *Surg. Obes. Relat. Dis.* **2008**, *4*, 383–388. [CrossRef] [PubMed]
11. Xu, C.; Yan, M.; Sun, Y.; Joo, J.; Wan, X.; Yu, C.; Wang, Q.; Shen, C.; Chen, P.; Li, Y.; et al. Prevalence of Helicobacter pylori infection and its relation with body mass index in a Chinese population. *Helicobacter* **2014**, *19*, 437–442. [CrossRef] [PubMed]
12. Alqarni, S.S.M. A Review of Prevalence of Obesity in Saudi Arabia. *J. Obes. Eat. Disord.* **2016**. [CrossRef]
13. Al-Moagel, M.A.; Evans, D.G.; Abdulghani, M.E.; Adam, E.; Evans, D.J., Jr.; Malaty, H.M.; Graham, D.Y. Prevalence of Helicobacter (formerly Campylobacter) pylori infection in Saudia Arabia, and comparison of those with and without upper gastrointestinal symptoms. *Am. J. Gastroenterol.* **1990**, *85*, 944–948. [PubMed]

14. Elsawaf, Z.M.; Albasri, A.M.; Hussainy, A.S.; Alhujaily, A.S. Histopathological pattern of benign endoscopic gastric biopsies in Western Saudi Arabia: A review of 1236 cases. *J. Pak. Med. Assoc.* **2017**, *67*, 252–255. [PubMed]
15. Khan, A.R. An age- and gender-specific analysis of H. Pylori infection. *Ann. Saudi Med.* **1998**, *18*, 6–8. [CrossRef] [PubMed]
16. Mohamed, A.E.; Al-Karawi, M.A.; Al-Jumah, A.A.; Ahmed, A.M.; Shariq, S.; Yasawy, M.I.; Haleem, A. Helicobacter pylori: Prevalence in 352 consecutive patients with dyspepsia. *Ann. Saudi Med.* **1994**, *14*, 134–135. [CrossRef] [PubMed]
17. Rashed, R.S.; Ayoola, E.A.; Mofleh, I.A.; Chowdhury, M.N.; Mahmood, K.; Faleh, F.Z. Helicobacter pylori and dyspepsia in an Arab population. *Trop. Geogr. Med.* **1992**, *44*, 304–307. [PubMed]
18. Almadi, M.A.; Aljebreen, A.M.; Tounesi, F.A.; Abdo, A.A. Helicobacter pylori prevalence among medical students in a high endemic area. *Saudi Med. J.* **2007**, *28*, 896–898. [PubMed]
19. Meyer, S.B.; Thompson, C.; Hakendorf, P.; Horwood, C.; McNaughton, D.; Gray, J.; Ward, P.R.; Mwanri, L.; Booth, S.; Kow, L.; et al. Bariatric surgery revisions and private health insurance. *Obes. Res. Clin. Pract.* **2017**, *11*, 616–621. [CrossRef] [PubMed]
20. Csendes, A.; Burgos, A.M.; Smok, G.; Beltran, M. Endoscopic and histologic findings of the foregut in 426 patients with morbid obesity. *Obes. Surg.* **2007**, *17*, 28–34. [CrossRef] [PubMed]
21. Azagury, D.; Dumonceau, J.M.; Morel, P.; Chassot, G.; Huber, O. Preoperative work-up in asymptomatic patients undergoing Roux-en-Y gastric bypass: Is endoscopy mandatory? *Obes. Surg.* **2006**, *16*, 1304–1311. [CrossRef] [PubMed]
22. Ramaswamy, A.; Lin, E.; Ramshaw, B.J.; Smith, C.D. Early effects of Helicobacter pylori infection in patients undergoing bariatric surgery. *Arch. Surg.* **2004**, *139*, 1094–1096. [CrossRef] [PubMed]
23. Schirmer, B.; Erenoglu, C.; Miller, A. Flexible endoscopy in the management of patients undergoing Roux-en-Y gastric bypass. *Obes. Surg.* **2002**, *12*, 634–638. [CrossRef] [PubMed]
24. Vanek, V.W.; Catania, M.; Triveri, K.; Woodruff, R.W., Jr. Retrospective review of the preoperative biliary and gastrointestinal evaluation for gastric bypass surgery. *Surg. Obes. Relat. Dis.* **2006**, *2*, 13–17. [CrossRef] [PubMed]
25. Mechanick, J.I.; Youdim, A.; Jones, D.B.; Garvey, W.T.; Hurley, D.L.; McMahon, M.M.; Heinberg, L.J.; Kushner, R.; Adams, T.D.; Shikora, S.; et al. Clinical practice guidelines for the perioperative nutritional, metabolic, and nonsurgical support of the bariatric surgery patient—2013 update: Cosponsored by American Association of Clinical Endocrinologists, The Obesity Society, and American Society for Metabolic & Bariatric Surgery. *Obesity* **2013**, *21* (Suppl. 1), S1–S27.
26. Sauerland, S.; Angrisani, L.; Belachew, M.; Chevallier, J.M.; Favretti, F.; Finer, N.; Fingerhut, A.; Garcia Caballero, M.; Guisado Macias, J.A.; Mittermair, R.; et al. Obesity surgery: Evidence-based guidelines of the European Association for Endoscopic Surgery (EAES). *Surg. Endosc.* **2005**, *19*, 200–221. [CrossRef] [PubMed]
27. WHO. *BMI Classification*; WHO: Geneva, Switzerland, 2000.
28. Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults–The Evidence Report. National Institutes of Health. *Obes. Res.* **1998**, *6* (Suppl. 2), S51–S209.
29. Ekesbo, R.; Nilsson, P.M.; Lindholm, L.H.; Persson, K.; Wadstrom, T. Combined seropositivity for *H. pylori* and *C. pneumoniae* is associated with age, obesity and social factors. *J. Cardiovasc. Risk* **2000**, *7*, 191–195. [CrossRef] [PubMed]
30. Gerig, R.; Ernst, B.; Wilms, B.; Thurnheer, M.; Schultes, B. Preoperative nutritional deficiencies in severely obese bariatric candidates are not linked to gastric Helicobacter pylori infection. *Obes. Surg.* **2013**, *23*, 698–702. [CrossRef] [PubMed]
31. Arslan, E.; Atilgan, H.; Yavasoglu, I. The prevalence of Helicobacter pylori in obese subjects. *Eur. J. Intern. Med.* **2009**, *20*, 695–697. [CrossRef] [PubMed]
32. Karima, T.M.; Bukhari, S.Z.; Ghais, M.A.; Fatani, M.I.; Hussain, W.M. Prevalence of Helicobacter pylori infection in patients with peptic ulcer diseases. *Saudi Med. J.* **2006**, *27*, 621–626. [PubMed]
33. Morad, N.A.; Ahmed, M.E.; Al-Wabel, A.; Foli, A.K. Helicobacter pylori associated dyspepsia in 208 patients from southern Saudi Arabia. *Ann. Saudi Med.* **1993**, *13*, 340–343. [CrossRef] [PubMed]

34. Erim, T.; Cruz-Correa, M.R.; Szomstein, S.; Velis, E.; Rosenthal, R. Prevalence of Helicobacter pylori seropositivity among patients undergoing bariatric surgery: A preliminary study. *World J. Surg.* **2008**, *32*, 2021–2025. [CrossRef] [PubMed]
35. Smith, S.; Jolaiya, T.; Fowora, M.; Palamides, P.; Ngoka, F.; Bamidele, M.; Lesi, O.; Onyekwere, C.; Ugiagbe, R.; Agbo, I.; et al. Clinical and Socio- Demographic Risk Factors for Acquisition of Helicobacter pylori Infection in Nigeria. *Asian Pac. J. Cancer Prev.* **2018**, *19*, 1851–1857. [PubMed]
36. Carabotti, M.; D'Ercole, C.; Iossa, A.; Corazziari, E.; Silecchia, G.; Severi, C. Helicobacter pylori infection in obesity and its clinical outcome after bariatric surgery. *World J. Gastroenterol.* **2014**, *20*, 647–653. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

A Fuzzy Comprehensive Assessment and Hierarchical Management System for Urban Lake Health: A Case Study on the Lakes in Wuhan City, Hubei Province, China

Teng Wang [1,†], Jingjing Yan [2,†], Jinlong Ma [2], Fei Li [2,*], Chaoyang Liu [2], Ying Cai [2], Si Chen [3], Jingjing Zeng [2] and Yu Qi [2]

1 Hubei Research Center of Water Affair, Hubei University of Economics, Wuhan 430073, China; wangtenghbue@163.com
2 School of Information and Safety Engineering, Zhongnan University of Economics and Law, Wuhan 430073, China; Yanjing@zuel.edu.cn (J.Y.); majinlong@zuel.edu.cn (J.M.); lcy@zuel.edu.cn (C.L.); 1993cy@zuel.edu.cn (Y.C.); jjzeng@zuel.edu.cn (J.Z.); z0004447@zuel.edu.cn (Y.Q.)
3 School of Resources and Environment, Hubei University, Wuhan 430062, China; kathryncs123@hotmail.com
* Correspondence: lifei@zuel.edu.cn
† These authors contributed equally to this work.

Received: 13 October 2018; Accepted: 18 November 2018; Published: 22 November 2018

Abstract: Environmental assessment of eutrophication or heavy metals in urban lakes is an important reference for identifying the pollution degree and formulating pollution prevention strategies. At present, the most research on lake health states is often evaluated from a single angle for toxic metals pollution or eutrophication using the standard comparison method for both, the comprehensive trophic level index (TLI), and the health risk assessment for toxic metals. Moreover, the above deterministic methods probably lead to biased or unreliable assessment due to the randomness and fuzziness in environment system caused by natural change and human activities. In this paper, a fuzzy comprehensive lake health assessment method (FCLHAM) was established to evaluate comprehensive lake health states more comprehensively and accurately, which integrates quantitative eutrophication and health risk considerations. To test and verify FCLHAM, 21 lakes, scientifically selected from the total 143 lakes in the Chinese Wuhan city as study case, were investigated and analyzed for their state of eutrophication and the health risk posed by heavy metals. According to the FCLHAM, the average comprehensive lake health state decreased in the sequence of L20 (considerate risk level) > L1–L17, L19, L21 (moderate risk level) > L18 (low risk level). Based on the result, lakes were classified into three categories: general management (L18), enhanced management (L1–L17, L19, L21), and priority management (L20). If the 143 lakes in Wuhan were classified by the "area-region-function" classification, they would be assigned to the same category as the representative lakes of the same type. At this point, we will attribute all of Wuhan's lakes to the three types. Depending on the characteristics of each type, a targeted approach to different types of management for each type of lake is a more efficient way to manage many of Wuhan's lakes. This management mode also serves as an effective reference for the environmental management of urban lakes both at home and abroad. In other words, according to the FCLHAM, a hierarchical management system based on lake characteristics classification was obtained.

Keywords: urban lake; comprehensive nutrition status Index; heavy metals; health risk; fuzzy comprehensive method

1. Introduction

Eutrophication is a common water pollution phenomenon caused by excessive nutrients and is accompanied by ecological problems such as algal blooms, oxygen depletion, aquatic organism death, and aquatic ecosystem deterioration [1]. With the increase of population and rapid urbanization, nutrient loading in lots of urban lakes has increased and exceeded their corresponding natural carrying capacities [2–4]. Especially, urban lake eutrophication poses a serious threat to regional economic development, the ecological environment, and drinking water security [5,6]. Furthermore, heavy metal pollutants are discharged into urban lakes through industrial, agricultural, and domestic waste-water discharges, precipitation, and the release of contaminated sediment [7–9]. To a certain extent, toxic metal accumulation in urban lakes can do serious harm to the "water—aquatic plant—aquatic animal" system which will affect human health directly or indirectly through drinking water, food chains, etc. [10].

In recent years, various works have been done to explore the heavy metal or eutrophication distributions and sources in urban lakes [11,12], as well as toxic metal eco-risk and health risk levels [13,14], environmental management strategies [15,16], and corresponding remediation technologies [17]. Obviously, eutrophication and heavy metal pollution play key roles in urban lake health. Environmental assessment of eutrophication or heavy metals in urban lakes is an important reference for identifying the degree of pollution and formulating pollution prevention strategies. Unfortunately, most research on lake risk is often evaluated from a single angle for toxic metals pollution or eutrophication using the standard comparison method [18–20] for both, the comprehensive trophic level index (TLI) [21,22], and the health risk assessment for toxic metals [23,24]. Moreover, the above deterministic methods probably lead to biased or unreliable assessments due to the randomness and fuzziness in environment systems caused by natural change and human activities [25–28]. Frankly, when there are different theoretical foundations, the evaluation results and conclusions differ to some extent, which makes it challenging for decision-makers to make scientific and synthetic management decision under consideration of separate assessment methods. Therefore, in order to implement appropriate environmental management strategies and measures, it is important to explore a feasible lake health assessment method synthetically considering lake eutrophication and corresponding health risks posed by heavy metals, together with evaluation of systematic uncertainty.

The objectives of this study were (i) to investigate and analyze the state of eutrophication and the health risk posed by heavy metals in the selected lakes from the total 143 lakes in the Chinese Wuhan city as a study case; (ii) to develop a fuzzy comprehensive lake health assessment method (FCLHAM) integrating quantitative eutrophication and health risk consideration; (iii) to test and verify FCLHAM by assessing the integrated lake health states of the studied lakes; and (iv) to introduce a novel hierarchical management system for lakes based on the FCLHAM results.

2. Materials and Methods

2.1. Study Area

Known for having over 100 lakes, Wuhan city (113°41′ E–115°05′ E; 29°58′ N–31°22′ N) is located in the interior of China, in the eastern part of Jianghan Plain. Located on the northern side of the Northern Tropic, the sub-tropical monsoon humid climate zone has a mean annual temperature of 15.8 °C–17.5 °C and an annual rainfall of 1150–1450 mm. Wuhan city has 143 lakes and the total area of lakes is 803.2 km^2, which is the highest of all Chinese cities. Because of the high ecological value and the large number of lakes, the local economy is increasing rapidly. Additionally, various functions of these lakes play significant roles in the development of the city. But, as a result of the excessive increase of economy and other factors, the number of lakes is rapidly decreasing and some lakes are polluted to a certain degree, bringing adverse effects to sustainable urban development.

2.2. Sample Collection and Analysis Methods

One hundred and forty-three lakes in Wuhan (Figure S1) were divided into large lakes ($\geq$20 km^2), medium lakes (10 km^2–20 km^2), and small lakes ($\leq$10 km^2) with the proportion of 3%, 6%, and 91%, respectively. The large and medium-sized lakes were small in number and important to the development of Wuhan city, so they were all included in the typical lakes. Other typical lakes in the small lake category were selected by combing the region and function further. Based on regionalization, the lakes could be divided into Central District, Dongxihu District, Caidian District, Hannan District, Jiangxia District, Huangpi District, Xinzhou District, Economic Development Zone, and the High and New Technology Development Zone of Donghu lake. The lakes have five main water functions: regulation, irrigation, water supply, aquaculture or planting, landscape or entertainment, and reservation. Most lakes had multiple functions, especially large lakes, while small lakes were relatively simple in function. According to the preliminary classification of region and function, the small lakes were selected comprehensively to contain all city regions and functions. The selection method is presented in Figure S2. Therefore, under the premise of ensuring the results are representative and scientific, a total of 21 typical lakes (Figure 1) were screened and selected from 143 lakes in Wuhan city based on comprehensive consideration of their characteristics of area, function, and region. Figure 1 was made based on our investigation's data and the ArcGIS software version 9.3 (Environmental Systems Research Institute, Redlands, CA, USA) (https://www.arcgis.com/index.html). The 21 selected typical lakes were Tangxun Lake (L1), Niushan Lake (L2), Baoxie Lake(L3), Luhu Lake (L4), Shenshan Lake (L5), Qingling Lake (L6), Huangjia Lake (L7), Yanxi Lake (L8), Wuhu Lake (L9), Houhu Lake (L10), Jingyin Lake (L11), Donghu Lake (L12), Nanhu Lake (L13), Moshui Lake (L14), Tanghu Lake (L15), Lanni Lake (L16), Zhongshan Lake (L17), Guanlian Lake (L18), Tonghu Lake (L19), Xiaozha Lake (L20), and Hougong Lake (L21).

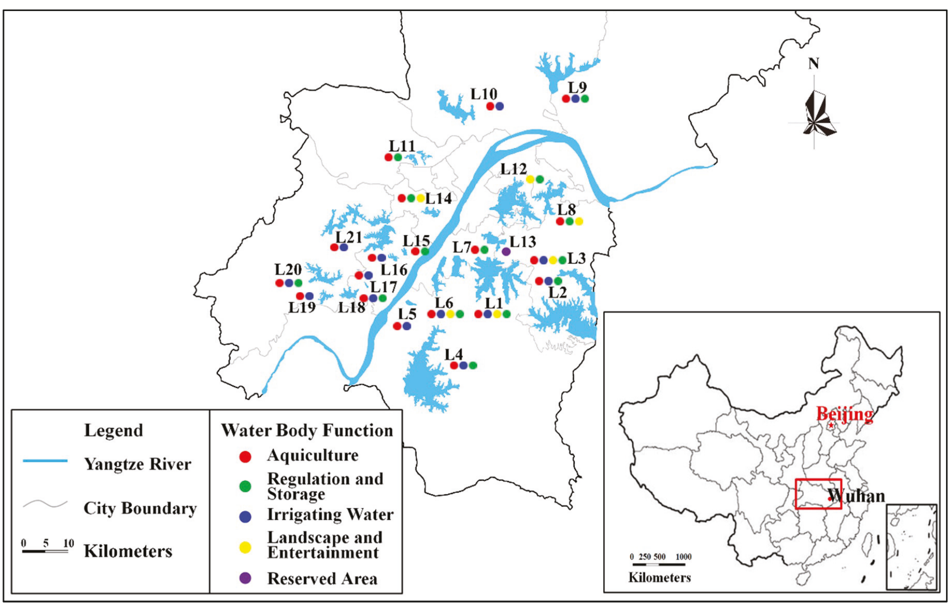


Figure 1. The geolocation of the 21 lakes in Wuhan (ArcGIS software version 9.3 (https://www.arcgis.com/index.html)).

The design layout and sampling methods of lake water samples were strictly according to the Chinese Technical Specifications Requirements for Monitoring of Surface Water and Waste Water

(HJ91-2002) and the Chinese Water Quality Technical Regulation on the Design of Sampling Programs (HJ495-2009), combined with the area size and hydrological characteristics of the studied lake. All water samples were collected into polytetrafluoroethylene bottles which were rinsed with lake water at least three times before using. Afterwards, water samples were acidified to pH 1–2 with H_2SO_4 and then kept in thermostats with ice bags, and transferred to the laboratory within 24 h. The water temperature (T), pH, dissolved oxygen (DO), electrical conductivity (EC), and Chl-*a* were measured by a portable water quality analyzer (HQ40d, HACH, Loveland, CO, USA) and the transparency (SD) was measured by the lead method in situ. In the laboratory, the concentration of TP was measured by Molybdenum Antimony Spectrophotometry (GB11893-89, China), the concentration of TN was measured by Alkaline Potassium Persulfate Digestion UV Spectrophotometry (HJ636-2012, China), and the concentration of COD_{Mn} was measured by Water Quality—Determination of Permanganate Index (GB11892-89, China). In addition, the total amounts of Cu, Zn, Cr, Cd, Mn, Fe, Ni, and Pb were measured with Atomic Absorption Spectroscopy (AAS ZEEnit 700P, Jena, Germany), the total amount of As and Hg were measured by Atomic Fluorescence Spectrometry (AFS-9730, Haiguang Instrument Co. Ltd., Beijing, China), both of the determinations were according to Water Quality—Digestion of Total Metals-Nitric Acid Digestion Method (HJ677–2013, China) and Water Quality—Determination of Mercury, Arsenic, Selenium, Bismuth and Antimony—Atomic Fluorescence Spectrometry (HJ694-2014, China).

To ensure the accuracy and reliability of analysis, parallel samples and blank samples were used to analyze error and at least 10% of each batch of samples was taken as parallel samples. If the relative deviation of the results was less than 20%, the analysis results were considered reliable. The analysis results of blank samples should be lower than the detection limits, so as to eliminate the pollution that might generate between processing procedures and determinations. The standard curve would be drawn for each sample analysis, and the correlation coefficient of the standard curve was not below 0.995.

2.3. Comprehensive Trophic Level Index (TLI) Method

TLI is one of the comprehensive eutrophication evaluation methods taking chlorophyll a (Chl-*a*), total phosphorus (TP), total nitrogen (TN), transparency (SD), and permanganate index (COD_{Mn}) as the evaluation indicators [28,29]. It was widely used in trophic state assessment of lakes and rivers due to its diversity and applicability in evaluation indicators [30,31]. TLI takes Chl-*a* as the benchmark parameter, obtaining the corresponding weights of all parameters depending on the correlation degree between the benchmark parameter and the other parameters, and then obtains the TLI by a weighted algorithm [29]. The TLI model is as follows [32]:

$$TLI(\Sigma) = \sum_{j=1}^{n} W_j \times TLI(j) \tag{1}$$

where $TLI(\Sigma)$ is comprehensive trophic level index. W_j represents the corresponding weight of parameter *j*. $TLI(\Sigma)$ represents trophic state index of parameter *j*. *n* is the numbers of evaluation parameters.

With Chl-*a* as the benchmark parameter, the normalized correlation weight of parameter *j* is as follows [32]:

$$W_j = {r_{ij}}^2 / \sum_{j=1}^{n} {r_{ij}}^2 \tag{2}$$

where W_j represents the corresponding weight of parameter *j*. r_{ij} represents the correlation coefficient between benchmark parameter and parameter *j*. *n* is the numbers of evaluation parameters. Based on a eutrophication survey of Chinese lakes, the correlation coefficients r_{ij} between the benchmark parameter (Chl-*a*) and other parameters are $r_{\text{Chl-}a} = 1$, $r_{TN} = 0.82$, $r_{TP} = 0.84$, $r_{SD} = 0.83$ and $r_{CODMn} = 0.83$ [28,32].

Trophic level indexes of each parameter are calculated as Equations (3)–(7).

$$TLI(\mathrm{Chl}-\alpha) = 10 \times (2.5 + 1.086 \ln \rho_{\mathrm{Chl}-\alpha}) \quad (3)$$

$$TLI(TP) = 10 \times (9.436 + 1.624 \ln \rho_{TP}) \quad (4)$$

$$TLI(TN) = (10 \times 5.453 + 1.694 \ln \rho_{TN}) \quad (5)$$

$$TLI(SD) = 10 \times (5.118 - 1.94 \ln \rho_{SD}) \quad (6)$$

$$TLI(COD_{MN}) = 10 \times \left(0.109 + 2.66 \ln \rho_{COD_{MN}}\right) \quad (7)$$

where $\rho_{\text{Chl-}a}$ represents concentration of Chl-*a* (mg/m^3), and ρ_{TP} ρ_{TN} $\rho_{COD_{Mn}}$ represent concentrations of *TP*, *TN*, and COD_{Mn} (mg/L), respectively. ρ_{SD} represents transparency (m). The trophic state of the lakes is graded using continuous numbers from 0 to 100, as shown in Table 1 [28,32].

Table 1. Classification of eutrophication levels.

TLI(Σ)	[0, 30)	[30, 50]	(50, 60]	(60, 70]	(70, 100)
Grades	I	II	III	IV	V
Trophic state	Oligotrophic	Medium trophic	Light eutrophication	Medium eutrophication	Severe eutrophication

2.4. Health Risk for Heavy Metals in Lakes

Health risk assessment is described as processes used to estimate event probability and probable degree of adverse health effects over a specific period [33–35]. Risk level of environmental pollutants to human beings depends on the body's exposure dose to the pollutants and the toxicity of the pollutants. There are two main pathways for human exposure to trace elements in water: ingestion and dermal absorption, ignoring exposure via inhalation [35,36]. The exposure dose can be calculated by Equations (8) and (9) [37,38].

$$ADD_{ing} = \frac{C_W \times IR \times EF \times ED}{BW \times AT} \quad (8)$$

where ADD_{ing} (μg/(kg·day)) represents the exposure dose through ingestion. In this study, the ingestion mainly refers to the intake through water from studied lakes. C_w is the mean concentration of trace element in water (μg/L). *IR* is the intake rate of water, including direct drinking rate and indirect drinking rate (L/day). *EF* is the exposure frequency to pollutants (day/year). *ED* is the exposure duration, and it means the length of time over which contact with the contaminant lasts (year). *BW* represents the body weight (kg). *AT* is the average time (day). For carcinogenic risk, *AT* is the average life expectancy of people [37,38].

$$ADD_{derm} = \frac{C_W \times SA \times K_P \times ET \times EF \times ED \times 10^{-3}}{BW \times AT} \quad (9)$$

where ADD_{derm} (μg/(kg·day)) represents the exposure dose through dermal absorption. *SA* is the exposure area of skin (cm^2). K_p is the dermal permeability coefficient of pollutants in water (cm/h), in this study, 0.001 cm/h for Cu, Cd, and As, 0.0001 cm/h for Pb, 0.002 cm/h for Cr, and 0.0006 cm/h for Zn [14,35], and *ET* is the exposure time (h/day). In this study, *ET* is 0.6 h/day. For the meanings of C_w, *EF*, *ED*, *BW*, and *AT*, please refer to Equation (8).

The health risks caused by environmental pollutants can be divided into carcinogenic risk and non-carcinogenic risks according to their properties. In general, carcinogens are of greater risk than non-carcinogens. Therefore, the risk of cancer caused by lake water is used as the assessment medium of lake health risk.

Carcinogenic risk is the product of daily exposure dose and cancer slope factor, which is shown in Equation (10). Under the assumption that there is no antagonism and synergism between pollutants, the integrated carcinogenic risk can also be identified as the sum of carcinogenic risks exposure by various pollutants via different pathways as shown in Equation (11). The EPA believes that carcinogenic risk value of human being is acceptable within 1×10^{-4}, while the maximum acceptable risk value recommended by International Commission on Radiological Protection (ICRP) is 5×10^{-5} [14]. The significant difference between the two evaluation standards may mislead the decision makers in their final judgment. Furthermore, it should be noted that there is currently no official and uniform standard of acceptable risk value in China and many developing countries, which may lead to uncertainty and incomparability among different decision-makers. Therefore, risk classification was carried out in this study in order to make the evaluation results clearer and more intelligible. Risk levels were rated as five levels based on the Delphi method, assessment criteria of USEPA and ICRP, as well as existing research (Table 2) [38].

$$CR_i = ADD_i \times CSF_i \tag{10}$$

$$CR = \sum_i^n CR_i \tag{11}$$

where CR_i is the carcinogenic risk of trace elements through ingestion or dermal absorption, dimensionless. ADD_i (µg/(kg·day)) is the daily exposure dose of carcinogenic pollutants. CSF_i (kg·day/µg) is the cancer slope factor of carcinogenic pollutants. CR is the sum of CR_i. i is the pathways of exposure. n is the kinds of trace elements.

Table 2. Levels and values of assessment standards.

Risk Grades		Range of *CR*	Acceptability
Grade I	Extremely low risk	$<10^{-6}$	Completely accept
Grade II	Low risk	$[10^{-6}, 10^{-5})$	Do not mind about the risk
Grade III	Low-medium risk	$[10^{-5}, 5 \times 10^{-5})$	Care about the risk
Grade IV	Medium risk	$[5 \times 10^{-5}, 10^{-4})$	Care about the risk and willing to invest
Grade V	High risk	$>10^{-4}$	Pay attention to the risk and take action to solve it

2.5. Fuzzy Comprehensive Lake Health Assessment Method (FCLHAM)

As one of the most important human habitats, lakes provide a variety of service functions, and their health status is closely related to the survival and development of human beings. How to comprehensively and scientifically evaluate the health status of lakes is becoming an important concern in the field of environmental science and ecology. And it has extremely important application value for the monitoring and management of lakes. Therefore, a fuzzy assessment method needs to be developed to efficiently identify comprehensive lake health states. Based on *TLI*, Health Risk Assessment framework for heavy metals, and fuzzy theory, it is of significance to explore a novel assessment method synthetically considering heavy metals' health risk, eutrophication risk, and fuzziness of the assessment system. Based on the fuzzy comprehensive evaluation theory [39,40], the comprehensive lake health state was defined as follows:

$$Risk = f(Risk_E, Risk_H) \tag{12}$$

where $Risk_E$ represents the eutrophication risk of studied lakes, which is assessed by TLI. $Risk_H$ represents the health risk for the heavy metals of studied lakes. And f represents the comprehensive lake health state calculation functions.

Fuzzy language recognition theory in fuzzy mathematics was used to identify the risk in this model. The comprehensive lake health state can be calculated as follows:

$$Risk = \widetilde{C} \cdot \widetilde{R} = (C_1, C_2) \cdot \begin{pmatrix} A_1 & A_2 & A_3 & A_4 \\ B_1 & B_2 & B_3 & B_4 \end{pmatrix} \tag{13}$$

where $\widetilde{C} \cdot \widetilde{R}$ characterize the f in Equation (13). C is the weight values of $Risk_E$ and $Risk_H$. C_1 and C_2 was determined as 0.4 and 0.6 by the Delphi method, which indicated that the risk of lakes depends on the health risk of heavy metals more than the eutrophication risk by expert advices [41,42]. $\widetilde{R}$ is membership matrix for levels of $Risk_E$ and $Risk_H$. A_1, A_2, A_3, A_4, and A_5 represent membership degrees of five levels of $Risk_E$ (Table 1), and B_1, B_2, B_3, B_4, and B_5 represent membership degrees of five levels of $Risk_H$ (Table 2).

Therefore, the comprehensive lake health state can be represented as a matrix with one row and five columns. The calculated comprehensive lake health state were divided into five levels as follows: (1) level I, low risk; (2) level II, moderate risk; (3) level III, considerable risk; (4) level IV, high risk; (5) level V, very high risk. The membership degree of each assessment factor plays a key role in the fuzzy comprehensive risk assessment, which is the basis of the foundation of comprehensive fuzzy assessment. According to Tables 1 and 2, the membership function of $Risk_E$ and $Risk_H$ was established, and the membership degree of each level can be calculated by the following formulas [39,40]:

(1) $Risk_E$

$$u_1(r) = \begin{cases} 1, r \in [0,30) \\ (50-r)/20, r \in [30,50) \\ 0, r \in [50,+\infty) \end{cases} \tag{14}$$

$$u_2(r) = \begin{cases} 0, r \in [0,30)\ or\ [60,+\infty) \\ (r-30)/20, r \in [30,50) \\ (60-r)/10, r \in [50,60) \end{cases} \tag{15}$$

$$u_3(r) = \begin{cases} 0, r \in [0,50) or [70,+\infty) \\ (r-50)/10, r \in [50,60) \\ (70-r)/10, r \in [60,70) \end{cases} \tag{16}$$

$$u_4(r) = \begin{cases} 0, r \in [0,60) \\ (r-60)/10, r \in [60,70) \\ 0, r \in [70,+\infty) \end{cases} \tag{17}$$

$$u_5(r) = \begin{cases} 0, r \in [0,70) \\ (r-70)/30, r \in [70,100) \\ 1, r \in [100,+\infty) \end{cases} \tag{18}$$

(2) $Risk_H$

$$u_1(r) = \begin{cases} 1, r \in [0,10^{-6}) \\ (10^{-5}-r)/(10^{-5}-10^{-6}), r \in [10^{-6},10^{-5}) \\ 0, r \in [10^{-5},+\infty) \end{cases} \tag{19}$$

$$u_2(r) = \begin{cases} 0, r \in [0,10^{-6})\ or\ [5\times10^{-5},+\infty) \\ (r-10^{-6})/(10^{-5}-10^{-6}), r \in [10^{-6},10^{-5}) \\ (5\times10^{-5}-r)/(4\times10^{-5}), r \in [10^{-5},5\times10^{-5}) \end{cases} \tag{20}$$

$$u_3(r) = \begin{cases} 0, r \in [0,10^{-5})\ or\ [10^{-4},+\infty) \\ (r-10^{-5})/(4\times10^{-5}), r \in [10^{-5},5\times10^{-5}) \\ (10^{-4}-r)/(10^{-4}-5\times10^{-5}), r \in [5\times10^{-5},10^{-4}) \end{cases} \tag{21}$$

$$u_4(r) = \begin{cases} 0, r \in [0, 5 \times 10^{-5}) \\ (r - 5 \times 10^{-5}) / (10^{-4} - 5 \times 10^{-5}), r \in [5 \times 10^{-5}, 10^{-4}) \\ 0, r \in [10^{-4}, +\infty) \end{cases} \quad (22)$$

$$u_5(r) = \begin{cases} 0, r \in [0, 10^{-4}) \\ 1, r \in [10^{-4}, +\infty) \end{cases} \quad (23)$$

3. Results and Discussion

3.1. Basic Parameters and Trace Element Concentrations in Surface Water from Studied Lakes

Table 3 provides a statistical summary of the water quality parameters measured in the 21 lakes in Wuhan. The pH of lakes was basically between 9 and 10, and the highest was in L16, the lowest in L1. According to the five levels of standard limited values stipulated in the Chinese Environmental Quality Standard for Surface Water (GB3838-2002), the concentrations of Chl-*a* and EC in lakes were within their target water quality standard. With the exception of L1 and L3, which were slightly below the target, the DO of the other 19 lakes met the target water quality standards. However, the concentrations of TN and TP did not reach the standard. The TP concentrations of 18 of the 21 lakes did not reach the target water quality. There are four lakes (L10; L13; L15, L21) with higher TP concentrations than the Class V water quality standards, far from reaching the target water quality standards of GB3838-2002. The TN concentrations ranged from 2.16 to 5.48 mg·L^{-1}, with the highest value in L15 and lowest in L4. Not only did all the studied lakes not meet their target water quality standards, but they also did not reach the limit of the Class V water quality standard of GB3838-2002, indicating that 21 lakes have been heavily polluted with nitrogen and have shown significant eutrophication pollution characteristics. In addition, the SD values of 16 of the 21 lakes did not reach the target water quality standard limit, and the COD_{Mn} concentrations of 12 lakes did not reach the target water quality standard limit, which also showed the serious pollution of lake water.

The detected heavy metal concentrations of 21 lakes in Wuhan are listed in Table 4. Cr was not detected in water samples. The concentration range of As is 1.237–12.148 μg·L^{-1}. Except for L20, the concentrations of As in the studied lakes was within the permissible limits of USEPA, WHO, and Chinese Ministry of Health (2007). The concentrations of Cd were all within the permissible limits of China, WHO, and USEPA. The results indicated that the concentration of carcinogenic heavy metals in lakes was not very high, and other related risks of lake health need to be further explored.

Table 3. Basic parameters and target water quality of 21 lakes in Wuhan.

Lake		pH	DO ($mg \cdot L^{-1}$)	Cond ($\mu s \cdot cm^{-1}$)	Chl-*a* ($mg \cdot m^{-3}$)	TP ($mg \cdot L^{-1}$)	TN ($mg \cdot L^{-1}$)	SD (m)	COD_{Mn} ($mg \cdot L^{-1}$)	Target Quality
L1		8.94	4.41	409.33	4.51	0.19	3.84	0.12	6.94	III
L2		9.07	6.76	147.80	1.10	0.03	2.29	0.25	6.23	II
L3		9.10	4.14	228.50	1.45	0.08	2.50	0.25	7.15	II
L4		9.33	8.24	239.00	1.62	0.04	2.16	0.27	5.86	II
L5		9.46	8.99	217.80	0.44	0.07	3.13	0.23	4.65	/
L6		9.54	11.23	332.50	9.74	0.24	3.96	0.15	5.50	III
L7		9.46	6.50	279.00	0.44	0.18	3.65	0.21	7.65	III
L8		9.40	8.02	322.50	2.40	0.06	2.26	0.20	6.66	III
L9		9.48	10.34	172.20	0.69	0.11	2.42	0.24	6.08	III
L10		9.58	16.98	312.50	8.27	0.25	3.61	0.11	6.62	III
L11		9.78	13.17	336.00	2.12	0.18	3.56	0.24	6.46	IV
L12		9.96	14.53	276.00	5.29	0.12	3.33	0.25	6.31	III
L13		9.90	12.00	506.50	7.39	0.38	4.05	0.19	8.19	IV
L14		9.56	9.10	335.50	9.09	0.15	3.96	0.14	7.27	IV
L15		9.88	14.23	414.00	6.53	0.44	5.48	0.05	8.58	IV
L16		10.02	12.55	293.00	6.36	0.13	3.65	0.10	7.42	IV
L17		9.74	13.09	321.00	14.28	0.15	2.48	0.10	7.73	/
L18		9.64	11.50	326.00	7.72	0.03	2.56	0.09	7.35	IV
L19		9.64	9.35	244.00	1.32	0.01	4.80	0.15	6.58	III
L20		9.50	8.30	246.00	3.02	0.16	3.44	0.14	7.19	III
L21		9.67	11.37	313.50	3.80	0.22	3.12	0.13	6.92	III
Chinese standards [a]	Class I	6–9	≥7.5	≤2000	≤1.0	≤0.01	≤0.2	≥15.0	≤2.0	/
	Class II		≥6.0		≤4.0	≤0.025	≤0.5	≥4.0	≤4.0	
	Class III		≥5.0		≤10	≤0.05	≤1.0	≥2.5	≤6.0	
	Class IV		≥3.0		≤50	≤0.1	≤1.5	≥1.5	≤10	
	Class V		≥2.0		≤65	≤0.2	≤2.0	≥0.5	≤15	

[a] The standard values of the Chinese Environmental Quality Standards for Surface Water (GB3838-2002).

Table 4. The concentrations of heavy metals in 21 lakes in Wuhan.

Lake	Cr ($\mu g \cdot L^{-1}$)	As ($\mu g \cdot L^{-1}$)	Cd ($\mu g \cdot L^{-1}$)
L1	<0.1 [b]	3.353	0.033
L2	<0.1	3.165	0.056
L3	<0.1	6.133	0.012
L4	<0.1	3.521	0.035
L5	<0.1	1.237	0.014
L6	<0.1	5.759	0.022
L7	<0.1	2.359	0.017
L8	<0.1	3.951	0.015
L9	<0.1	2.194	<0.1
L10	<0.1	5.534	0.01
L11	<0.1	3.057	0.009
L12	<0.1	3.872	0.004
L13	<0.1	5.812	0.02
L14	<0.1	3.516	0.018
L15	<0.1	4.792	0.021
L16	<0.1	5.095	0.026
L17	<0.1	2.941	0.005
L18	<0.1	2.089	0.021
L19	<0.1	2.062	0.009
L20	<0.1	12.148	0.161
L21	<0.1	4.125	0.005
WHO [a]	50	10	3
USEPA [b]	100	10	5
Chinese standards [c]	50	10	5

[a] WHO, 2008; [b] USEPA, 2009; [c] Chinese Ministry of Health, 2007.

3.2. Eutrophication State Analysis of the Studied Lakes

The comprehensive trophic level index (TLI) of 21 lakes in Wuhan are calculated and shown in Table S1. L2 and L5 had relatively better water quality with TLI values of 49.14 and 49.44 (corresponding to the medium trophic condition). Unfortunately, most of the studied lakes presented eutrophication to different extents. Particularly, the TLI of lake L15 was higher than the other lakes, which indicated the most severe eutrophication status. The TLI values in the lakes L3, L4, L7, L8, L9, L11, L12, and L19 varied from 49.14 to 72.96, indicating light eutrophication. Furthermore, the other 10 lakes reached medium eutrophication states, decreasing in the order of L13 > L10 > L17 > L6 > L14 > L16 > L1 > L21 > L20 > L18. The main exceeding standard factors of each lake include TN, TP, SD, and COD_{Mn}. This result accords with the data listed in Table 3, which indicates that the cause of eutrophication in lakes is the result of the combined effect of reducing substances and nutrients in the water [14,43].

3.3. Health Risk Assessment for Heavy Metals in the Studied Lakes

As the results show in Table S2, the values of CR_{Cr} and CR_{Cd} in all lakes were both below 10^{-6}, indicating that there was no carcinogenic risk of Cr and Cd. However, there was a certain carcinogenic risk of As because the CR_{As} values of 21 lakes exceeded 10^{-6}, decreasing in the order of: medium risk (L20 > L3 > L13 > L6 > L10 > L16) > low risk (L15 > L21 > L8 > L12 > L4 > L14 > L1 > L2 > L11 > L17 > L7 > L9 > L18 > L19 > L5). The maximum and minimum values of CR_{As} were 2.44×10^{-5} and 2.48×10^{-6}, which were 24.4 times and 2.5 times than the lowest risk limit. Moreover, Table S2 indicated that the risk levels of lakes L3, L6, L10, L13, L16, and L20 were Grade III, which indicates that the pollution of carcinogenic heavy metals in these lakes should be given certain attention by relevant local departments. The risk level of the remaining lakes was Grade II, indicating that carcinogenic heavy metal pollution was not very serious, and not a current health risk concern.

3.4. Results of the Fuzzy Comprehensive Lake Health Assessment Method (FCLHAM)

To sum up, we can see that some differences surely existed between the results of eutrophication and health risk assessment, which may confuse the decision-maker because these methods unilaterally focus on evaluating the eutrophication level or the health risks of heavy metals of the lakes. For example, the eutrophication level of L15 is very high and belongs to the severely eutrophic category. However, the result of health risk assessment of L15 is low risk. FCLHAM assigns weights to the results of TLI and CR in order to evaluate comprehensive lake health states more comprehensively and accurately. Then, comprehensive lake health state can be calculated. According to $u_i(r)$ arithmetic calculation, the assessment matrix is shown in Table 5. The five numbers contained in the assessment matrix represent the degree to which Risk belongs to the level represented by $u_i(r)$. According to the maximum membership principle, the closer the membership of $u_i(r)$ to 1, the higher the degree to which Risk belongs to the level represented by $u_i(r)$. As is shown in Table 5, average comprehensive lake health state decreased in the sequence of L20 (considerate risk level) > L1~L17, L19, L21 (moderate risk level) > L18 (low risk level). It indicated that L20 needs increased human, material, and financial resource investment and to be given priority regarding its governance. L1~L17, L19, and L21 need more attention, governance, and oversight to maintain their current state. Although L5's risk level was close to the low risk level, its low risk level (0.508) and moderate risk level (0.492) were too close to each other. Under the principle of maximum risk protection, L5 was determined as moderate risk. If the absolute difference between the memberships of two adjacent risk levels is less than 10%, the final risk level can be determined as the higher level.

Table 5. The fuzzy matrix of $Risk_E$ and $Risk_H$.

Lake	$Risk_E$	$Risk_H$	Risk	Membership Level	
L1	(0,0,0.583,0.417,0)	(0.345,0.655,0,0,0)	(0.207,0.393,0.2332,0.1668,0)	II	Moderate risk
L2	(0.043,0.957,0,0,0)	(0.375,0.625,0,0,0)	(0.2422,0.7578,0,0,0)	II	Moderate risk
L3	(0,0.585,0.415,0,0)	(0,0.941,0.059,0,0)	(0,0.7986,0.2014,0,0)	II	Moderate risk
L4	(0,0.932,0.068,0,0)	(0.307,0.693,0,0,0)	(0.1842,0.7886,0.0272,0,0)	II	Moderate risk
L5	(0.028,0.972,0,0,0)	(0.828,0.172,0,0,0)	(0.508,0.492,0,0,0)	I	Low risk
L6	(0,0,0.484,0.516,0)	(0,0.958,0.042,0,0)	(0,0.5748,0.2188,0.2064,0)	II	Moderate risk
L7	(0,0.443,0.557,0,0)	(0.576,0.424,0,0,0)	(0.3456,0.4316,0.2228,0,0)	II	Moderate risk
L8	(0,0.494,0.506,0,0)	(0.222,0.778,0,0,0)	(0.3732,0.6644,0.2024,0,0)	II	Moderate risk
L9	(0,0.764,0.236,0,0)	(0.622,0.378,0,0,0)	(0.3732,0.5324,0.0944,0,0)	II	Moderate risk
L10	(0,0,0.342,0.658,0)	(0,0.971,0.029,0,0)	(0,0.5826,0.1542,0.2632,0)	II	Moderate risk
L11	(0,0.134,0.866,0,0)	(0.425,0.575,0,0,0)	(0.255,0.3986,0.3464,0,0)	II	Moderate risk
L12	(0,0.052,0.948,0,0)	(0.246,0.754,0,0,0)	(0.1476,0.4732,0.3792,0,0)	II	Moderate risk
L13	(0,0,0.308,0.692,0)	(0,0.956,0.044,0,0)	(0,0.5736,0.1496,0.2768,0)	II	Moderate risk
L14	(0,0,0.485,0.515,0)	(0.317,0.683,0,0,0)	(0.1902,0.4098,0.194,0.206,0)	II	Moderate risk
L15	(0,0,0,0,0.0986)	(0.031,0.969,0,0,0)	(0.0186,0.5814,0,0,0.03944)	II	Moderate risk
L16	(0,0,0.51,0.49,0)	(0,0.991,0.009,0,0)	(0,0.5946,0.2094,0.196,0)	II	Moderate risk
L17	(0,0,0.358,0.642,0)	(0.453,0.547,0,0,0)	(0.2718,0.3282,0.1432,0.2568,0)	II	Moderate risk
L18	(0,0,0.829,0.171,0)	(0.634,0.366,0,0,0)	(0.3804,0.2196,0.3316,0.0684,0)	I	Low risk
L19	(0,0.886,0.114,0,0)	(0.646,0.354,0,0,0)	(0.3876,0.5668,0.0456,0,0)	II	Moderate risk
L20	(0,0,0.99,0.01,0)	(0,0.621,0.379,0,0)	(0,0.3726,0.6234,0.004,0)	III	Considerate risk
L21	(0,0,0.698,0.302,0)	(0.189,0.811,0,0,0)	(0.1134,0.4866,0.2792,0.1208,0)	II	Moderate risk

Through comparative analysis, we can see: (i) only the heavy eutrophication of L15 also had a high health risk for heavy metals, with a level II risk assessment, indicating that the risk should be noticed. Comparing the results of FCLHAM, L15 belongs to the level II (moderate risk) category, which should attract attention. (ii) All 10 lakes at moderate eutrophication and 8 lakes in mild eutrophication were all rated at level II or level III, while the corresponding results of the FCLHAM levels are also level II or level III, except for L18, whose level was I (low risk). (iii) All three assessment results of L2 and L5, which were in the medium level of nutrition, were level II. That also provided further proof of the reliability of FCLHAM. The first two methods of L5 were level II, and the results of FCLHAM was level

I. The examples of L5 and L18 illustrated the more hierarchical and scientific results of FCLHAM than the other two evaluation methods, which makes up for the deficiency of the deterministic assessment.

3.5. Classification and Control Countermeasures of Studied Lakes

It is understood that lakes in Wuhan are currently subject to cross-sectoral management by functional classification. With the exception of the water sector, other departments, such as fishery, transportation, environmental protection, health, land, forestry, tourism, health, and other related departments, have a certain management function for lakes. However, the pollution treatment of lakes depends on the management and promotion of the environmental protection department, so it is necessary to consider the lakes' function and nutritional status in the lake classification. Based on the results of FCLHAM, the 21 studied lakes were classified into three categories: general management, enhanced management, and priority management (Figure 2). Figure 2 was made based on the results of FCLHAM with ArcGIS software version 9.3 (https://www.arcgis.com/index.html). The low risk lake (L18) which was rated level I corresponds to common management; enhanced management corresponds to the moderate risk lakes (L1–L17, L19, L21); and priority management corresponds to the considerable risk lake (L20).

According to the characteristics of each kind of lake, we give countermeasures of lake management, specifically: for the common management lakes, its eutrophication level and health risk for heavy metal are both low, so at present, it does not need a great deal of manpower and material resources for remediation, but observation and supervision policies should be implemented so that the lake stays in good condition. The enhanced management lakes, whose risk is in the middle, should be given attention and some treatment measures should be taken. The priority management lakes, whose levels of eutrophication and the health risks for heavy metals are both high, require greater human, material, and financial resource expenditure for remediation to prevent threats to the physical health of the surrounding residents. In addition, the corresponding policies and the necessary source investigation should also be carried out to protect the lake after treatment.

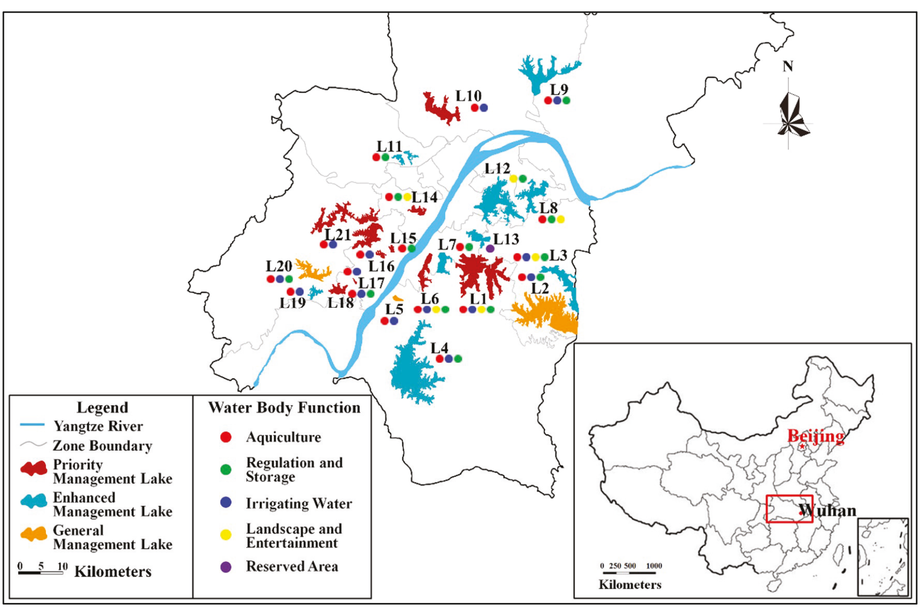


Figure 2. Management classification chart of 21 investigated lakes (ArcGIS software version 9.3 (https://www.arcgis.com/index.html)).

If the 143 lakes in Wuhan were classified by the "area-region-function" classification, they would be assigned to the same category as the representative lakes of the same type. At this point, we will attribute all of Wuhan's lakes to the three types of common management lakes, enhanced management lakes, and priority management lakes. Depending on the characteristics of each type, a targeted approach to different types of management for each type of lake is a more efficient way to manage many of Wuhan's lakes. Based on FCLHAM, a novel hierarchical management system for urban lake health based on lake characteristics classification was obtained (Figure 3). Here, we have tested and verified the rationality, efficiency, and science of this novel hierarchical management system for innovative lake management of Wuhan. Therefore, this management mode can serve as an effective reference for the environmental management of urban lakes both at home and abroad, to manage urban lake health hierarchically and efficiently.

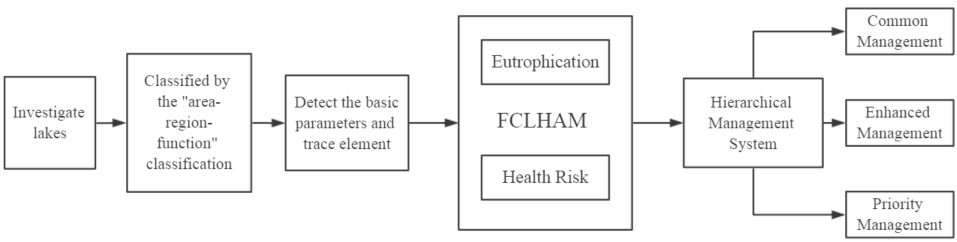


Figure 3. Workflow of the established hierarchical management system.

4. Conclusions

FCLHAM was established to evaluate lake health states more comprehensively and accurately by integrating quantitative eutrophication and health risk considerations. To test and verify FCLHAM, Wuhan was taken as an example to carry out a practical test. The state of eutrophication and the health risk posed by heavy metals in 21 of the 143 lakes in Wuhan city were investigated and analyzed. Under two different evaluation methods, the results of the same lake were different, and some even deviated greatly, such as L15 and L20, which is very disadvantageous for the managers of urban lake health to administer effective lake management. FCLHAM solves this problem for decision makers and offers the evaluation results of comprehensive consideration of the state of eutrophication and the health risk posed by heavy metals. The evaluation results of FCLHAM are as follows: L20 (considerate risk level) > L1–L17; L19; L21 (moderate risk level) > L18 (low risk level). According to the results, the studied lakes were grouped into three categories (general management lakes, enhanced management lakes, and priority management lakes) and effective protection and management measures were provided with respect to the characteristics of each type of lake. According to the characteristics of lake clustering, all lakes in Wuhan are classified into the above-mentioned three types. According to the characteristics of each type, solutions are put forward for the management each kind of lake. Therefore, FCLHAM offers a novel hierarchical management system for urban lake health based on lake characteristics classification, which can serve as an effective reference for the environmental management of urban lakes both at home and abroad.

Supplementary Materials: The following are available online at http://www.mdpi.com/1660-4601/15/12/2617/s1, Table S1: Eutrophication evaluation results of 21 studied lakes, Table S2: Health risk indicators for carcinogenic health for 21 studied lakes, Figure S1: Distribution map of lake in Wuhan city, Figure S2: The workflow of typical lake selection method.

Author Contributions: T.W. and J.Y. contributed to prepared datasets, performed the statistical analysis, and drafted the manuscript. J.M. contributed to prepared datasets and interpretation of the analysis. F.L. organized this study, conducted the study design, and revised the manuscript. C.L. contributed to interpretation of the analysis, and revision of the manuscript. Y.C. prepared datasets and performed the statistical analysis. S.C. contributed to revision of the manuscript. J.Z. contributed to the result analysis and English polish of the manuscript. Y.Q. contributed to interpretation of the analysis.

Funding: This study was financially supported by the Humanities and Social Sciences Foundation of Ministry of Education of China (Youth Fund: 17YJCZH081), the Nature Science Foundation of Hubei Province (2018CFB722), and the Soft Science Project Foundation of Hubei Province (2018ADC144).

Acknowledgments: The authors would like to express their deep gratitude to Master Zhenzhen Qiu and everyone else who helped us about this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mcdowell, R.W.; Moreau, P.; Salmon-Monviola, J.; Durand, P.; Leterme, P.; Merot, P. Contrasting the spatial management of nitrogen and phosphorus for improved water quality: Modelling studies in New Zealand and France. *Eur. J. Agron.* **2014**, *57*, 52–61. [CrossRef]
2. Ferreira, J.G.; Bricker, S.B.; Simas, T.C. Application and sensitivity testing of a eutrophication assessment method on coastal systems in the United States and European Union. *J. Environ. Manag.* **2007**, *82*, 433–445. [CrossRef] [PubMed]
3. Xu, Q.J.; Zhu, Y.Z.; Zheng, B.H.; Zhao, W. Comparative study on lake eutrophication control standards in Eastern and Yunnan-Guizhou lake region of China. *China Environ. Sci.* **2011**, *31*, 2046–2051. (In Chinese)
4. Smith, V.H.; Wood, S.A.; Mcbride, C.G.; Atalah, J.; Hamilton, D.P.; Abell, J. Phosphorus and nitrogen loading restraints are essential for successful eutrophication control of Lake Rotorua, New Zealand. *Inland Waters* **2016**, *6*, 273–283. [CrossRef]
5. Smith, V.H.; Schindler, D.W. Eutrophication science: Where do we go from here? *Trends Ecol. Evol.* **2008**, *24*, 201–207. [CrossRef] [PubMed]
6. Vorosmarty, C.J.; Mclntyre, P.B.; Gessner, M.O.; Dudgeon, D.; Prusevich, A.; Green, P.; Glidden, S.; Bunn, S.E.; Sullivan, C.A.; Liermann, C.R.; et al. Global threats to human water security and river biodiversity. *Nature* **2010**, *467*, 555–561. [CrossRef] [PubMed]
7. Li, F.; Huang, J.H.; Zeng, G.M.; Yuan, X.; Li, X.; Liang, J.; Wang, X.; Tang, X.; Bai, B. Spatial risk assessment and sources identification of heavy metals in surface sediments from the Dongting Lake, Middle China. *J. Geochem. Explor.* **2013**, *132*, 75–83. [CrossRef]
8. Li, F.; Zhang, J.D.; Liu, C.Y.; Xiao, M.; Wu, Z. Distribution, bioavailability and probabilistic integrated ecological risk assessment of heavy metals in sediments from Honghu Lake, China. *Process Saf. Environ. Prot.* **2018**, *116*, 169–179. [CrossRef]
9. Liu, C.Y.; Zhang, J.D.; Li, F.; Yang, J.; Qiu, Z.; Cai, Y.; Zhu, L.; Xiao, M.; Wu, Z. Spatial health risk assessment and statistical source identification of the trace elements in surface water from the Honghu Lake, China. *J. Cent. South Univ.* **2018**, *25*, 1598–1611. [CrossRef]
10. Sun, C.; Chen, Z.L.; Zhang, G.; Shi, G.; Bi, C. Health risk Assessment of heavy Metals in main drinking Water sources in Shanghai. *Res. Environ. Sci.* **2009**, *22*, 60–65.
11. Li, F.; Qiu, Z.Z.; Zhang, J.D.; Liu, C.; Cai, Y.; Xiao, M.; Zhu, L. Temporal variation of major nutrients and probabilistic eutrophication evaluation based on stochastic-fuzzy method in Honghu Lake, Middle China. *Sci. China Tech. Sci.* **2018**. [CrossRef]
12. Liu, X.P.; Jiang, J.Y.; Yan, Y.; Dai, Y.; Deng, B.; Ding, S.; Su, S.; Sun, W.; Li, Z.; Gan, Z. Distribution and risk assessment of metals in water, sediments, and wild fish from Jinjiang River in Chengdu, China. *Chemosphere* **2018**, *196*, 45–52. [CrossRef] [PubMed]
13. Väänänen, K.; Leppänen, M.T.; Chen, X.P.; Akkanen, J. Metal bioavailability in ecological risk assessment of freshwater ecosystems: From science to environmental management. *Ecotoxicol. Environ. Saf.* **2018**, *147*, 430–446. [CrossRef] [PubMed]
14. Li, F.; Qiu, Z.; Zhang, J.D.; Liu, C.; Cai, Y.; Xiao, M. Spatial distribution and fuzzy health risk assessment of trace elements in surface water from Honghu Lake. *Int. J. Environ. Res. Public Health* **2017**, *14*, 1011. [CrossRef] [PubMed]
15. Jiang, X.; Wang, W.; Wang, S.; Zhang, B.; Hu, J. Initial identification of heavy metals contamination in Taihu Lake, a eutrophic lake in China. *J. Environ. Sci.* **2012**, *24*, 1539–1548. [CrossRef]
16. Iqbal, J.; Shah, M.H.; Akhter, G. Characterization, source apportionment and health risk assessment of trace metals in freshwater Rawal Lake, Pakistan. *J. Geochem. Explor.* **2013**, *125*, 94–101. [CrossRef]

17. Copetti, D.; Valsecchi, L.; Capodaglio, A.G.; Tartari, G. Direct measurement of nutrient concentrations in freshwaters with a miniaturized analytical probe: Evaluation and validation. *Environ. Monit. Assess.* **2017**, *189*, 144–153. [CrossRef] [PubMed]
18. Wu, F.F.; Wang, X. Eutrophication Evaluation Based on Set Pair Analysis of Baiyangdian Lake, North China. *Procedia Environ. Sci.* **2012**, *13*, 1030–1036. [CrossRef]
19. Rajeshkumar, S.; Liu, Y.; Zhang, X.; Ravikumar, B.; Bai, G.; Li, X. Studies on seasonal pollution of heavy metals in water, sediment, fish and oyster from the Meiliang Bay of Taihu Lake in China. *Chemosphere* **2018**, *191*, 626–638. [CrossRef] [PubMed]
20. Yazidi, A.; Saidi, S.; Mbarek, N.B.; Darragi, F. Contribution of GIS to evaluate surface water pollution by heavy metals: Case of Ichkeul Lake (Northern Tunisia). *J. Afr. Earth Sci.* **2017**, *134*, 166–173. [CrossRef]
21. Huo, S.; Ma, C.; Xi, B.D.; Su, J.; Zan, F.; Ji, D.; He, Z. Establishing eutrophication assessment standards for four lake regions, China. *J. Environ. Sci.* **2013**, *25*, 2014–2022. [CrossRef]
22. Ding, J.; Cao, J.; Xu, Q.; Xi, B.; Su, J.; Gao, R.; Huo, S.; Liu, H. Spatial heterogeneity of lake eutrophication caused by physiogeographic conditions: An analysis of 143 lakes in China. *J. Environ. Sci.* **2015**, *30*, 140–147. [CrossRef] [PubMed]
23. Li, F.; Zhang, J.D.; Liu, W.; Liu, J.; Huang, J.; Zeng, G. An exploration of an integrated stochastic-fuzzy pollution assessment for heavy metals in urban topsoil based on metal enrichment and bioaccessibility. *Sci. Total Environ.* **2018**, *644*, 649–660. [CrossRef] [PubMed]
24. Capodaglio, A.G.; Muraca, A.; Becchi, G. Accounting for water quality effects of future urbanization: Diffuse pollution loads estimates and control in Mantua's lakes (Italy). *Water Sci. Technol.* **2003**, *47*, 291–298. [CrossRef] [PubMed]
25. Li, F.; Huang, J.; Zeng, G.; Yuan, X.; Liang, J.; Wang, X. Multimedia health impact assessment: A study of the scenario-uncertainty. *J. Cent. South Univ.* **2012**, *19*, 2901–2909. [CrossRef]
26. Hu, B.B.; Liu, B.Q.; Zhou, J.; Guo, J.; Sun, Z.; Meng, W.; Guo, X.; Duan, J. Health risk assessment on heavy metals in urban street dust of Tianjin based on trapezoidal fuzzy numbers. *Hum. Ecol. Risk Assess.* **2016**, *22*, 678–692. [CrossRef]
27. Li, L.; Xu, Z.R.; Zhang, C.; Bao, J.; Dai, X. Quantitative evaluation of heavy metals in solid residues from sub- and super-critical water gasification of sewage sludge. *Bioresour. Technol.* **2012**, *121*, 169–175. [CrossRef] [PubMed]
28. Li, Z.Z.; Li, X.D.; Li, F.; Gao, Z.; Yan, W.; Liang, J.; Zeng, G. Improved assessment model for comprehensive trophic state index based on dynamic cluster analysis and blind theory. *Chin. J. Environ. Eng.* **2015**, *9*, 2021–2026. (In Chinese)
29. Wang, J.L.; Fu, Z.S.; Qiao, H.X.; Liu, F.X. Assessment of eutrophication and water quality in the estuarine area of Lake Wuli, Lake Taihu, China. *Sci. Total Environ.* **2019**, *650*, 1392–1402. [CrossRef] [PubMed]
30. Lu, J.; Yang, Z.; Zhang, Y. Algae functional group characteristics in reservoirs and lakes with different trophic levels in Northwestern semi-humid and semi-arid regions in China. *J. Environ. Sci.* **2018**, *64*, 166–173. [CrossRef] [PubMed]
31. Zhi, G.Z.; Chen, Y.N.; Liao, Z.L.; Walther, M.; Yuan, X. Comprehensive assessment of eutrophication status based on Monte Carlo–triangular fuzzy numbers model: Site study of Dongting Lake, Mid-South China. *Environ. Earth Sci.* **2016**, *75*, 1011. [CrossRef]
32. China National Environmental Monitoring Center. *Evaluation Methods and Classification Technical Regulations for Eutrophication Assessment of Lakes (Reservoirs) ([2001]090)*; China National Environmental Monitoring Center: Beijing, China, 2001. (In Chinese)
33. Li, F.; Cai, Y.; Zhang, J. Spatial Characteristics, Health Risk Assessment and Sustainable Management of Heavy Metals and Metalloids in Soils from Central China. *Sustainability* **2018**, *10*, 91. [CrossRef]
34. Huang, J.; Li, F.; Zeng, G.; Liu, W.; Huang, X.; Xiao, Z.; Wu, H.; Gu, Y.; Li, X.; He, X.; et al. Integrating hierarchical bioavailability and population distribution into potential eco-risk assessment of heavy metals in road dust: A case study in Xiandao District, Changsha city, China. *Sci. Total Environ.* **2016**, *541*, 969–976. [CrossRef] [PubMed]
35. Li, F.; Zhang, J.D.; Jiang, W.; Liu, C.Y.; Zhang, Z.M.; Zhang, C.D.; Zeng, G.M. Spatial health risk assessment and hierarchical risk management for mercury in soils from a typical contaminated site, China. *Environ. Geochem. Health* **2017**, *39*, 923–934. [CrossRef] [PubMed]

36. Saha, N.; Rahman, M.S.; Ahmed, M.B.; Zhou, J.L.; Ngo, H.H.; Guo, W. Industrial metal pollution in water and probabilistic assessment of human health risk. *J. Environ. Manag.* **2017**, *185*, 70–78. [CrossRef] [PubMed]
37. US Environmental Protection Agency (USEPA). *Quantitative Uncertainty Analysis of Superfund Residential Risk Pathway Models for Soil and Groundwater: White Paper*; USEPA: Washington, DC, USA, 1996.
38. USEPA. *Risk Assessment Guidance for Superfund Volume I: Human Health Evaluation Manual (Part E, Supplemental Guidance for Dermal Risk Assessment)*; US Environment Protection Agency: Washington, DC, USA, 2004.
39. Zhang, Q.W.; Yang, X.H.; Zhang, Y.; Zhong, M. Risk assessment of groundwater contamination: A multilevel fuzzy comprehensive evaluation approach based on drastic model. *Sci. World J.* **2013**, *2013*, 610390. [CrossRef] [PubMed]
40. Liu, Y.L.; Huang, X.L.; Duan, J.; Zhang, H. The assessment of traffic accident risk based on grey relational analysis and fuzzy comprehensive evaluation method. *Nat. Hazards* **2017**, *88*, 1409–1422. [CrossRef]
41. Linstone, H.A.; Turoff, M. *The Delphi Method: Techniques and Applications*; Addison-Wesley Educational Publishers Inc.: Boston, MA, USA, 1975.
42. Steurer, J. The delphi method: An efficient procedure to generate knowledge. *Skelet. Radiol.* **2011**, *40*, 959–961. [CrossRef] [PubMed]
43. Liu, W.Z.; Zhang, Q.F.; Liu, G.H. Lake eutrophication associated with geographic location, lake morphology and climate in China. *Hydrobiologia* **2010**, *644*, 289–299. [CrossRef]

International Journal of
Environmental Research and Public Health

Article

Healthy Lifestyle: Relationship between Mediterranean Diet, Body Composition and Physical Fitness in 13 to 16-Years Old Icelandic Students

Pablo Galan-Lopez [1], Francis Ries [1], Thordis Gisladottir [2], Raúl Domínguez [3] and Antonio J. Sánchez-Oliver [4,5,*]

1 Faculty of Educational Sciences, University of Seville, 41013 Sevilla, Spain; pgalan1@us.es (P.G.-L.); fries@us.es (F.R.)
2 Research Center for Sport and Health Sciences, School of Education, University of Iceland, 105 Reykjavík, Iceland; thg@hi.is
3 Faculty of Health Sciences of Universidad Isabel I, Universidad Isabel I, 09004 Burgos, Spain; raul.dominguez@ui1.es
4 Faculty of Sports Sciences of Universidad Pablo de Olavide, Universidad Pablo de Olavide, 41013 Sevilla, Spain
5 Human Motricity and Sports Performance Area, University of Seville, 41013 Sevilla, Spain
* Correspondence: asanchez@upo.es; Tel.: +34-656-305-480

Received: 27 September 2018; Accepted: 21 November 2018; Published: 24 November 2018

Abstract: Childhood and adolescent obesity are currently among the greatest challenges for public health. Physical activity, physical fitness, and adherence to the Mediterranean diet (MD), representing powerful indicators of healthy lifestyles, are shown as determinant factors in the prevention and treatment of obesity. The aim of the present study has been to analyse the relationship between health-related physical fitness components, body composition, and adherence to MD in 387 Icelandic adolescents of 13–16-years old (54% boys). The ALPHA Fitness Test was used to measure physical fitness and body composition. The KIDMED questionnaire was used to assess the adherence to MD among participants. Associations between variables were tested according to gender and age using linear regression models and analysis of variance. Participants with high/medium adherence to MD showed significantly higher endurance scores in both the boys and the girls. Gender differences were found. The boys in high/medium MD categories had significantly lower fat percentages and ran a 4×10 m sprint faster than the girls. The girls scored higher than the boys in endurance and speed-agility tests. It can be concluded that a high and medium adherence to MD is associated with high and very high endurance in both the girls and the boys.

Keywords: adolescents; physical fitness; Mediterranean diet; body composition

1. Introduction

Noncommunicable diseases, cardiovascular diseases, cancer, respiratory diseases, obesity, and diabetes are the main causes of death worldwide [1]. These deaths are caused, in large part, by poor diet, physical inactivity, and the consumption of alcohol and tobacco [2,3], thus establishing lifestyle as a good predictor of health and morbidity and mortality [4,5].

Childhood and adolescence are important periods of life, since many physiological and psychological transformations take place at these ages. Similarly, lifestyle and healthy or unhealthy behaviours are established during these years, which may influence adult behaviour and health status [6,7]. In children and young people, the prevalence of overweight and obesity has increased in recent years [8], causing premature deaths and increasing the risk of cardiovascular and metabolic diseases [9]. Childhood and adolescent obesity are a transcendental challenge for public health, both in its magnitude and in its consequences [10]. Lifestyle intervention is the most common treatment strategy in children and adolescents with obesity. Although there are studies with significant effects of lifestyle treatment in children and adolescents with obesity [11], additional research is required to draw conclusions about this type of population [12,13].

Among the habits that lead to a healthy lifestyle are physical activity (PA) and eating healthily [14]. Defined as 'any bodily movement produced by skeletal muscles that results in energy expenditure' [15], PA is a vital part of a healthy lifestyle and has been extensively documented and associated with health benefits in children and adolescents. Some of the benefits include reductions in blood cholesterol, hypertension, metabolic syndrome, obesity, and associated health problems such as diabetes mellitus type 2, cardiovascular diseases, or bone health problems in this population [16–22]. Moreover, physical fitness, mainly cardiorespiratory fitness, muscular fitness, and motor ability, have been shown to be powerful markers of health in young people [23]. Several PA tracking studies have been analysing activity patterns in childhood and adolescence and the risk of maintaining sedentary behaviours [24]. A low level of PA is associated with metabolic risk factors in young people that can also persist until adulthood [25].

Food intake in adolescence is a significant predictor of intake in adulthood [26]. In the context of overall dietary patterns, the MD has been accepted as one of the healthiest dietary patterns in the world [27], showing significant protection concerning mortality and morbidity when there is high adherence to it [28–31]. The MD has shown health benefits by reducing cardiovascular diseases, type 2 diabetes, certain types of cancer, and some neurodegenerative diseases [27,28]. Studies focusing on the influence of the MD in children and adolescents have increased in recent years [32]. In general, adherence to the MD has been shown to be associated with physical benefits and high levels of health-related quality of life, reducing the different factors associated with obesity, among others [33,34].

There is growing evidence that health behaviours are grouped. For example, the combination of regular PA and healthy eating habits helps to maintain and improve health and physical and mental well-being [35]. During youth, healthy eating combined with regular PA increases the likelihood of a healthy pattern of consistent physical maturation [36]. In addition, there are independent and combined associations between physical fitness, physical activity, and adherence to the MD with quality of life related to health in children, adolescents, and adults [33,37–40], with significant improvements in joint interventions [41–43].

However, despite all the benefits mentioned above, the current data show unhealthy patterns of eating and PA during the transition from childhood to adolescence [44], substantially contributing to the global burden of morbidity, mortality, and disability [45], and increasing the prevalence of overweight and obesity at those ages [46]. As mentioned before, obesity and hypertension, among others, have been largely attributed to unhealthy diets and a decrease in PA [47]. Thus, PA and nutrition are shown as fundamental pillars in the prevention and control of obesity [48,49].

The aim of the present study is to analyse the independent associations between health-related physical fitness components, body composition, and adherence to the MD of adolescents from Reykjavik (Iceland).

2. Materials and Methods

2.1. Study Sample and Design

The study design was cross sectional, including 13–16-year-old students from the capital area of Reykjavik, Iceland. A total of 439 participants (235 boys and 204 girls) were selected for the present research. Finally, 387 adolescents (209 boys (54%), Mage = 13.57, SD = 1.13 and 178 girls (46%), Mage = 13.38, SD = 1.14) took part in it, which yields a participation rate of 88.15%.

The subjects were recruited from two different schools in Reykjavik: Seljaskóli School (54.2%) and Ölduselsskóli School (45.8%). The inclusion criteria for this research were: male and female participants between 13 and 16 years old who had submitted the signed informed consent by their parents/guardians. Related to their health status, participants in the present investigation were those subjects who, due to their state of health, could participate regularly in the subject of Physical Education. The participants did not have any type of cognitive or physical/motor limitations The National Committee of Bioethics of Iceland approved the present study (Ref.: VSNb2017030026/03.01), which ensured that all the procedures related to research involving human beings would be carried out in complete safety. Written informed consent was obtained from the parents of the participants. The adolescents were asked for verbal consent, while they were informed that participation was voluntary and that they could leave the study at any time.

2.2. Instruments

Adherence to a Mediterranean Diet Questionnaire (KIDMED): The KIDMED questionnaire, previously validated, was used to evaluate the adherence to MD in adolescents (http://www.aulamedica.es/nh/pdf/9828.pdf). The questionnaire consists of 16 items, where 12 questions assume a positive score for adherence to MD (consumption of yogurt and dairy products, consumption of legumes, use of olive oil, consumption of vegetables, fruits, fish, cereals, rice, pasta, and nuts) and four questions assume a negative score (consumption of fast food, not having breakfast daily, consuming sweets several times a day, consuming industrial pastries). Affirmative answers to questions that represent a negative connotation in relation to the MD are worth −1 point and affirmative answers to questions that represent a positive aspect in relation to the MD are worth +1 point. Negative answers do not score [50–52].

Therefore, this index can range from 0 (minimum adherence) to 12 (maximum adherence). The sum of the values of this questionnaire gives rise to the KIDMED index, which is classified into three categories: From 8 to 12: an optimal MD (high adherence); from 4 to 7: a need to improve the food pattern to adapt it to the Mediterranean model (average adherence); and from 0 to 3: a very low-quality diet (low adherence) [50–52].

Alpha Fitness test battery: Physical fitness and anthropometric variables were assessed by a modified version of the extended ALPHA fitness test battery, (Ref: 2006120)). Skin folds were omitted for limited time reasons and the 4 × 10 m speed-agility test was added to the version in order to have more information about physical fitness. The protocol marked on the ALPHA-Fitness Battery for measurement was followed at all times [53].

Body composition: The height of the subjects was recorded barefoot with an accuracy of 0.1 cm using a portable stadiometer (Seca 213, Seca, Hamburg, Germany). The weight of the participants was measured with an accuracy of 0.10 kg, the subjects wore light clothing, and a portable electronic scale was used. Body fat percentage (BF%) was measured by bioelectrical impedance (Tanita Inner Scan BC-543, Tanita, Tokyo, Japan). Body mass index (BMI) was calculated from the ratio of body weight (kg) to body height (m^2). Waist circumference was measured with a non-flexible measurement tape (Seca 201, Seca, Hamburg, Germany) with the adolescent standing upright and with an accuracy of 0.1 cm. The measuring point was the narrowest part of the space between the lowest rib and the anterior superior iliac spine at the end of normal expiration.

Cardiovascular fitness was assessed with the multistage 20 m shuttle run test (Leger et al., 1988) [54]. In this test, the participants had to run a distance of 20 m, adjusting their speed to the rhythm of the audio signals that were emitted from a previously recorded CD. The subjects finished the test when they could not reach the line a second time concurrent with the audio, or when the subject stopped due to fatigue. The initial speed was 8.5 km/h, with this being increased by 0.5 km/h per minute [55].

Lower body explosive muscle strength was assessed using a standing long jump. The participants, placed behind the jumping line with their feet together, pushed hard and jumped as far away as possible, contacting the ground with both feet simultaneously and in a vertical position. The distance was measured from the rearmost heel to the jumping line and was always performed on a non-slippery surface.

Upper body maximal muscle strength was measured by means of handgrip strength using a hand dynamometer with an adjustable grip (TKK 5401 Grip D, Takey, Tokyo, Japan). The examiner showed the correct way of execution and adjusted the grip measure according to the size of the hand [56]. The test was performed twice, and the best result was recorded, calculating the average of the two hands. The subjects were verbally encouraged to "squeeze as hard as possible" and to exert the maximum effort for at least two seconds (s). Speed-agility was tested using the 4 × 10 shuttle run test. The examiner showed the correct way of execution. The test was performed twice, and the best result was recorded (s). The participants had to run, as fast as possible, the distance between the two lines placed 10 m away, change a series of sponges (three times), and run back to the starting line.

2.3. Methodology

All participants performed the test battery and the KIDMED questionnaire during the time corresponding to their physical education classes. The different tests were organised as a circuit and the participants carried out all the tests consecutively, except for the cardiovascular fitness test, which was performed by several students at the same time and on a different day. The development and performance of all physical tests lasted one hour for each class of 20–25 students.

2.4. Data Analysis

Quantitative variables were presented as means (M) and standard deviations (SD), while frequencies and percentages (%) were used for qualitative variables. After verifying the normality of the variables by means of the Kolmogorov-Smirnoff test, a Student-T test for independent samples was used to perform a comparative analysis of the quantitative variables of body composition and physical fitness between the boys and the girls. In addition, to check for possible differences in the proportion of subjects that are in the different categories in the % of body fat for an adolescent population established by Moreno et al. (2006) [57], a Chi-square test was performed to check for the possible differences between genders. In order to analyse the degree of adherence to an MD (low, medium or high), both in the boys and the girls, in relation to the different variables of body composition and physical fitness, after checking homoskedasticity by means of Levene's test, a one-way ANOVA and, in the case of statistically significant differences, a Bonferroni post-hoc test, was performed. Furthermore, a Chi-square test was performed to check for the possible differences in each of the questions that compose the KIDMED questionnaire grounded on the normative levels of body fat % established by Moreno et al. (2006) [57]. The level of statistical significance was set as $p < 0.05$. All statistical analyses were carried out using the SPSS statistical package (version 18.0, SPSS Inc., Chicago, IL, USA).

3. Results

The characteristics of the participants, including age, body composition, and physical fitness, are shown in Table 1. Although anthropometric data show a statistically higher weight (+7.5%, $p = 0.010$) and height (+3.7%, $p < 0.001$) in the boys, there were no statistically significant differences in relation to BMI ($p = 0.241$) between the boys and the girls. In addition, the boys showed a lower percentage of body fat (−33.4%, $p < 0.001$), but a significantly higher waist circumference (+5.8%, $p < 0.001$).

However, by relativising the average data of the boys and the girls to the different levels of body composition established by Moreno et al. (2006, 2007) [57,58], it is confirmed that the values of BMI, waist and % of body fat are classified as 'medium' in each of the parameters (see Figure 1).

Table 1. Anthropometric characteristics and physical fitness variables (n = 387).

Variables	Mean ± SD	Boys (n = 209)	Girls (n = 178)	p-Value
Age (year)	13.48 ± 1.14	13.57 ± 1.13	13.38 ± 1.14	0.097
Weight (kg)	57.29 ± 13.53	59.20 ± 14.21	55.05 ± 12.36	0.010 *
Height (m)	1.64 ± 0.10	1.67 ± 0.11	1.61 ± 0.81	<0.001 *
BMI (kg/m^2)	22.26 ± 4.72	22.17 ± 4.91	22.37 ± 4.49	0.241
Body fat (%)	21.37 ± 8.79	17.37 ± 8.44	26.07 ± 6.61	<0.001 *
Waist (cm)	72.05 ± 10.12	73.91 ± 10.41	69.86 ± 9.34	<0.001 *
Handgrip (kg)	25.95 ± 7.01	28.18 ± 7.90	23.33 ± 4.58	<0.001 *
Jump (cm)	171.94 ± 1.51	182.94 ± 28.60	159.02 ± 25.51	<0.001 *
4 × 10 m (s)	11.80 ± 1.40	11.53 ± 1.59	12.11 ± 1.06	<0.001 *
Endurance (CRF)	6.73 ± 2.52	7.52 ± 2.62	5.80 ± 2.04	<0.001 *
KIDMED Index	5.83 ± 2.31	5.82 ± 2.18	5.84 ± 2.46	0.940

Note: SD = Standard Deviation, BMI = Body Mass Index, Waist = Waist Circumference, CRF = Cardiorespiratory Fitness. * Expresses statistically significant differences between boys and girls (p < 0.05).

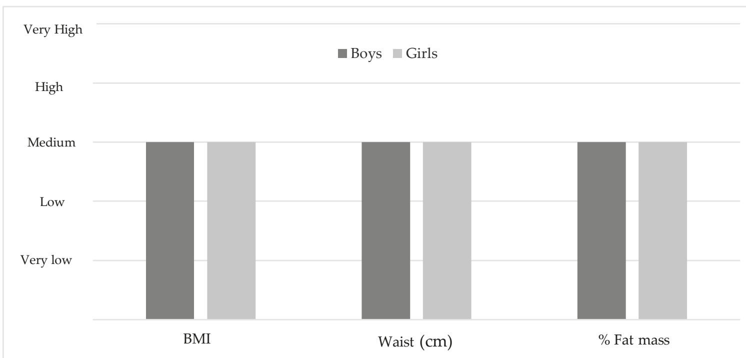


Figure 1. Classification of the mean levels of body composition variables based on the average levels established by Moreno et al. (2006, 2007) [57,58].

In relation to the physical fitness components, the boys showed significantly higher performances in each of the tests (p < 0.001). However, when categorising the average levels of the boys and the girls to the different levels of physical fitness established by Ortega et al. (2011) [59], the average value of the boys is established as an average value and the girls show high levels. In endurance, the boys show a high level and the girls very high (see Figure 2).

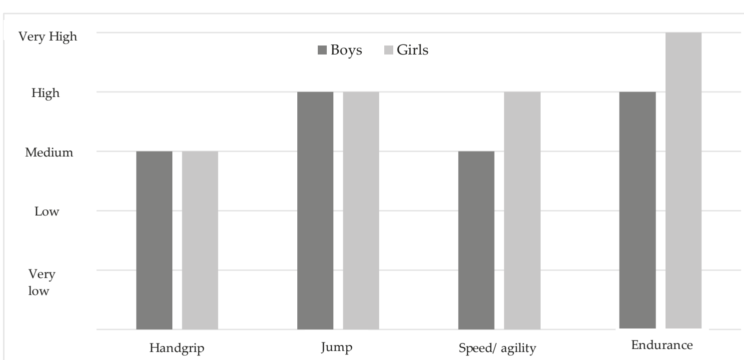


Figure 2. Mean levels of physical fitness variables based on the average levels established by Ortega et al. (2011) [59].

In addition, Table 1 highlights no differences in the KIDMED-index when comparing both the boys and the girls. Analysing the participants in the different categories established based on the percentage of body fat %, statistically significant differences were observed between both ($p = 0.003$), where more girls showed high and very high values and a lower % ranked in low and medium (see Figure 3).

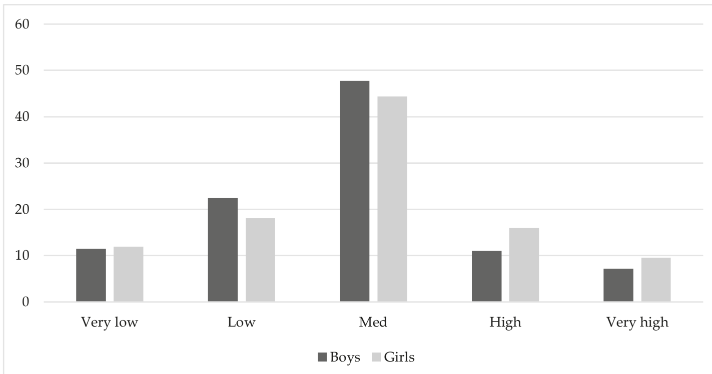


Figure 3. Proportion of boys and girls in each of the Alpha Fitness categories based on the percentage of body fat.

When performing a stratification of the sample based on the degree of adherence to MD, 14.99% showed low adherence, 60.72% an average level, and 24.29% a high level. When comparing body composition in relation to the degree of adherence to MD, it was found that, although there were no statistically significant differences in girls, boys showed a higher % of body fat among those who had a low adherence (21.84%) in comparison to a medium (16.79%) or a high adherence (16.21%) ($p = 0.006$) (see Table 2). When categorising the average levels with the classification established by Moreno et al. (2006, 2007) [57,58], the boys and the girls with different levels of adherence to MD present a medium level in the % of body fat and waist circumference (see Figures 4 and 5). On the contrary, when comparing the normative levels of the mean levels of these variables with the categorisation established by Moreno et al. (2006, 2007), it is verified that both genders present average levels in waist circumference and body fat percentage [57,58]. However, in the BMI variable in the boys, even though there were no differences ($p > 0.05$), the average level of subjects with low adherence to MD is rated as a high level, whereas those who show a medium or high adherence present an average level.

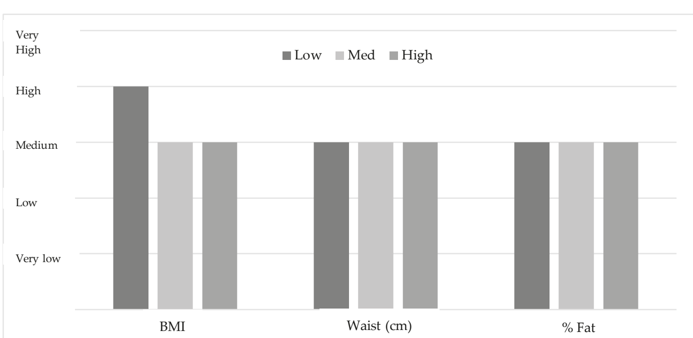


Figure 4. Mean levels of body composition variables in the boys based on the grouping by Moreno et al. (2006, 2007) [57,58].

Table 2. Anthropometric characteristics and physical fitness variables of the sample stratified according to adherence to MD in both boys and girls (Data presented as Mean ± SD).

Variables	Boys				Girls			
	Low (n = 58/387)	Med (n = 235/387)	High (n = 94/387)	p-Value	Low (n = 58/387)	Med (n = 235/387)	High (n = 94/387)	p-Value
Age (year)	13.37 ± 1.30	13.61 ± 1.10	13.58 ± 1.13	0.562	13.18 ± 0.98	13.29 ± 1.11	13.07 ± 1.11	0.678
Weight (kg)	61.15 ± 15.62	59.72 ± 14.51	56.80 ± 12.48	0.332	57.34 ± 9.90	55.31 ± 12.78	52.83 ± 12.67	0.308
High (m)	1.64 ± 0.10	1.67 ± 0.11	1.66 ± 0.11	0.338	1.62 ± 0.60	1.61 ± 0.09	1.60 ± 0.07	0.586
BMI (kg/m^2)	23.90 ± 6.06	22.07 ± 4.71	21.41 ± 4.52	0.080	22.34 ± 5.01	22.52 ± 4.40	22.01 ± 4.43	0.824
Body fat (%)	21.84 ± 10.01 [a]	16.79 ± 7.92	16.21 ± 8.03	0.006 *	26.50 ± 6.52	25.93 ± 6.69	26.16 ± 6.60	0.919
Waist (cm)	76.95 ± 11.75	73.87 ± 10.68	72.24 ± 8.56	0.142	70.48 ± 8.57	70.29 ± 9.54	68.33 ± 9.36	0.479
Handgrip average (kg)	27.99 ± 8.43	28.81 ± 7.88	26.75 ± 7.59	0.286	24.36 ± 3.69	23.40 ± 4.47	22.45 ± 5.31	0.225
Jump (cm)	177.2 ± 31.4	185.0 ± 27.7	181.3 ± 29.1	0.362	155.75 ± 25.68	160.02 ± 26.42	158.62 ± 23.33	0.730
4 × 10 m (s)	12.42 ± 2.00 [a]	11.29 ± 1.21	11.59 ± 1.96	<0.002 *	12.46 ± 0.90	12.04 ± 1.16	12.07 ± 1.17	0.174
Endurance (CRF)	6.28 ± 2.64 [a]	7.68 ± 2.43	7.85 ± 2.88	<0.018 *	4.63 ± 1.32 [a]	6.04 ± 2.12	5.94 ± 1.99	0.004 *

Note: SD = Standard Deviation, BMI = Body Mass Index, Waist = Waist Circumference, CRF = Cardiorespiratory Fitness. * Expresses statistically significant differences between groups ($p < 0.05$). [a] Expresses statistically significant differences between Low and Med and High.

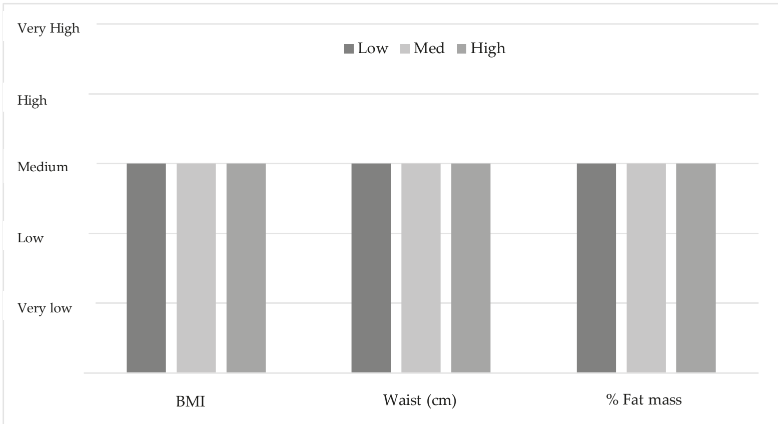


Figure 5. Mean levels of body composition variables in the girls based on the grouping by Moreno et al. (2006, 2007) [57,58].

The results of the fitness tests showed a higher performance in endurance, both for the boys and the girls, among the participants with a medium or high adherence to MD than those with low adherence ($p < 0.05$) (see Table 2). Moreover, when comparing the endurance classification of the average levels in the boys with a low adherence with the normative levels established by Ortega et al. (2011), these are medium [59], while those that have a medium or high adherence are high (see Figures 6 and 7).

Furthermore, the boys with a high and medium score in the MD have a significantly lower fat percentage and run faster on the 4 ×1 0 m sprint test in comparison to those with a low adherence to MD (p = 0.002), with no such difference being found in the girls. However, the average levels of the girls with a low level of adherence to MD were medium, whereas the girls with a medium or a high adherence were classified as high according to Ortega et al. (2011) (ALPHA fitness test) [59]. Regarding the jump test, no statistically significant differences were found in relation to adherence to MD. The same happened with the hand grip test, although the average levels of manual grip strength are classified as low in the boys and the girls with a high adherence to MD compared to the medium levels of participants with a low and medium adherence to MD.

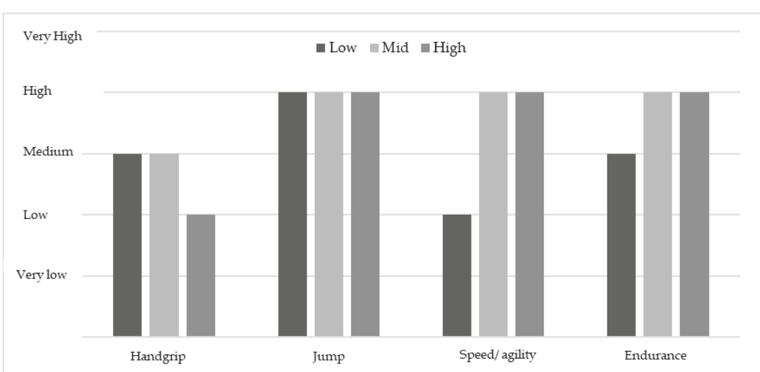


Figure 6. Mean levels of the physical fitness variables in the boys based on the average levels established by Ortega et al. (2011) [59].

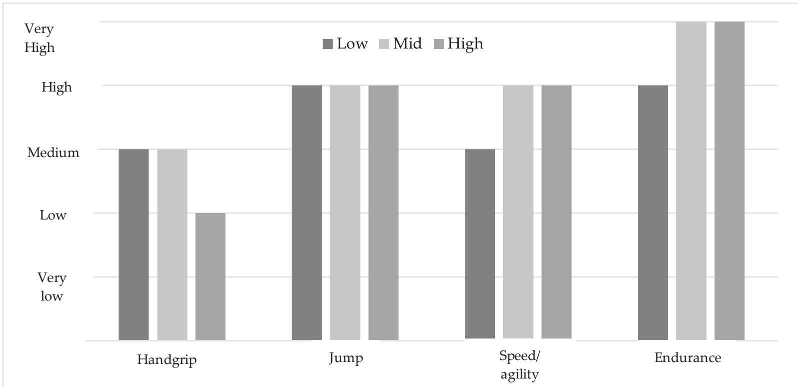


Figure 7. Mean levels of the physical fitness variables in the girls based on the average levels established by Ortega et al. (2011) [59].

4. Discussion

The present study is the first to analyse and describe the health-related physical fitness together with the adherence to MD of Icelandic adolescents.

As expected after analysing the body composition of the participants, significant differences were obtained between the boys and the girls in weight, height, % of body fat, and waist circumference (see Table 1). These results are similar to those obtained in several studies on the adolescent population [60–63], in which girls had higher levels of adiposity, whereas boys had higher weight, height, and waist circumference values. Both the boys and the girls show average values of BMI, % of body fat, and waist circumference (see Figure 1).

The weights, heights, and BMIs of the present study sample are comparable to the reference values provided by Ortega et al. (2011) [59]. Unlike what was found by Wärnberg et al. (2006) [64], there is no prevalence of obesity in the adolescent participants of this study as the BMI, waist circumference, and % of body fat are considered medium values (see Figure 3).

In relation to the performance in the physical fitness tests, the boys obtained significantly better results in manual dynamometry, the long jump, the 4×10 m sprint, and the endurance test (see Table 1). These results are similar to those of previous studies carried out [55,65–67]. However, when analysing the mean values of the different tests in relation to Alpha Fitness categories, the girls show the highest scores (see Figure 2). These results differ from several studies where boys score higher on these tests [66,68]. It is possible that three hours of mandatory PE classes and swimming lessons in secondary education may contribute to obtaining these results [69].

As mentioned before, MD is considered one of the healthiest dietary patterns [70], with benefits on a physical and mental level, among others [71–73]. The results of the present study show a low/poor MD adherence (14.99%) and are superior in an average (60.72%) and high adherence (24.29%). In addition, no significant differences were observed in the adherence to the diet according to the gender of the participants. These results are also superior to those of studies conducted in Mediterranean countries [74–76]. These results are different with respect to a recent study that analysed, in a similar way, the adherence to MD in non-Mediterranean countries, showing worse final results than ours with a poor (39%), medium (47.7%), and high (13.3%) adherence [73].

Several recent articles directly associate the adherence to MD with the weight and the BMI of the participants [77,78]. These data are similar to those obtained here, as the participants with the highest adherence to MD are those who also show regular weights and BMIs. Those with a low adherence to MD show a high BMI, although this is not significant. It should be noted that these values are significant when considering fat % (see Table 2), so it would be substantial to see which variable, BMI or fat %, has more importance in the three subgroups (low, medium, and high) that result from a better adherence to the MD (see Figures 4 and 5).

In contrast to the results obtained by Ozen et al. (2015) [79], which showed important differences in relation to a high and low adherence in the population analysed, with a clear tendency to abandon the MD, the results found in the present study display a tendency to maintain or even increase the MD patterns related to this type of diet, since a medium and high adherence gather 85% of the participants (see Figures 4 and 5).

These results from the KIDMED index are in line with the results obtained in other studies [77]. In addition, the results found are similar to studies carried out in southern European countries, where a large part of the sample is at a medium level of adherence to MD [80,81].

Regarding the relationship of the MD with the waist circumference, Bacopoulou et al. (2017), after studying more than 1600 subjects of a similar age to those of the present research, determined that the increase in adherence to MD was associated with a decrease in the perimeter of the waist, indicating a potential for school interventions to fight against abdominal obesity in adolescents [71]. This matches with the results of the present study and with the findings of Schröder et al. (2010), where more than 60% of the participating subjects presented a medium adherence to MD and medium, low, and very low waist circumference values [38].

As a novel aspect, the present study searched for relationships between health-related physical fitness, adherence to MD, and body composition (see Table 2 and Figures 6 and 7) in Icelandic adolescents. Significant differences were found in the tests of 4×10 m and endurance in the boys, and endurance in the girls with respect to those participants that show low adherence, compared with those with a medium or high adherence to MD. The disparity between genders in performance scores can be explained in the different processes of the adolescents' development. Girls experience development earlier than boys, which determines their ability to develop higher levels of strength, speed, and endurance [82].

The results mentioned in the previous paragraph are consistent with the conclusions of recent research. Muros et al. (2017), for example, found a positive relation between a high performance in the resistance test and a high adherence to MD [77]. Evaristo et al. (2018) not only demonstrated the relationship between a high adherence to MD and high levels of health-related physical fitness, but the subjects also showed high levels of health-related quality of life [70].

Despite the strength of the study, it is also important to acknowledge the limitations of the current research, which may restrict the generalisability of our findings and possible alternative interpretations. First, our data are cross-sectional and, therefore, do not enable us to infer the causal direction of our predictions. Nevertheless, they can be used as valuable indications to be considered for future research. Second, some of the data collected (KIDMED) were self-reported, which could lead to an error in the reports and recall bias due to the nature of the study. In addition, it must be borne in mind that the KIDMED questionnaire, although it was used to observe the adherence to MD does not contemplate the content and intake of nutrients consumed by the sample, which may be a confounding factor to be taken into account.

5. Conclusions

The adolescents participating in this study show medium/high levels of health-related physical fitness, with the girls obtaining slightly higher results. The participants' adherence to MD is classified as medium/high since 60% of the participants are in the middle level and almost 25% are in the high level.

The results found showed significant correlations between MD and the endurance test in the girls and the boys. A high adherence to MD also correlates with better results in endurance and agility speed tests in the boys.

This research shows the importance of developing and maintaining an adequate physical fitness and, together with a medium or high level of adherence to MD, it culminates in a better health-related quality of life in adolescents. Both the boys and the girls that showed a medium and high adherence to MD had the highest scores in the health-related physical fitness tests [50].

These results agree with those obtained in the Spanish and Portuguese adolescent population, since a high adherence to MD is related to higher levels of perceived quality of life, within which a good level of physical fitness is found. Moreover, a high adherence to MD is associated with a significant improvement in physical health and with lower obesity, a fact that is consistent with recent results [42].

Finally, this research appeals for the development of public health programmes, awareness campaigns, and the creation of PA and healthy eating environments for children and adolescents [64]. Not only an adequate diet is sufficient [65], but a minimum of daily physical activity practice is necessary to avoid the appearance of diseases derived from a sedentary lifestyle [40] and consequently, a poor quality of life [66].

Author Contributions: T.G., F.R. and A.J.S.-O. designed the study. P.G.-L. and T.G. collected the data. R.D. realised statistical work. P.G.-L., R.D., and A.J.S.-O. interpreted data. P.G.-L., R.D., and A.J.S.-O. drafted the manuscript. All the authors contributed to developing, editing, and approving the final version of the paper.

Funding: This work was supported by the fifth Research Programme 2013-2016 (PPI2015-IV.5 / PP2016-EBRV) of the University of Seville, Seville, Spain and the University of Iceland Research Fund, Reykjavik, Iceland.

Acknowledgments: The authors thank the participants and schools in the study.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Sánchez-Oliver, A.J.; García, C.M.; Gálvez-Ruiz, P.; González-Jurado, J.A. Mortality and economic expenses of cardiovascular diseases caused by physical inactivity in Spain. *J. Phys. Edu. Sport* **2018**, *18*, 1420–1426.
2. Diem, G.; Brownson, R.C.; Grabauskas, V.; Shatchkute, A.; Stachenko, S. Prevention and control of noncommunicable diseases through evidence-based public health: Implementing the NCD 2020 action plan. *Glob. Health Promot.* **2016**, *23*, 5–13. [CrossRef] [PubMed]
3. Riley, L.; Guthold, R.; Cowan, M.; Savin, S.; Bhatti, L.; Armstrong, T.; Bonita, R. The World Health Organization STEPwise Approach to Noncommunicable Disease Risk-Factor Surveillance: Methods, Challenges, and Opportunities. *Am. J. Public Health* **2016**, *106*, 74–78. [CrossRef] [PubMed]
4. Loef, M.; Walach, H. The combined effects of healthy lifestyle behaviors on all cause mortality: A systematic review and meta-analysis. *Prev. Med.* **2012**, *55*, 163–170. [CrossRef] [PubMed]
5. Tomba, E. Assessment of lifestyle in relation to health. In *The Psychosomatic Assessment: Strategies to Improve Clinical Practice*; Fava, G.A., Sonino, N., Wise, T.N., Eds.; Karger: Freiburg, Germany, 2012; Volume 32, pp. 72–96, ISBN 9783805598545.
6. Sawyer, S.M.; Afifi, R.A.; Bearinger, L.H.; Blakemore, S.-J.; Dick, B.; Ezeh, A.C.; Patton, G.C. Adolescence: A foundation for future health. *Lancet* **2012**, *379*, 1630–1640. [CrossRef]
7. Twig, G.; Yaniv, G.; Levine, H.; Leiba, A.; Goldberger, N.; Derazne, E.; Ben-Ami Shor, D.; Tzur, D.; Afek, A.; Shamiss, A.; et al. Body-Mass Index in 2.3 Million Adolescents and Cardiovascular Death in Adulthood. *N. Engl. J. Med.* **2016**, *374*, 2430–2440. [CrossRef] [PubMed]
8. Cattaneo, A.; Monasta, L.; Stamatakis, E.; Lioret, S.; Castetbon, K.; Frenken, F.; Manios, Y.; Moschonis, G.; Savva, S.; Zaborskis, A.; et al. Overweight and obesity in infants and pre-school children in the European Union: A review of existing data. *Obes. Rev.* **2010**, *11*, 389–398. [CrossRef] [PubMed]
9. Franks, P.W.; Hanson, R.L.; Knowler, W.C.; Sievers, M.L.; Bennett, P.H.; Looker, H.C. Childhood Obesity, Other Cardiovascular Risk Factors, and Premature Death. *N. Engl. J. Med.* **2010**, *362*, 485–493. [CrossRef] [PubMed]
10. Peng, W.; Goldsmith, R.; Berry, E.M. Demographic and lifestyle factors associated with adherence to the Mediterranean diet in relation to overweight/obesity among Israeli adolescents: Findings from the Mabat Israeli national youth health and nutrition survey. *Public Health Nutr.* **2017**, *20*, 883–892. [CrossRef] [PubMed]
11. Mårild, S.; Gronowitz, E.; Forsell, C.; Dahlgren, J.; Friberg, P. A controlled study of lifestyle treatment in primary care for children with obesity. *Pediatr. Obes.* **2013**, *8*, 207–217. [CrossRef] [PubMed]

12. Mead, E.; Brown, T.; Rees, K.; Azevedo, L.B.; Whittaker, V.; Jones, D.; Olajide, J.; Mainardi, G.M.; Corpeleijn, E.; O'Malley, C.; et al. Diet, physical activity and behavioural interventions for the treatment of overweight or obese children from the age of 6 to 11 years. *Cochrane Database Syst. Rev.* **2017**, *22*, CD012651. [CrossRef] [PubMed]
13. Al-Khudairy, L.; Loveman, E.; Colquitt, J.L.; Mead, E.; Johnson, R.E.; Fraser, H.; Olajide, J.; Murphy, M.; Velho, R.M.; O'Malley, C.; et al. Diet, physical activity and behavioural interventions for the treatment of overweight or obese adolescents aged 12 to 17 years. *Cochrane Database Syst. Rev.* **2017**, *22*, CD012691. [CrossRef] [PubMed]
14. Suárez-Carmona, W.; Sánchez-Oliver, A.; González-Jurado, J. Pathophysiology of obesity: Current view. *Rev. Chil. Nutr.* **2017**, *44*, 226–233.
15. Caspersen, C.J.; Powell, K.E.; Christensen, G.M. Physical Activity, Exercise, and Physical Fitness: Definitions and Distinctions for Health- Related Research. *Public Heal. Rec.* **1985**, *100*, 126–131.
16. Longmuir, P.; Colley, R.; Wherley, V.; Tremblay, M. Risks and benefits of childhood physical activity. *Lancet Diabetes Endocrinol.* **2014**, *2*, 861–862. [CrossRef]
17. Tremblay, M.S.; Gray, C.E.; Akinroye, K.; Harrington, D.M.; Katzmarzyk, P.T.; Lambert, E.V.; Liukkonen, J.; Maddison, R.; Ocansey, R.T.; Onywera, V.O.; et al. Physical Activity of Children: A Global Matrix of Grades Comparing 15 Countries. *J. Phys. Act. Heal.* **2014**, *11*, S113–S125. [CrossRef] [PubMed]
18. Tremblay, M.S.; Barnes, J.D.; González, S.A.; Katzmarzyk, P.T.; Onywera, V.O.; Reilly, J.J.; Tomkinson, G.R.; Global Matrix 2.0 Research Team. Introduction to the Global Matrix 2.0: Report Card Grades on the Physical Activity of Children and Youth Comparing 38 Countries. *J. Phys. Act. Health* **2016**, *13*, S85–S86. [CrossRef] [PubMed]
19. Berrington de Gonzalez, A.; Hartge, P.; Cerhan, J.R.; Flint, A.J.; Hannan, L.; MacInnis, R.J.; Moore, S.C.; Tobias, G.S.; Anton-Culver, H.; Freeman, L.B.; et al. Body-Mass Index and Mortality among 1.46 Million White Adults. *N. Engl. J. Med.* **2010**, *363*, 2211–2219. [CrossRef] [PubMed]
20. de Onis, M.; Blossner, M.; Borghi, E. Global prevalence and trends of overweight and obesity among preschool children. *Am. J. Clin. Nutr.* **2010**, *92*, 1257–1264. [CrossRef] [PubMed]
21. World Health Organization. *Global Action plan on Physical Activity 2018–2030: More Active People for a Healthier World*; WHO: Geneva, Switzerland, 2018.
22. Golden, N.H.; Abrams, S.A.; Committee on Nutrition. Optimizing bone health in children and adolescents. *Pediatrics* **2014**, *134*, e1229-43. [CrossRef] [PubMed]
23. Ortega, F.B.; Ruiz, J.R.; Castillo, M.J.; Sjöström, M. Physical fitness in childhood and adolescence: A powerful marker of health. *Int. J. Obes.* **2008**, *32*, 1–11. [CrossRef] [PubMed]
24. Janz, K.F.; Burns, T.L.; Levy, S.M. Tracking of activity and sedentary behaviors in childhood: The Iowa bone development study. *Am. J. Prev. Med.* **2005**, *29*, 171–178. [CrossRef] [PubMed]
25. Ekelund, U.; Anderssen, S.A.; Froberg, K.; Sardinha, L.; Andersen, L.B.; Brage, S.; European Youth Heart Study Group. Independent associations of physical activity and cardiorespiratory fitness with metabolic risk factors in children: The European youth heart study. *Diabetologia* **2007**, *50*, 1832–1840. [CrossRef] [PubMed]
26. Lake, A.A.; Mathers, J.C.; Rugg-Gunn, A.J.; Adamson, A.J. Longitudinal change in food habits between adolescence (11–12 years) and adulthood (32–33 years): The ASH30 study. *J. Public Health* **2006**, *28*, 10–16. [CrossRef] [PubMed]
27. Serra-Majem, L.; Roman, B.; Estruch, R. Scientific evidence of interventions using the Mediterranean diet: A systematic review. *Nutr. Rev.* **2006**, *64*, S27–S47. [CrossRef] [PubMed]
28. Sofi, F.; Abbate, R.; Gensini, G.F.; Casini, A. Accruing evidence about benefits of adherence to the Mediterranean diet on health: An updated systematic review and meta-analysis. *Am. J. Clin. Nutr.* **2010**, *92*, 1189–1196. [CrossRef] [PubMed]
29. Sofi, F.; Cesari, F.; Abbate, R.; Gensini, G.F.; Casini, A.; Transmission, P.P. Adherence to Mediterranean diet and health status: Meta-analysis. *BMJ* **2014**, *337*, 333–340. [CrossRef] [PubMed]
30. Martinez-Gonzalez, M.A.; Bes-Rastrollo, M.; Serra-Majem, L.; Lairon, D.; Estruch, R.; Trichopoulou, A. Mediterranean food pattern and the primary prevention of chronic disease: Recent developments. *Nutr. Rev.* **2009**, *67* (Suppl. 1), S111–S116. [CrossRef] [PubMed]
31. Schwingshackl, L.; Schwedhelm, C.; Galbete, C.; Hoffmann, G. Adherence to mediterranean diet and risk of cancer: An updated systematic review and meta-analysis. *Nutrients* **2017**, *9*, 1063. [CrossRef] [PubMed]

32. Novak, D.; Štefan, L.; Prosoli, R.; Emeljanovas, A.; Mieziene, B.; Milanović, I.; Radisavljević-Janić, S. Mediterranean Diet and its Correlates among Adolescents in Non-Mediterranean European Countries: A Population-Based Study. *Nutrients* **2017**, *9*, 177. [CrossRef] [PubMed]
33. Iaccarino Idelson, P.; Scalfi, L.; Valerio, G. Adherence to the Mediterranean Diet in children and adolescents: A systematic review. *Nutr. Metab. Cardiovasc. Dis.* **2017**, *27*, 283–299. [CrossRef] [PubMed]
34. Martin-Calvo, N.; Chavarro, J.E.; Falbe, J.; Hu, F.B.; Field, A.E. Adherence to the Mediterranean dietary pattern and BMI change among US adolescents. *Int. J. Obes.* **2016**, *40*, 1103–1108. [CrossRef] [PubMed]
35. World Health Organization. *Global Strategy on Diet, Physical Activity and Health Physical Inactivity: A Global Public Health Problem Physical Inactivity*; WHO: Geneva, Switzerland, 2017.
36. Sallis, J.F.; Glanz, K. Physical activity and food environments: Solutions to the obesity epidemic. *Milbank Q.* **2009**, *87*, 123–154. [CrossRef] [PubMed]
37. Evaristo, O.S.; Moreira, C.; Lopes, L.; Abreu, S.; Agostinis-Sobrinho, C.; Oliveira-Santos, J.; Póvoas, S.; Oliveira, A.; Santos, R.; Mota, J. Associations between physical fitness and adherence to the Mediterranean diet with health-related quality of life in adolescents: Results from the LabMed Physical Activity Study. *Eur. J. Public Health* **2018**, *28*, 631–635. [CrossRef] [PubMed]
38. Agostinis-Sobrinho, C.; Santos, R.; Rosário, R.; Moreira, C.; Lopes, L.; Mota, J.; Martinkenas, A.; García-Hermoso, A.; Correa-Bautista, J.; Ramírez-Vélez, R.; et al. Optimal Adherence to a Mediterranean Diet May not Overcome the Deleterious Effects of Low Physical Fitness on Cardiovascular Disease Risk in Adolescents: A Cross-Sectional Pooled Analysis. *Nutrients* **2018**, *10*, 815. [CrossRef] [PubMed]
39. Ramírez-Vélez, R.; Correa-Bautista, J.E.; Ojeda-Pardo, M.L.; Sandoval-Cuellar, C.; García-Hermoso, A.; Carrillo, H.A.; González-Ruíz, K.; Prieto-Benavides, D.H.; Tordecilla-Sanders, A.; Martinkėnas, A.; et al. Optimal Adherence to a Mediterranean Diet and High Muscular Fitness Are Associated with a Healthier Cardiometabolic Profile in Collegiate Students. *Nutrients* **2018**, *10*, 511. [CrossRef] [PubMed]
40. Agostinis-Sobrinho, C.; Ramírez-Vélez, R.; García-Hermoso, A.; Rosário, R.; Moreira, C.; Lopes, L.; Martinkenas, A.; Mota, J.; Santos, R. The combined association of adherence to Mediterranean diet, muscular and cardiorespiratory fitness on low-grade inflammation in adolescents: A pooled analysis. *Eur. J. Nutr.* **2018**, *3*, 1–8. [CrossRef] [PubMed]
41. Albert Pérez, E.; Mateu Olivares, V.; Martínez-Espinosa, R.M.; Molina Vila, M.D.; Reig García-Galbis, M. New Insights about How to Make an Intervention in Children and Adolescents with Metabolic Syndrome: Diet, Exercise vs. Changes in Body Composition. A Systematic Review of RCT. *Nutrients* **2018**, *10*, 878. [CrossRef] [PubMed]
42. Salas-Salvadó, J.; Díaz-López, A.; Ruiz-Canela, M.; Basora, J.; Fitó, M.; Corella, D.; Serra-Majem, L.; Wärnberg, J.; Romaguera, D.; Estruch, R.; et al. PREDIMED-Plus investigators. Effect of a lifestyle intervention program with energy-restricted Mediterranean diet and exercise on weight loss and cardiovascular risk factors: One-year results of the PREDIMED-Plus trial. *Diabetes Care.* **1808**, dc180836, (Epub ahead of print). [CrossRef]
43. Ros, E.; Martínez-González, M.A.; Estruch, R.; Salas-Salvadó, J.; Fitó, M.; Martínez, J.A.; Corella, D. Mediterranean Diet and Cardiovascular Health: Teachings of the PREDIMED Study. *Adv. Nutr.* **2014**, *5*, 330S–336S. [CrossRef] [PubMed]
44. Organización Mundial de la Salud. *Estrategia Mundial Para la Salud de la Mujer, el Niño y el Adolescente (2016–2030)*; Organización Mundial de la Salud: Geneva, Switzerland, 2015.
45. Khatib, O. Noncommunicable diseases: Risk factors and regional strategies for prevention and care. *East. Mediterr. Health J.* **2004**, *10*, 778–788. [PubMed]
46. Lobstein, T.; Baur, L.; Uauy, R. Obesity in children and young people: A crisis in public health. *Obes. Rev. Suppl.* **2004**, *5*, 4–104. [CrossRef] [PubMed]
47. Popkin, B.M. Global nutrition dynamics: The world is shifting rapidly toward a diet linked with noncommunicable diseases. *Am. J. Clin. Nutr.* **2006**, *84*, 289–298. [CrossRef] [PubMed]
48. Waters, E.; de Silva-Sanigorski, A.; Burford, B.J.; Brown, T.; Campbell, K.J.; Gao, Y.; Armstrong, R.; Prosser, L.; Summerbell, C.D. Interventions for preventing obesity in children. *Cochrane Database Syst. Rev.* **2011**, *12*, CD001871. [CrossRef] [PubMed]

49. Wolfenden, L.; Jones, J.; Williams, C.M.; Finch, M.; Wyse, R.J.; Kingsland, M.; Tzelepis, F.; Wiggers, J.; Williams, A.J.; Seward, K.; et al. Strategies to improve the implementation of healthy eating, physical activity and obesity prevention policies, practices or programmes within childcare services. *Cochrane Database Syst. Rev.* **2016**, *10*, CD011779. [CrossRef] [PubMed]
50. Serra-Majem, L.; García-Closas, R.; Ribas, L.; Pérez-Rodrigo, C.; Aranceta, J. Food patterns of Spanish schoolchildren and adolescents: The enKid Study. *Public Health Nutr.* **2001**, *4*, 1433–1438. [CrossRef] [PubMed]
51. Serra-Majem, L.; Ribas, L.; Ngo, J.; Ortega, R.M.; García, A.; Pérez-Rodrigo, C.; Aranceta, J. Food, youth and the Mediterranean diet in Spain. Development of KIDMED, Mediterranean Diet Quality Index in children and adolescents. *Public Health Nutr.* **2004**, *7*, 931–935. [CrossRef] [PubMed]
52. García Cabrera, S.; Herrera Fernández, N.; Rodríguez Hernández, C.; Nissensohn, M.; Román-Viñas, B.; Serra-Majem, L. KIDMED test; prevalence of low adherence to the Mediterranean Diet in children and young; a systematic review. *Nutr. Hosp.* **2015**, *32*, 2390–2399. [PubMed]
53. Ruiz, J.R.; Castro-Piñero, J.; España-Romero, V.; Artero, E.G.; Ortega, F.B.; Cuenca, M.A.M.; Enez-Pavón, D.J.; Chillón, P.; Girela-Rejón, M.J.; Mora, J.; et al. Field-based fitness assessment in young people: The ALPHA health-related fitness test battery for children and adolescents. *Br. J. Sports Med.* **2011**, *45*, 518–524. [CrossRef] [PubMed]
54. Léger, L.A.; Mercier, D.; Gadoury, C.; Lambert, J. The multistage 20 metre shuttle run test for aerobic fitness. *J. Sports Sci.* **1988**, *6*, 93–101. [CrossRef] [PubMed]
55. Villa-González, E.; Ruiz, J.R.; Chillón, P. Associations between active commuting to school and health-related physical fitness in Spanish school-aged children: A cross-sectional study. *Int. J. Environ. Res. Public Health* **2015**, *12*, 10362–10373. [CrossRef] [PubMed]
56. España-Romero, V.; Artero, E.G.; Santaliestra-Pasias, A.M.; Gutierrez, A.; Castillo, M.J.; Ruiz, J.R. Hand Span Influences Optimal Grip Span in Boys and Girls Aged 6 to 12 Years. *J. Hand Surg. Am.* **2008**, *33*, 378–384. [CrossRef] [PubMed]
57. Moreno, L.A.; Mesana, M.I.; González-Gross, M.; Gil, C.M.; Fleta, J.; Wärnberg, J.; Ruiz, J.R.; Sarría, A.; Marcos, A.; Bueno, M.; et al. Anthropometric body fat composition reference values in Spanish adolescents. The AVENA Study. *Eur. J. Clin. Nutr.* **2006**, *60*, 191–196. [CrossRef] [PubMed]
58. Moreno, L.A.; Mesana, M.I.; González-Gross, M.; Gil, C.M.; Ortega, F.B.; Fleta, J.; Wärnberg, J.; León, J.F.; Marcos, A.; Bueno, M.; et al. Body fat distribution reference standards in Spanish adolescents: The AVENA Study. *Int. J. Obes.* **2007**, *31*, 1798–1805. [CrossRef] [PubMed]
59. Ortega, F.; Artero, E.; Ruiz, J.; España-Romero, V.; Jiménez-Pavón, D.; Vicente-Rodriguez, G.; Moreno, L.; Manios, Y.; Béghin, L.; Ottevaere, C.; et al. Physical fitness levels among European adolescents: The HELENA study. *Br. J. Sports Med.* **2011**, *45*, 20–29. [CrossRef] [PubMed]
60. Barker, A.R.; Gracia-Marco, L.; Ruiz, J.R.; Castillo, M.J.; Aparicio-Ugarriza, R.; González-Gross, M.; Kafatos, A.; Androutsos, O.; Polito, A.; Molnar, D.; et al. Physical activity, sedentary time, TV viewing, physical fitness and cardiovascular disease risk in adolescents: The HELENA study. *Int. J. Cardiol.* **2018**, *254*, 303–309. [CrossRef] [PubMed]
61. Gualteros, J.A.; Torres, J.A.; Umbarila-Espinosa, L.M.; Rodríguez-Valero, F.J.; Ramírez-Vélez, R. A lower cardiorespiratory fitness is associated to an unhealthy status among children and adolescents from Bogotá, Colombia. *Endocrinol. Nutr.* **2015**, *62*, 437–446. [CrossRef] [PubMed]
62. Joensuu, L.; Syväoja, H.; Kallio, J.; Kulmala, J.; Kujala, U.M.; Tammelin, T.H. Objectively measured physical activity, body composition and physical fitness: Cross-sectional associations in 9- to 15-year-old children. *Eur. J. Sport Sci.* **2018**, *18*, 882–892. [CrossRef] [PubMed]
63. Syväoja, H.J.; Kankaanpää, A.; Kallio, J.; Hakonen, H.; Kulmala, J.; Hillman, C.H.; Pesonen, A.-K.; Tammelin, T.H. The relation of physical activity, sedentary behaviors, and academic achievement is mediated by fitness and bedtime. *J. Phys. Act. Health* **2018**, *15*, 135–143. [CrossRef] [PubMed]
64. Wärnberg, J.; Ruiz, J.R.; Ortega, F.B.; Romeo, J.; González-Gross, M.; Moreno, L.A.; García-Fuentes, M.; Gómez, S.; Nova, E.; Díaz, L.E. Estudio AVENA (Alimentación y valoración del estado nutricional en adolescentes). Resultados obtenidos 2003–2006. *Pediatr. Integr.* **2006**, *1*, 50–55.
65. García-Sánchez, A.; Burgueño-Menjibar, R.; López-Blanco, D.; Ortega, F.B. Condición física, adiposidad y autoconcepto en adolescentes. Estudio piloto. *Rev. Psicol. Deport.* **2013**, *22*, 453–461.

66. Secchi, J.D.; García, G.C.; España-Romero, V.; Castro-Piñero, J. Condición física y riesgo cardiovascular futuro en niños y adolescentes argentinos: Una introducción de la batería ALPHA. *Arch. Argent. Pediatr.* **2014**, *112*, 132–140. [PubMed]
67. Fernández, I.; Canet, O.; Giné-Garriga, M. Assessment of physical activity levels, fitness and perceived barriers to physical activity practice in adolescents: Cross-sectional study. *Eur. J. Pediatr.* **2017**, *176*, 57–65. [CrossRef] [PubMed]
68. Buchan, D.S.; Boddy, L.M.; Young, J.D.; Cooper, S.M.; Noakes, T.D.; Mahoney, C.; Shields, J.P.H.; Baker, J.S. Relationships between Cardiorespiratory and Muscular Fitness with Cardiometabolic Risk in Adolescents. *Res. Sport. Med.* **2015**, *23*, 227–239. [CrossRef] [PubMed]
69. Ministry of Education, Science and Culture. *Icelandic National Curriculum Guide for Compulsory Schools –with Subjects Areas*; Ministry of Education, Science and Culture: Reykjavik, Iceland, 2014; p. 250, ISBN 978-9935-436-24-5.
70. Bacopoulou, F.; Landis, G.; Rentoumis, A.; Tsitsika, A.; Efthymiou, V. Mediterranean diet decreases adolescent waist circumference. *Eur. J. Clin. Invest.* **2017**, *47*, 447–455. [CrossRef] [PubMed]
71. Muros, J.; Salvador, F.; Zurita, F.; Gamez, V.K.E. The association between healthy lifestyle behaviors and health-related quality of life among adolescents. *J. Pediatr.* **2017**, *93*, 406–412. [CrossRef] [PubMed]
72. Martino, F.; Puddu, P.E.; Lamacchia, F.; Colantoni, C.; Zanoni, C.; Barillà, F.; Martino, E.; Angelico, F. Mediterranean diet and physical activity impact on metabolic syndrome among children and adolescents from Southern Italy: Contribution from the Calabrian Sierras Community Study (CSCS). *Int. J. Cardiol.* **2016**, *225*, 284–288. [CrossRef] [PubMed]
73. Martino, F.; Puddu, P.E.; Pannarale, G.; Colantoni, C.; Zanoni, C.; Martino, E.; Barillà, F. Metabolic syndrome among children and adolescents from Southern Italy: Contribution from the Calabrian Sierras Community Study (CSCS). *Int. J. Cardiol.* **2014**, *177*, 455–460. [CrossRef] [PubMed]
74. Muros, J.J.; Cofre-Bolados, C.; Arriscado, D.; Zurita, F.; Knox, E. Mediterranean diet adherence is associated with lifestyle, physical fitness, and mental wellness among 10-y-olds in Chile. *Nutrition* **2017**, *35*, 87–92. [CrossRef] [PubMed]
75. Mistretta, A.; Marventano, S.; Antoci, M.; Cagnetti, A.; Giogianni, G.; Nolfo, F.; Rametta, S.; Pecora, G.; Marranzano, M. Mediterranean diet adherence and body composition among Southern Italian adolescents. *Obes. Res. Clin. Pract.* **2017**, *11*, 215–226. [CrossRef] [PubMed]
76. Ozen, A.E.; Bibiloni, M.D.M.; Murcia, M.A.; Pons, A.; Tur, J.A. Adherence to the Mediterranean diet and consumption of functional foods among the Balearic Islands' adolescent population. *Public Health Nutr.* **2015**, *18*, 659–668. [CrossRef] [PubMed]
77. Grosso, G.; Galvano, F. Mediterranean diet adherence in children and adolescents in southern European countries. *NFS J.* **2016**, *3*, 13–19. [CrossRef]
78. Santomauro, F.; Lorini, C.; Tanini, T.; Indiani, L.; Lastrucci, V.; Comodo, N.; Bonaccorsi, G. Adherence to Mediterranean diet in a sample of Tuscan adolescents. *Nutrition* **2014**, *30*, 1379–1383. [CrossRef] [PubMed]
79. Schröder, H.; Mendez, M.A.; Ribas-Barba, L.; Covas, M.-I.; Serra-Majem, L. Mediterranean diet and waist circumference in a representative national sample of young Spaniards. *Int. J. Pediatr. Obes.* **2010**, *5*, 516–519. [CrossRef] [PubMed]
80. Lloyd, R.S.; Oliver, J.L. The youth physical development model: A new approach to long-term athletic development. *Strength Cond. J.* **2012**, *34*, 61–72. [CrossRef]
81. de Oliveira, B.F.; de Carvalho, L.V.; de Souza Mourão, D.; da Costa, R.D.; de Castro, H.A.; Artaxo, P.; Junger, W.L.; Hacon, S. Environmental Exposure Associated with Oxidative Stress Biomarkers in Children and Adolescents Residents in Brazilian Western Amazon. *J. Environ. Prot.* **2018**, *09*, 347–367. [CrossRef]
82. Damian, M.; Oltean, A.; Damian, C. The impact of sedentary behavior on health and the need for physical activity in children and adolescents. *Rev. Rom. Pentru Educ. Multidimens.* **2018**, *10*, 71. [CrossRef]

International Journal of
Environmental Research and Public Health

MDPI

Article

Utilization of Formal and Informal Care by Community-Living People with Dementia: A Comparative Study between Sweden and Italy

Carlos Chiatti [1,*], Danae Rodríguez Gatta [1], Agneta Malmgren Fänge [1], Valerio Mattia Scandali [2], Filippo Masera [3] and Connie Lethin [1,4] on behalf of the UP-TECH and TECH@HOME research groups

1. Department of Health Sciences, Faculty of medicine, Lund University, Box 157, SE-221 00 Lund, Sweden; danaerodriguezg@gmail.com (D.R.G.); agneta.malmgren_fange@med.lu.se (A.M.F.)
2. Department of Biomedical Sciences, Polytechnic University of Marche, 60030 Ancona, Italy; valeriomattiascandali@gmail.com
3. Department of Health Care Planning, Regional Health Agency of Marche Region, 60015 Ancona, Italy; filippo.masera@regione.marche.it
4. Department of Clinical Sciences, Clinical Memory Research Unit, Faculty of Medicine, Lund University, SE-214 28 Malmö, Sweden; connie.lethin@med.lu.se

* Correspondence: carlos.chiatti@med.lu.se; Tel.: +46-70-213-04-83

Received: 13 October 2018; Accepted: 22 November 2018; Published: 28 November 2018

Abstract: *Background:* Dementia is a public health priority with a dramatic social and economic impact on people with dementia (PwD), their caregivers and societies. The aim of this study was to contribute to the knowledge on how utilization of formal and informal care varies between Sweden and Italy. *Methods:* Data were retrieved from two trials: TECH@HOME (Sweden) and UP-TECH (Italy). The sample consisted of 89 Swedish and 317 Italian dyads (PwD and caregivers). Using bivariate analysis, we compared demographic characteristics and informal resource utilization. Multiple linear regression was performed to analyze factors associated with time spent on care by the informal caregivers. *Results:* Swedish participants utilized more frequently health care and social services. Informal caregivers in Italy spent more time in caregiving than the Swedish ones (6.3 and 3.7 h per day, respectively). Factors associated with an increased time were country of origin, PwD level of dependency, living situation, use of formal care services and occupation. *Conclusions:* Care and service utilization significantly varies between Sweden and Italy. The level of formal care support received by the caregivers has a significant impact on time spent on informal care. Knowledge on the factors triggering formal care resources utilization by PwD and their caregivers might further support care services planning and delivery across different countries.

Keywords: dementia; health services; resource utilization; dementia care; informal care; formal care; cognitive disorder

1. Introduction

Dementia is a disease currently affecting around 50 million people worldwide, corresponding to about 5% of the older population worldwide [1]. It is a public health priority for which most of the modern welfare states are not fully prepared [2]. Dementia onset is associated with age and low education in early life [1,3]. Other causal determinants include hypertension at midlife, smoking and diabetes, which are often related to lifestyle factors [3]. People with dementia face a disabling condition and gradually reach the need for extensive support by informal and formal care in activities of daily living (ADL), e.g., showering, dressing and managing finances [3], restricting their participation in the society leading to adverse events such as institutionalization and inappropriate use of health care and social services (care and services) [3]. The disease has no cure and affects considerably also the life of families and significant others, leading to an overall negative economic and health impact for the society [4]. Indeed, empirical data show that everywhere in Europe families and other informal caregivers are those who provide the bulk of care to their dependent older relatives, thus facing the most significant share of the disease burden [5].

The World Health Organization (WHO) strongly recommends addressing this challenge by creating specific dementia care plans in each country and accordingly reallocating resources [1]. In line with these recommendations, high-income countries in recent years endeavored to achieve effective health prevention, timely diagnosis and an overall increase of the availability of health care and social services (care and services) [6]. Access to long-term care services is particularly relevant due to the dementia characteristic trajectories during which people might require a mix of care and services to ensure independent living and quality of life [7]. The care is mainly related to the timely diagnosis and symptom management and services are mainly focusing on supporting independence in ADL. The availability of care and services may differ significantly across countries, e.g., in terms of type and amount of services provided and the way they are delivered [8]. Recent evidence suggests that this intervention area could include interventions to improve lifestyle, e.g., in terms of physical activity and nutrition, as this might prevent complications and even slow down the progression of the disease [9]. As the majority of people with dementia live at home, due to the deinstitutionalization and ageing-in-place policies, informal caregivers, mainly family members, make up the cornerstone of the whole care system [3,4]. However, as a consequence of their difficult roles, they experience a high level of caregiver burden which put them in a condition of psychological, physical and financial strain [1,3], being at risk of declining health and increased care needs. As the on-going trends suggest a future increase of the prevalence of people with dementia worldwide, there is an urgent need to improve the effectiveness and efficiency of the dementia care systems in all countries. It is especially relevant to achieve an appropriate balance between the formal and informal care contribution to guarantee care situations respectful of the quality of life and social rights of both people with dementia and their caregivers.

In this respect, several studies have analyzed the resource utilization in dementia care [10] across different geographical and cultural contexts, also using a comparative approach [11]. These studies have revealed how widely the use of care and services can vary across countries, and how different the conditions of the informal caregivers can be. Nonetheless, many studies have adopted a pure cost-analysis approach, providing mostly data on the global financial consequences of the disease (e.g., in terms of total public or private healthcare expenditure), without giving details on the characteristics and intensity of the services driving the costs and their relation with the informal care network. Comparative analysis in this area would be particularly interesting within the European Union context, and the similarities and differences in how the dementia care systems are organized [12]. Sweden and Italy for instance represent different welfare systems and cultures (Nordic vs. Mediterranean), facing similar challenges in organizing dementia care and support for informal caregivers. Both Sweden and Italy are high income countries with similar life expectancy at birth, estimated at 83 and 82 years for Italy and Sweden, respectively [13,14]. Sweden is a Nordic welfare state with a long tradition of service provision to support people with dementia and their caregivers [15–17]. The population size is around 10 million inhabitants, representing one of the oldest populations in the world [18,19], out of which approximately 160,000 people have a dementia diagnosis [20]. Dementia is one of the leading causes of mortality in Sweden [14]. The majority of people with dementia live at home with support from informal caregivers and have a wide range of available care and services [20]. The care of people with dementia in Sweden is a shared responsibility between the counties and the municipalities. The county councils are responsible for healthcare according to the Health and Medical Services Act [21]. The municipalities are responsible for the elderly citizens, 65 years or older, according to the Social Services Act [22]. All care is financed through taxes. National guidelines for dementia care and the provision of care and services have been developed to ensure that the care provided by county councils and municipalities are equal independently of where one lives [21]. Italy has a population of approximately 60 million people and the second highest proportion of older people (65 years old or more) in Europe and about 1 million people with dementia [23,24]. The healthcare system is regionally decentralized, and Italy was the first country to introduce nationwide specialized memory clinics for dementia care [16,25]. Here, while home help is provided by the municipalities, nursing home care is the responsibility of the Local Health District and largely funded by the National Health Fund. However, significant disparities exist among Italian regions regarding the availability of these services [25]. Most often, families opt for privately paid care workers due to the scarce services available for dementia care, and thus they face a significant impact on household budgets due to out-of-pocket expenditures [23,26]. This in part urged the creation of a national plan that aims to promote and improve interventions for people with dementia and their families [25].

In terms of resources allocated to healthcare, the total health expenditure in Italy reaches about 9% and in Sweden about 11% of GDP, while out-of-pocket expenditure of the Italian population accounts for 23% of their current health expenditure, compared to 15% for Sweden [27]. Due to the different allocations of resources within the healthcare budget and the different levels of resources available in the area of care and services, it is estimated that, on average, a Swedish citizen has at least three times more funding available for long-term care compared to an Italian citizen [15].

This study was based on the experience of two recent research projects, namely TECH@HOME and UP-TECH, which have investigated, among other outcomes, the economic impact of dementia in Sweden and Italy. The similarities of the study designs and samples give the opportunity to investigate two different care systems responding to similar challenges concerning welfare systems sustainability. Such comparison might contribute to a better understanding of the current dementia care and service systems in Europe, but also to hypothesize how differences in the levels of formal care provision might impact on the dynamics of informal care levels, and vice versa [28]. The overarching aim of this study was thus to contribute to the knowledge gap on utilization of formal and informal dementia care across European welfare states and on how such differences might impact on the daily arrangements of people with dementia and their caregivers. Specific aims of the study were to evaluate: (a) the characteristics of people with dementia and their caregivers in Sweden and Italy; (b) the level of resources utilization in the two countries; and (c) the factors associated with the time spent in caregiving by the informal caregivers of the people with dementia.

2. Materials and Methods

2.1. Study Design and Sample

This study had a comparative cross-sectional design, drawing on data from two studies. TECH@HOME [29] is an ongoing prospective study conducted in southern Sweden between 2016 and 2019, focusing aspects of physical and mental health, health related quality of life and use of health care and social services among PwD and their informal caregivers, after installation of a sensors-based monitoring system in their homes. The UP-TECH (Chiatti et al.) [30] project is a multi-component, randomized controlled trial recently conducted in the Italian Marche region during 2014–2015. The main objective was to reduce caregiver burden of family caregivers of patients with Alzheimer's disease (AD) and to maintain patients with AD at home for as long as possible, through the use of new technologies and of case-management approaches. Inclusion criteria for the two studies were very close, although not completely overlapping. In TECH@HOME, the inclusion criteria were: a diagnosis of major neurocognitive disorders with a mild to moderate severity; a Mini Mental State Examination (MMSE-SR) between 14 and 24 [31]; a Global Deterioration Scale (GDS) score between 1 and 5; living at home; being able to understand and communicate in Swedish; and having an informal caregiver [29]. Dyads were enrolled through the primary health care centers and the municipalities. Inclusion criteria of UP-TECH were [30]: Alzheimer's diagnosis at an intermediate stage; and MMSE between 10 and 20. Potential participants were recruited at the Alzheimer Evaluation Units [30].

To improve the comparability of the two samples, we selected in this study only people with dementia with a MMSE between 14 and 24 and their caregivers. The final sample resulted in 89 Swedish and 317 Italian dyads.

2.2. Data and Measures

Basic socio-demographic data such as gender, age, marital status and living situation of the person with dementia were available for both the TECH@HOME and the UP-TECH samples, as well as data on the age, gender, occupation, education and caring situation of the caregivers. In addition, the questionnaires administered by the research nurses in both studies included reliable and valid instruments to address several clinical (related to diseases and other clinical conditions) and functional dimensions (related to the individual cognitive and physical functioning, e.g., ADL, IADL and MMSE). The rationale for also evaluating these dimensions in our study is that literature has vastly confirmed the existence of multiple explanatory factors behind the use of health and social care resources. The Andersen–Newman healthcare utilization model [32], for instance, suggests that use of resources could depend on the actual needs of a person as well as on the socio-environmental characteristics of the individual and his/her family. The procedures used for data collection are thoroughly described in details elsewhere [29,30]. In both studies, information on cognitive function, as measured by the MMSE [31], was available. The MMSE is a 30-point questionnaire extensively used in clinical and research settings to measure cognitive impairment. A score greater or equal to 24 points indicates intact functioning, while below this threshold can indicate severe (≤9), moderate (10–18) or mild (19–23) cognitive impairment. Data on ADL were collected by means of the Interrai ADL Hierarchy Scale [33]. This scale includes four items rating the functional status in relation to self-performance (i.e., personal hygiene, toilet use, locomotion, and eating), which are summarized in a hierarchical scale that ranges from 0 (no impairment) to 6 (totally dependent). Dependence in IADL were assessed using the IADL Involvement Scale, which is based on seven IADL-related items, summarized in a scale that ranges from 0 to 48, with higher scores indicating greater dependency [34]. The Hospital Anxiety and Depression scale (HADS) was used in both studies to assess caregiver's level of psychological health [35]. HADS includes 14 items; seven items related to anxiety and seven items related to depression. Each item is assessed on a Likert scale that ranges from 0 to 3 and the overall scores range from 0 to 21 in both anxiety and depression where 0–7 is consistent with absence of the conditions [35].

2.3. Use of Resources

The instrument used for data collection in TECH@HOME was the Resource Utilization in Dementia (RUD) [36]. This instrument has been extensively used especially for cost analysis and has a widespread use in global settings. The RUD instrument collects information regarding the level of resource utilization by frequency in hours and days, from both the person with dementia and the caregiver, together with other demographic and health status information. The UP-TECH study used an ad-hoc developed Resource Utilization Form, which had several sections overlapping with those included in the RUD, therefore providing comparable data for the analysis [30].

Some variables have been be recoded to obtain identical unit of measurements: for example, outpatient visits in Italy were measured using a six-month timeframe, while in the Swedish questionnaire they were measured using a 30-day period. Time spent in care activities by informal caregivers in TECH@HOME has been assessed using the specific section of the RUD instrument, while in UP-TECH one item assessed the hours spent in caregiving activities by the primary caregiver during the day. A second item retrieved the time spent by other secondary caregivers.

2.4. Data Analysis

In a first step, samples were compared according to socio-demographic characteristics, clinical and functional measures. Differences were investigated using bivariate analysis. Statistical significance for categorical variables was assessed using the Chi-squared or the Fisher's exact test. For continuous variables, independent T-test and Mann–Whitney test were used to compare means depending on the distribution of the investigated variable (normal vs. non-normal distribution). In a second step, differences in the level of use of formal care services and in the amount of informal care provided were evaluated following a similar statistical procedure. Finally, multiple linear regression models were built to evaluate factors associated with informal caregiving time. Potential factors were tested in the model if they had a p-value ≤ 0.25 following the results of the bivariate analysis and inserted in the model using a step-forward. Different models were tried, and variables were dropped from the model when they decreased the Adjusted R-squared, had no significant results or had no impact in the rest of the parameters. The coefficients described in the tables are based on 1000 bootstrap samples. Statistical significance was considered with a p-value ≤ 0.05. The statistical software package used was STATA (Statacorp, College Station, TX, USA) [37].

3. Results

3.1. Demographic Characteristics of Persons with Dementia and Informal Caregivers

Mean age among study participants with dementia was higher in Italy compared to Sweden (81.5 vs. 78.4, $p < 0.001$). Nonetheless, the largest age group in both countries was that of people between 80 and 89 years old (Table 1). Most people in both countries were female and living with their spouses. However, people with dementia in Sweden tend to live alone more frequently (47.2% in Sweden vs. 15.1% in Italy, $p < 0.001$). In Italy, it was frequent that a person with dementia lives with an adult child, which was really uncommon in the Swedish context (only in 1.1% of the cases). From a clinical and functional point of view, the Italian participants seem characterized by an overall worse condition. Cognitive function was slightly lower in Italy where the mean MMSE value is 17.8 (SD = 2.3) (vs. 19 ± 3.7 in Sweden). Additionally, the Italian group was characterized by a higher number of comorbidities as well as by higher level of dependency measured using the ADL and IADL scale.

The majority of informal caregivers were women, although in Italy women were more represented than in the Swedish context (68.8% vs. 52.3%). Caregivers in Italy were younger, and more often an adult child living with the person with dementia. In Sweden, the informal caregivers more often were spouses, which tend to be more frequently engaged in a working occupation.

Table 1. Baseline characteristics of persons with dementia and informal caregivers.

	Swedish Sample (n = 89)		Italian Sample (n = 317)		
	%	Mean (SD)	%	Mean (SD)	p *
Persons with dementia					
Age in years, (%)	-	78.4 (7.8)	-	81.5 (5.8)	0.001
Early old age (<69)	11.2	-	3.5	-	0.009
Middle old age (70–79)	39.3	-	32.5	-	
Later old age (80–89)	44.9	-	59.3	-	
Very old age (>90)	4.5	-	4.7	-	
Gender					
Men	27.0	-	30.3	-	0.544
Women	73.0	-	69.7	-	
Living situation					
Living alone	47.2	-	15.1	-	<0.001
Husband/wife	49.4	-	50.2	-	
Child	1.1	-	21.5	-	
Other	2.2	-	13.2	-	
Comorbidity, n of diseases	-	1.7 (1.1)	-	2.2 (1.4)	0.001

Table 1. *Cont.*

	Swedish Sample (n = 89)		Italian Sample (n = 317)		
	%	Mean (SD)	%	Mean (SD)	p *
Cognitive function, MMSE score (0–30) [a]	-	19 (3.7)	-	17.8 (2.3)	<0.001
Mild (≥20)	53.0	-	54 (17.0)	-	<0.001
Moderate (10–19)	47.0	-	263 (83.0)	-	
ADL scale (0–6) [b]	-	0.2 (0.5)	-	1.4 (1.5)	<0.001
IADL scale (0–48) [c]	-	25.2 (12.4)	-	33.3 (14.2)	<0.001
Informal caregivers					
Age-groups (%) [d]					
<39–54	29.6	-	39.1	-	0.051
55–69	39.8	-	31.2	-	
>70	30.7	-	29.7	-	
Gender [d]					
Men	47.7	-	31.2	-	0.004
Women	52.3	-	68.8	-	
Marital status [d]					
Married	74.2	-	79.8	-	0.031
Divorced/widowed	2.2	-	7.9	-	
Never married	22.5	-	12.3	-	
Occupation [e]					
Employee/self–employed	50.0	-	46.1	-	0.001
Job seeking	4.7	-	3.8	-	
Retired	36.0	-	36.0	-	
Sickness/activity allowance/sick leave [f]	5.9	-	2.2	-	
Other	3.5	-	12.0	-	
Level of education [d]					
Elementary school	14.8	-	7.6	-	0.074
Gymnasium	36.4	-	33.8	-	
University	48.9	-	58.7	-	
Relationship to the person with dementia [d]					
Wife/Husband	43.2	-	30.9	-	0.038
Child	50.0	-	54.6	-	
Other	6.8	-	14.5	-	
Living with the person with dementia					
Yes	45.5	-	63.7	-	0.002

Data source for Sweden: TECH@HOME questionnaire. Data source for Italy: UP-TECH questionnaire; ADL, activities of daily living; IADL, instrumental activities of daily living; n, number of observations; MMSE, mini mental state examination; p value of significance; SD, standard deviation. [a] missing values, n = 6; [b] missing values, n = 2; [c] missing values, n = 4; [d] missing values, n = 1; [e] missing values, n = 3; [f] long term sick leave. * $p < 0.05$ was regarded as significant; significant p-values are underlined. Underlined values indicate positive results, e.g., 0–30.

3.2. Resource Utilization by Persons with Dementia and Their Caregivers

Results showed that the Swedish participants more frequently used healthcare and social service than the Italian sample, except for emergency care (Table 2). However, hospital admissions were more frequent in Sweden compared to Italy (5.6 vs. 1.6, $p < 0.001$). The most frequently used services in Sweden were day care centers (used by 39.8% of the sample), social services (65.9%) and home care (38.6%). In Italy, the main support for the people with dementia and their caregiver was the use of privately paid home help services (used by 28.7% of the sample).

Table 2. Resource utilization the last 30 days of healthcare and social services by persons with dementia.

	Swedish Sample (n = 89)		Italian Sample (n = 317)		
	%	Mean (SD)	%	Mean (SD)	p *
Hospital admission (%)					
Yes	5.6	-	1.6	-	<u>0.045</u>
Number of admissions	-	1.6 (1.3)	-	-	
Emergency ward admission [b]					
Yes	4.6	-	6.3	-	0.798
No	95.4	-	93.7	-	
Number of visits	-	1.3 (0.5)	-	1 (0)	0.477
Outpatient care					
Yes	38.2	-	9.1	-	<0.001
Number of visits	-	1.7 (1.3)	-	1 (0)	<u>0.001</u>
Municipal (SWE) or District (ITA) nurse [c]					
Yes	12.6	-	2.8	-	<u>0.001</u>
Number of visits	-	3.36 (2.7)	-	3 (3.5)	0.664
Day care centre					
Yes	39.8	-	3.8	-	<u><0.001</u>
Number of visits	-	11.4 (5.4)	-	1 (0)	<u><0.001</u>
Social care services [c]					
Yes	65.9	-	6.9	-	<u><0.001</u>
Home help (SWE) or Private Carer (ITA) [b]					
Yes	38.6	-	28.7	-	0.074

Data source for Sweden: TECH@HOME questionnaire. Data source for Italy: UP-TECH questionnaire. ITA, Italy; p, p-value; SD, standard deviation; SWE, Sweden. [a] Median and interquartile range is only provided in case the variable was non-normally distributed: [b] missing values, n = 1; [c] missing values, n = 2. * $p < 0.05$ was regarded as significant; significant p-values are underlined.

3.3. Time Spent in Informal Caregiving

Italian informal caregivers spent significantly more time in all caring activities than Swedish ones and had a mean of 6.27 h per day (vs. 3.5 h per day spent by Swedish caregivers) (Table 3). In Italy, the number of hours spent in caregiving activities is higher when the person with dementia is a male (7.47 vs. 5.76, $p = 0.006$) and the caregiver is a female (6.49 vs. 5.81, p value is not significant). In both countries, the hours of caregiving provided increase with the level of dependency in ADL and IADL. Living with the person with dementia is not associated with a higher amount of care provided either in Sweden or in Italy.

Table 3. Time spent in informal caregiving and psychological health of informal caregivers.

	Swedish Sample (n = 89)			Italian sample (n = 317)		
	N (%)	Mean (SD)	p *	N (%)	Mean (SD)	p *
Patients characteristics						
Age in years			0.205			0.230
<69	10 (11.4)	3.22 (3.63)		11 (3.5)	3.96 (3.49)	
70–79	35 (39.8)	3.92 (3.27)		103 (32.5)	6.26 (6.98)	
80–89	39 (44.3)	3.74 (5.03)		188 (59.3)	6.23 (6.13)	
>90	4 (4.5)	1.80 (0.91)		15 (4.7)	8.64 (7.72)	
Gender			0.351			<u>0.006</u>
Male	24 (27.3)	3.60 (2.65)		96 (30.3)	7.47 (6.72)	
Female	64 (62.7)	3.69 (4.55)		221 (69.7)	5.76 (6.26)	

Table 3. *Cont.*

	Swedish Sample (n = 89)			Italian sample (n = 317)		
	N (%)	Mean (SD)	p *	N (%)	Mean (SD)	p *
MMSE, score (0–30)			0.006			0.244
Mild (≥20)	43 (52.4)	2.98 (3.63)		54 (17.0)	6.11 (6.80)	
Moderate (10–19)	39 (47.6)	4.44 (4.55)		263 (83.0)	6.31 (6.38)	
ADL scale score (0–6)			0.304			0.002
0	74 (84.1)	3.24 (3.30)		125 (39.4)	4.73 (5.26)	
1	8 (9.1)	6.69 (7.49)		72 (22.7)	7.26 (7.13)	
2	4 (4.5)	5.59 (7.65)		61 (19.2)	7.26 (7.17)	
3 or more	2 (2.3)	3.50 (0.71)		59 (18.7)	7.32 (6.55)	
IADL scale score (0–48)			<0.001			<0.001
1–22	32 (38.1)	2.66 (3.70)		74 (23.3)	3.30 (3.48)	
23–35	34 (40.5)	3.31 (2.42)		69 (21.8)	6.79 (6.71)	
36–43	13 (15.5)	5.91 (4.35)		81 (25.6)	6.62 (6.68)	
44–48	5 (5.9)	9.35 (8.93)		93 (29.3)	7.97 (7.10)	
Formal Caregiver						
No	53 (60.9)	3.62 (3.82)	0.910	226 (71.3)	6.79 (6.93)	0.062
Yes	34 (39.1)	3.68 (4.62)		91 (28.7)	4.99 (4.81)	
Caregivers characteristics						
Gender			0.515			0.697
Male	42 (47.7)	3.76 (4.41)		99 (31.2)	5.81 (5.56)	
Female	46 (52.3)	3.58 (3.86)		218 (68.8)	6.49 (6.80)	
Status in employment			0.036			<0.001
Employed	43 (50.0)	3.28 (4.38)		146 (46.1)	3.35 (3.09)	
Not employed	43 (50.0)	3.74 (3.01)		171 (53.9)	8.77 (7.44)	
Living conditions			0.015			0.170
Living with the PwD	40 (45.5)	3.98 (3.18)		200 (63.7)	6.33 (6.22)	
Not living with the PwD	48 (54.5)	3.40 (4.77)		114 (36.3)	6.23 (6.91)	
HADS Score Depression (0–21)			0.017			
Normal (0–7)	75 (86.3)	2.95 (3.20)		169 (53.3)	5.95 (6.30)	0.218
Mild (8–10)	9 (10.3)	7.54 (5.61)		79 (24.9)	7.74 (7.65)	
Moderate (11–14)	3 (3.4)	8.78 (9.79)		47 (14.9)	5.55 (5.28)	
Severe (15–21)	0 (0.0)			22 (6.9)	5.10 (4.03)	
HADS Score Anxiety (0–21)			0.354			0.732
Normal (0–7)	83 (95.4)	3.58 (4.14)		202 (63.7)	6.55 (6.60)	
Mild (8–10)	4 (4.6)	4.65 (3.79)		57 (18.0)	5.68 (6.43)	
Moderate (11–14)	0 (0.0)			42 (13.3)	5.30 (5.00)	
Severe (15–21)	0 (0.0)			16 (5.0)	7.56 (7.79)	
Total sample	86 (100)	3.50 (3.74)	n.a.	317	6.27 (6.44)	n.a.

Data source for Sweden: TECH@HOME questionnaire. Data source for Italy: UP-TECH questionnaire. HADS, hospital anxiety and depression scale; p = p-value; n.a, not applicable; p, p-value; PwD, person with dementia; MMSE, mini mental state examination; SD, standard deviation. * $p < 0.05$ was regarded as significant; significant p-values are underlined. Underlining of values indicates positive results, e.g., 0–30.

The results of the multiple regression model showed that higher IADL dependence, working status of the caregiver, use of formal home care, and the country of origin were factors associated with the number of hours of informal care provided (Table 4). The lower the IADL dependence, the higher the amount of care provided (+4.6 h per day, among people with dementia with an IADL score between 44 and 48 ($p < 0.001$). Conversely, working carers provided 4.2 h less care compared to the non-working ones ($p < 0.001$). Compared to Italian couples not using any type of formal care support, those Italians using private care help provided - 1.8 h care/day. Being Swedish, even if not using home help from the municipality, was associated with a reduction of the number of daily hours of care equal to 2.6 ($p < 0.001$).

Table 4. Factors associated with hours of informal caregiving.

Independent Variables	Hours of Informal Caregiving (n = 392)		
	B (SE)	***p*-Value ***	**95% CI**
Constant	6.715 (2.03)	0.001	(2.74;10.686)
Relation with PwD (Partner/Spouse vs. Other)	0.461 (0.600)	0.443	(−0.716; 1.638)
Informal caregiver working (Yes vs. No)	−4.152 (0.553)	<0.001	(−5.237; −3.068)
Interaction between Country # Utilization of formal caregiving (Ref. Italy/not using formal caregiving)			
Italy/using formal caregiving	−1.874 (0.684)	0.006	(−3.215; −0.533)
Sweden/not using formal caregiving	−2.631 (0.707)	<0.001	(−4.017; −1.245)
Sweden/using formal caregiving	−1.078 (0.961)	0.262	(−2.963; 0.805)
IADL scale (1–48)			
23–36	2.794 (0.692)	<0.001	(1.436; 4.152)
37–43	3.273 (0.789)	<0.001	(1.725; 4.821)
44–48	4.655 (0.914)	<0.001	(2.863; 6.447)
Living with the PwD (Yes vs. No)	−0.001 (0.324)	0.995	(−0.638; 0.634)
ADL scale (one point increase)	−0.002 (0.178)	0.990	(−0.351; 0.346)
MMSE score (one point increase)	−0.061 (0.109)	0.572	(−0.276; 0.152)
R-squared	0.244		
Adjusted R-squared	0.222		

Data source: TECH@HOME and UP-TECH questionnaire. ADL, activities of daily living; B (SE), Beta coefficient (standard error); CI, confidence intervals; MMSE, mini mental state examination; PwD, person with dementia. IADL, instrumental activities of daily living. Country of origin: 1 = Sweden, 0 = Italy. Living with the person with dementia: 1 = no, 0 = yes. Utilization of formal caregiver: 1 = yes, 0 = no. Informal caregiver actively working: 1 = yes, 0 = no. * $p < 0.05$ was regarded as significant; significant p-values are underlined.

4. Discussion

4.1. Differences in Socio-Economic Conditions of the Dyads in the Two Countries

Findings of our study showed mainly "women caring for women", as in both countries they were the majority among caregivers and people with dementia. This trend has been seen in many other studies where female gender is highly represented [38–40]. However, the Swedish commitment with gender equality was revealed in our findings. In line with other studies focusing on long-term care in a Swedish context, only slight differences were seen between the proportion of male and female informal caregivers [17,18], differently from what was observed in Italy, where the ratio of female to male informal caregiver is about 7 to 3.

Diverse caring dynamics were seen in each country. In Italy, informal caregivers tend to live significantly more with the person with dementia than in the Swedish sample. Some European studies have shown that living with the PwD indicates higher time spent in caregiving [41–43] and higher risk of caregiver burden [41,44]. These precedents coincide with our findings. Italian caregivers had a higher time spent in caregiving activities and had a worse psychological health than the Swedish sample. Interestingly, when adding IADL dependence level to the model, the relation between the variable "living with the PwD" and "hours of informal care provided" changes. At a bivariate level, living together the PwD was associated to a higher number of hours of care provided. However, our regression suggests that this might be the result of the interaction between the living status and the IADL dependence level of the PwD. In other terms, caregivers caring for PwD with higher IADL impairment are more frequently living with them, compared with those caring for PwD with less level of dependency. This suggests that the choice of co-residence might be influenced by the level of dependency of the person with dementia.

In this respect, it is interesting to observe that Italian persons with dementia were more frequently living with their adult children. This difference may be explained due to a still high prevalence of intergenerational households in Italy and prevalent family-based care of the country [16,38,45]. This possibly reflects a history of a strong cultural norm of caring responsibilities, in spite of its current changing trend inside families [45]. It could also reflect a cost saving strategy. Households under economic strain and the financial situation of a country contribute in the decision of co-residence [46]. This could be intentionally opted for or more a response to the increasing uncertainties in the globalized world [46]. Conversely, in Sweden, people with dementia were more frequently living alone, probably due to the widespread support from the state, which could offset economic burden and allow them to sustain living on their own [46,47].

Surprisingly, even though Swedish caregivers represented an older age group, our findings showed that they were more actively working than the Italian caregivers. The extensive utilization of dementia care services might allow them to remain at the workplace and consequently spend fewer hours per day in caregiving. Conversely, the Italian sample was less actively working despite having more caregivers in a working age. As previously depicted in other studies, they might have to reduce their working hours or to stop working to care for the person with dementia [23,44]. In line with these results, our regression model indicated an association between caregivers actively working and less time spent in informal caregiving activities.

4.2. Differences in Informal Caregiving Intensity

The time spent in informal care was regarded as an important part of the analysis. In a pan-European study, Bremer and Cabrera [48] considered that an average of 4–8 h per day spent in caring for ADLs by informal caregivers constitutes a medium level of intensity of care. Similarly, Ory and Hoffman [49] utilized an intensity of care index where about 3 h per day is considered an intermediate level of care. According to these definitions, we could conclude that both countries significantly relied on informal care. The average time spent in caring activities was 3.9 h per day in the Swedish sample and 6.3 h per day in the Italian one. The age of the PwD was associated with informal caregiving time, and this is consistent with previous studies conducted in a European context [41,50]. Additionally, studies have reported an association between high severity of disease and worse health state of the informal caregiver [42]. Italian people with dementia had a higher level of severity of the disease and this could explain the higher proportion of psychological distress in their informal caregivers.

Interestingly, our regression showed that being from Sweden is associated with fewer mean hours spent in caregiving activities in comparison to the Italian sample. The impact of the country of origin on informal care might reflect differences in health care systems, and socioeconomic and cultural factors [40,51]. In our study, we found higher utilization of health and social care services in the Swedish sample, but lower levels of informal care provisions. These differences might be rooted in several explanatory factors. Firstly, Italy and Sweden represent two different types of welfare states, as suggested by Esping-Andersen work, in which different levels of complementarity between informal and formal care can be found [12,15]. Welfare states can be categorized as liberal, conservative and social democratic and differences among these typologies reflect their "(...) political ideologies with regard to stratification, de-commodification and the public-private mix of welfare" [12]. A social democratic model, such as the Swedish one, represents a state that ensures equal health and social service provision and funding for all citizens through a tax-based system [38]. In a conservative model, such as the Italian one, the state is partially responsible for service provision and funding through social insurance schemes, therefore relying on a strong family-based care. Raggi and Leonardi's work [52] is in line with this interpretation, as it identifies a north–south European gradient in care, in which Northern European countries are characterized by universal social policies, state support for families and a large public sector and Southern European one by a mix of universal private services and benefits together with a fragmented system between health and social care services. Our findings indeed show that the

Italian sample had a comparatively low utilization of services but strongly relied either in another informal caregiver or in paid care workers. Secondly, cultural norms could be considered another cause of divergence between countries. As Brandt and Haberkern pointed out [38], intergenerational help is subject to cultural norms, and according to these norms the state would have an effect in "crowding in" (more family support) or "crowding out" (less family support). For instance, in the Swedish context, intergenerational households are comparatively scarcer and adult children are less frequently caregivers than in the Italian context. The state here supports the families through services, thus creating a "crowding out" effect. Conversely, in Italy, more adult children are informal caregivers and the state plays a residual role, thus creating a "crowding in" effect [38,53]. Due to lack of services, Italian families might have slim alternatives of care and internalized caring responsibilities and this could also reflect their increased level of burden seen in our findings. Thirdly, help seeking behaviors constitute another potential source of explanation of the differences seen in resource utilization between the two countries in our study. According to Alzheimer's Disease International (ADI) "concerns regarding stigma may be one factor deterring or delaying help seeking" [6]. This could be explained by the lack of information and awareness of services available together with a need of improvement in tailored formal care delivery [54]. In the case of the Italian informal caregivers, a perception of self-sufficiency and a reliance on internal family support might hinder the demand for other alternatives of care [48]. Lastly, availability and accessibility of services influences the level of resource utilization. The dementia care pathway should be addressed through a continuum between health and social care services across "(...) the various stages of the disease as a seamless process, as needs for both types of care evolve" [8]. Fragmentation between these could potentially hinder utilization of resources. In this respect, formal services in Italy are often described as scarce and fragmented, hence this could explain the low resource utilization seen in our findings [15,55]. In the Swedish context, formal services seem to be highly available but its utilization is low, therefore an alignment with needs is necessary [56].

4.3. Strengths and Limitations

Among the strengths of our study is the possibility to compare countries due to a highly comparable sample of participants and research designs. Eligibility criteria and questionnaire structures were similar across the two clinical trials. However, cross-country comparison becomes challenging in presence of strong differences in how the services themselves are organized across welfare states. For instance, each country offers different types of social care services, in terms of contents and intensity, therefore their comparison should be taken with caution. Similarly, the variable that we have created for formal home help pointed at two different services in the two countries. Swedish formal caregivers have basic healthcare training while in Italy they are in most cases people without a formal qualification with a migration background [57]. Furthermore, the data available for analysis represent samples from specific regions of each country and significant variation could be found within and across them. Thus, the generalizability of the findings at country-level could be limited. In addition, given the cross-sectional design of our primary studies, other important factors that could be involved in determining the dynamics of care might have overlooked, i.e., generating a so-called ecological bias. Finally, informal care could entail support to several different activities, e.g., ADL, IADL or supervision [41]. Detailed comparisons of informal care time could not be performed in our study, since more details information where not included in the Italian data.

5. Conclusions

Our paper suggests that differences in the organization of health and social care systems across welfare states have broad impacts. They influence the time spent in informal care, the amount of reliance on family support, gender roles, working status and levels of psychological burden among populations. Our findings may facilitate the understanding of dementia care system variation across countries, thus having an impact for both policy and care management. This knowledge is also valuable for professionals and students in educational training, preparing for careers in the health care and social service area. Differences between Sweden and Italy are quite significant and rooted in diverse explanations. Balance between informal and formal care is essential for welfare states sustainability and similar resource utilization dynamics can provide crucial information regarding the informal–formal care mix [54]. In the future, further investigation could be made including analysis of the unmet needs. In this respect, further qualitative research could be necessary to explore the experiences of people during their debate on seeking support services or not in the dementia care pathway.

Author Contributions: D.R.G. and C.C. designed the study. C.L. and F.M. coordinated the data collection. C.C. and V.M.S. performed the statistical analyses in dialogue with all authors. D.R. and C.C. took responsibility for writing the paper and all authors interpreted data and critically revised the article. All authors approved the final version of the manuscript.

Funding: This study is funded by FORTE—Swedish Research Council for Health, Working Life and Welfare (Grant No. 2014-4913), the Faculty of Medicine at Lund University, the Department of Biomedical Sciences, Polytechnic University of Marche, Ancona, Italy; and the Regional Health Agency of Marche Region, Ancona, Italy.

Acknowledgments: Sincere thanks are extended to all the participants in the study and the nurses of the two studies collecting data. This study was conducted within the context of the Centre for Ageing and Supportive Environments (CASE) at Lund University, funded by the Swedish Council for Working Life and Social Research (FORTE). Members of the TECH@HOME Research Group are: Kajsa Hiller, Annica Karlsson, Therese Hector, Josefina Nilsson, Jenny Bengtsson, Przemizlaw Sozanski, Moa Wibom, Sonia Larsson, Knut Mårtensson, Carina Anderberg, Ulla Berggren, Christa Ohlson, Marie Persson, Per Mårten Johansson, Anna Liwander, Caroline Dahlgren Bergström, Maria H. Nilsson. Members of the UP-TECH Research Group are: the UP-TECH multidisciplinary research group: Diletta Baldassarri, Luigina Bitti, Aurora Carosi, Maila Sabbatini, Elisabetta Paolasini, Tania Fiori, Cinzia Fronzi, Simona Giacchetta, Valeria Giacomini, Laura Giovagnoli, Giuli Lattanzi, Anna Maria Manca, Ida Marinelli, Valeria Pigini, Marina Pignotti, Maria Chiara Proietti, Antonia Quarticelli, Martina Rovedi, Letizia Tasso, Valentina Valeri, Andrea Antonioli, Mariella Barabucci, Mara Bassani, Paola Bollettini, M Cristina Bruttapasta, Carla Buccolini, Rosa Carangella, Franca Carboni, Daniela Ceccolini, Patrizia D'Incecco, Mariangela Di Felice, Lara Dini, Giordano Gioia, Giuseppe Di Prima, Giancarlo Giusepponi, Cesarina Lanciotti, Alessia Loffreda, Oriana Luciani, Laura Mariani, Francesco Mastrorilli, Roberto Moroni, Izabela Piatkowska, Roberta Rucoli, Giampaola Scoccia, Nives Teodori, Tiziana Tonelli, Rossano Angeloni, Rosa Anna Bratti, Giuseppe Bonafede, Severino Lorenzetti, Cristina Paci, Giovanna Picciotti, Donella Pezzola, Vincenzo Rea, Vittorio Scialè, Mario Signorino, Francesca Sorvillo, Pia Francesca Tomassini, Alessia Ciccola, Lorena Cionfrini. Daniela Alessandrini, Antonio De Santis, Brunetta Formica, Giuliano Tacchi, Silvia Tortorelli, Gemma Marchegiani, Giovanni Pozzari, Giovanni Santarelli, Dario Bartolucci, Giorgio Caraffa, Filippo Cavallo, Lucia Di Furia, Antonio Lacetera, Claudio M Maffei, Lamberto Manzoli, Demetrio Postacchini, Fabrizia Lattanzio, Andrea Corsonello, Laura Cassetta, Franco Bonfranceschi, Joseph M. Rimland, Antonio Cherubini, Roberta Bevilacqua, Anna Rita Bonfigli, Silvia Bustacchini, Marina Capasso, Patrizia Civerchia, Federica Demma, Mirko Di Rosa, Cinzia Giuli, Marzio Marcellini, Paolo Marinelli, Gabriella Melchiorre, Maria Elena Moraca, Andrea Principi, Cristina Rocchetti, Osvaldo Scarpino, Liana Spazzafumo, Daniela Vincitorio, Krystian Bartulewicz, Paolo Olivetti, and Lorena Rossi.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. World Health Organization. *Global Action Plan on the Public Health Response to Dementia 2017–2025*; World Health Organization: Geneva, Switzerland, 2017.
2. World Health Organization. *Dementia: A Public Health Priority*; World Health Organization: Geneva, Switzerland, 2012.
3. World Health Organization. *The Epidemiology and Impact of Dementia. Current State and Future Trends*; World Health Organization: Geneva, Switzerland, 2015.
4. Alzheimer's Disease International. *World Alzheimer Report 2015. The Global Impact of Dementia. An Analysis of Prevalence, Incidence, Cost and Trends*; Alzheimer's Disease International: London, UK, 2015.
5. Chiatti, C.; Melchiorre, M.G.; Rosa, M.D.; Principi, A.; Santini, S.; Döhner, H.; Lamura, G. Family networks and supports in older age. In *Aging in European Societies: Healthy Aging in Europe*; Springer: Boston, MA, USA, 2013; pp. 133–150.
6. Alzheimer's Disease International. *World Alzheimer Report 2016. Improving Healthcare for People Living with Dementia. Coverage, Quality and Costs Now and in the Future*; Alzheimer's Disease International: London, UK, 2016.
7. Alzheimer's Disease International. *World Alzheimer Report 2013. Journey of Caring. An Analysis of Long-Term Care for Dementia*; Alzheimer's Disease International: London, UK, 2013.

8. Moïse, P.; Schwarzinger, M.; Um, M.-Y. *Dementia Care in 9 OECD Countries: A Comparative Analysis*; OECD Health Working Papers; OECD: Paris, France, 2004; p. 13.
9. Ngandu, T.; Lehtisalo, J.; Solomon, A.; Levälahti, E.; Ahtiluoto, S.; Antikainen, R.; Bäckman, L.; Hänninen, T.; Jula, A.; Laatikainen, T.; et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular riskmonitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): A randomised controlled trial. *Lancet* **2015**, *385*, 2255–2263. [CrossRef]
10. Bökberg, C.; Ahlström, G.; Leino-Kilpi, H.; Soto-Martin, M.E.; Cabrera, E.; Verbeek, H.; Saks, K.; Stephan, A.; Sutcliffe, C.; Karlsson, S. Care and service at home for persons with dementia in Europe. *J. Nurs. Scholarsh.* **2015**, *47*, 407–416. [CrossRef] [PubMed]
11. Wimo, A.; Guerchet, M.; Ali, G.C.; Wu, Y.T.; Prina, A.M.; Winblad, B.; Jönsson, L.; Liu, Z.; Prince, M. The worldwide costs of dementia 2015 and comparisons with 2010. *Alzheimers Dement.* **2017**, *13*, 1–7. [CrossRef] [PubMed]
12. Emmenegger, P.; Kvist, J.; Marx, P.; Petersen, K. Three worlds of welfare capitalism: The making of a classic. *J. Eur. Soc. Policy* **2015**, *25*, 3–13. [CrossRef]
13. World Health Organization. *Italy: WHO Statistical Profile*; World Health Organization: Geneva, Switzerland, 2015.
14. World Health Organization. *Sweden: WHO Statistical Profile*; World Health Organization: Geneva, Switzerland, 2015.
15. Carrera, F.; Pavolini, E.; Ranci, C.; Sabbatini, A. Long-term care systems in comparative perspective: Care needs, informal and formal coverage, and social impacts in European countries. In *Reforms in Long-Term Care Policies in Europe*; Ranci, C., Pavolini, E., Eds.; Springer-Verlag: New York, NY, USA, 2013; pp. 23–54.
16. Ferré, F.; de Belvis, A.G.; Valerio, L.; Longhi, S.; Lazzari, A.; Fattore, G.; Ricciardi, W.; Maresso, A. Italy: Health system review. *Health Syst. Trans.* **2014**, *16*, 1–168.
17. Meagher, G.; Szebehely, M. Long-term care in Sweden: Trends, actors, and consequences. In *Reforms in Long-Term Care Policies in Europe*; Ranci, C., Pavolini, E., Eds.; Springer-Verlag: New York, NY, USA, 2013; pp. 55–78.
18. Anell, A.; Glenngård, A.H.; Merkur, S. Sweden: Health system review. *Health Syst. Trans.* **2012**, *14*, 1–159.
19. World Bank. Data. Sweden. World Bank Group. 2018. Available online: https://data.worldbank.org/country/sweden?view=chart (accessed on 1 July 2018).
20. Swedish Council on Technology Assessment in Health Care. *Dementia—Etiology and Epidemiology: A Systematic Review*; The Swedish Council on Technology Assessment in Health Care: Stockholm, Sweden, 2008; Volume 1.
21. SFS 2017:30. Hälso- och sjukvårdslagen [Health and Medical Services Act]. Available online: https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/halso--och-sjukvardslag_sfs-2017-30 (accessed on 1 August 2018).
22. SFS 2001:453. Socialtjänstlagen [The Social Services Act]. Available online: https://ec.europa.eu/anti-trafficking/sites/antitrafficking/files/social_services_act_sweden_en_1.pdf (accessed on 1 August 2018).
23. De Cola, M.C.; Lo Buono, V.; Mento, A.; Foti, M.; Marino, S.; Bramanti, P.; Manuli, A.; Calabrò, R.S. Unmet needs for family caregivers of elderly people with dementia living in Italy: What do we know so far and what should we do next? *Inquiry* **2017**, *54*, 46958017713708. [CrossRef] [PubMed]
24. World Bank. Data. Italy. World Bank Group. 2018. Available online: https://data.worldbank.org/country/italy?view=chart (accessed on 1 July 2018).
25. Di Fiandra, T.; Canevelli, M.; Di Pucchio, A.; Vanacore, N. Italian Dementia National Plan Working Group. The Italian dementia national plan. Commentary. *Ann. Ist. Super. Sanita* **2015**, *51*, 261–264. [PubMed]
26. Scalmana, S.; Di Napoli, A.; Franco, F.; Vanacore, N.; Di Lallo, D.; Giarrizzo, M.L.; Guasticchi, G. Working Group for Dementia of Lazio Region. Use of health and social care services in a cohort of Italian dementia patients. *Funct. Neurol.* **2013**, *28*, 265–273. [PubMed]
27. World Health Organization. *Global Health Observatory Data Repository*; World Health Organization: Geneva, Switzerland, 2016. (updated 11 December 2017). Available online: http://apps.who.int/gho/data/node.main.GHEDCHEGDPSHA2011?lang=en (accessed on 21 July 2018).

28. Bieber, A.; Stephan, A.; Verbeek, H.; Verhey, F.; Kerpershoek, L.; Wolfs, C.; de Vugt, M.; Woods, R.T.; Røsvik, J.; Selbaek, G.; et al. Access to community care for people with dementia and their informal carers: Case vignettes for a European comparison of structures and common pathways to formal care. *Zeitschrift fur Gerontologie und Geriatrie* **2018**, *5*, 530–536. [CrossRef] [PubMed]
29. Malmgren, A.; Schmidt, S.M.; Nilsson, M.H.; Carlsson, G.; Liwander, A.; Dahlgren Bergström, C.; Olivetti, P.; Johansson, P.; Chiatti, C. TECH@HOME Research Group. The TECH@HOME study, a technological intervention to reduce caregiver burden for informal caregivers of people with dementia: Study protocol for a randomized controlled trial. *Trials* **2017**, *18*, 63. [CrossRef] [PubMed]
30. Chiatti, C.; Masera, F.; Rimland, J.M.; Cherubini, A.; Scarpino, O.; Spazzafumo, L.; Lattanzio, F. The UP-TECH project, an intervention to support caregivers of Alzheimer's disease patients in Italy: Study protocol for a randomized controlled trial. *Trials* **2013**, *14*, 155. [CrossRef] [PubMed]
31. Folstein, M.F.; Folstein, S.E.; McHugh, P.R. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* **1975**, *12*, 189–198. [CrossRef]
32. Andersen, R.; Newman, J.F. Societal and individual determinants of medical care utilization in the United States. *Milbank Mem. Fund Q. Health Soc.* **1973**, *51*, 95–124. [CrossRef] [PubMed]
33. Morris, J.N.; Fries, B.E.; Morris, S.A. Scaling ADLs within the MDS. *J. Gerontol. A Biol. Sci. Med. Sci.* **1999**, *54*, M546–M553. [CrossRef] [PubMed]
34. Morris, J.N.; Carpenter, I.; Berg, K.; Jones, R.N. Outcome measures for use with home care clients. *Can. J. Aging* **2000**, *19*, 87–105. [CrossRef]
35. Zigmond, A.; Snaith, R. The hospital anxiety and depression scale. *Acta Psychiatr. Scand.* **1983**, *67*, 361–370. [CrossRef] [PubMed]
36. Wimo, A.; Gustavsson, A.; Jonsson, L.; Winblad, B.; Hsu, M.A.; Gannon, B. Application of resource utilization in dementia (RUD) instrument in a global setting. *Alzheimers Dement.* **2013**, *9*, 429–435. [CrossRef] [PubMed]
37. Hamilton, L.C. *Statistics with Stata (Updated for Version 10)*; Thomson Brooks/Cole: Belmont, CA, USA, 2008.
38. Brandt, M.; Haberkern, K.; Szydlik, M. Intergenerational help and care in Europe. *Eur. Sociol. Rev.* **2009**, *25*, 585–601. [CrossRef]
39. Janssen, E.P.; de Vugt, M.; Köhler, S.; Wolfs, C.; Kerpershoek, L.; Handels, R.L.; Orrell, M.; Woods, B.; Jelley, H.; Stephan, A.; et al. Caregiver profiles in dementia related to quality of life, depression and perseverance time in the European Actifcare study: The importance of social health. *Aging Ment. Health* **2017**, *21*, 49–57. [CrossRef] [PubMed]
40. Sutcliffe, C.; Giebel, C.; Bleijlevens, M.; Lethin, C.; Stolt, M.; Saks, K.; Soto, M.E.; Meyer, G.; Zabalegui, A.; Chester, H.; et al. Caring for a person with dementia on the margins of long-term care: A perspective on burden from 8 European countries. *J. Am. Med. Dir. Assoc.* **2017**, *18*, 967–973. [CrossRef] [PubMed]
41. Wimo, A.; von Strauss, E.; Nordberg, G.; Sassi, F.; Johansson, L. Time spent on informal and formal care giving for persons with dementia in Sweden. *Health Policy* **2002**, *61*, 255–268. [CrossRef]
42. Haro, J.M.; Kahle-Wrobleski, K.; Bruno, G.; Belger, M.; Dell'Agnello, G.; Dodel, R.; Jones, R.W.; Reed, C.C.; Vellas, B.; Wimo, A.; et al. Analysis of burden in caregivers of people with Alzheimer's disease using self-report and supervision hours. *J. Nutr. Health Aging* **2014**, *18*, 677–684. [CrossRef] [PubMed]
43. Neubauer, S.; Holle, R.; Menn, P.; Grossfeld-Schmitz, M.; Graesel, E. Measurement of informal care time in a study of patients with dementia. *Int. Psychogeriatr.* **2008**, *20*, 1160–1176. [CrossRef] [PubMed]
44. Aguglia, E.; Onor, M.L.; Trevisiol, M.; Negro, C.; Saina, M.; Maso, E. Stress in the caregivers of Alzheimer's patients: An experimental investigation in Italy. *Am. J. Alzheimers Dis. Other Demen.* **2004**, *19*, 248–252. [CrossRef] [PubMed]
45. Cangiano, A. Elder care and migrant labor in Europe: A demographic outlook. *Popul. Dev. Rev.* **2014**, *40*, 131–154. [CrossRef]
46. Isengard, B.; Szydlik, M. Living apart (or) together? Coresidence of elderly parents and their adult children in Europe. *Res. Aging* **2012**, *34*, 449–474. [CrossRef]
47. Gunnarsson, E. The welfare state, the individual and the need for care: Older people´s views. *Int. J. Soc. Welf.* **2009**, *18*, 252–259. [CrossRef]
48. Bremer, P.; Cabrera, E.; Leino-Kilpi, H.; Lethin, C.; Saks, K.; Sutcliffe, C.; Soto, M.; Zwakhalen, S.M.; Wübker, A. RightTimePlaceCare Consortium. Informal dementia care: Consequences for caregivers' health and health care use in 8 European countries. *Health Policy* **2015**, *119*, 1459–1471. [CrossRef] [PubMed]

49. Ory, M.; Hoffman, R.R.; Yee, J.L.; Tennstedt, S.; Schulz, R. Prevalence and impact of caregiving: A detailed comparison between dementia and nondementia caregivers. *Gerontologist* **1999**, *39*, 177–185. [CrossRef] [PubMed]
50. Hajek, A.; Brettschneider, C.; Ernst, A.; Posselt, T.; Wiese, B.; Prokein, J.; Weyerer, S.; Werle, J.; Fuchs, A.; Pentzek, M.; et al. Longitudinal predictors of informal and formal caregiving time in community-dwelling dementia patients. *Soc. Psychiatry Psychiatr. Epidemiol.* **2016**, *51*, 607–616. [CrossRef] [PubMed]
51. Germain, S.; Adam, S.; Olivier, C.; Cash, H.; Ousset, P.J.; Andrieu, S.; Vellas, B.; Meulemans, T.; Reynish, E.; Salmon, E.; et al. Does cognitive impairment influence burden in caregivers of patients with Alzheimer's disease? *J. Alzheimers Dis.* **2009**, *17*, 105–114. [CrossRef] [PubMed]
52. Raggi, A.; Leonardi, M. Burden and cost of neurological diseases: A European North-South comparison. *Acta Neurol. Scand.* **2015**, *132*, 16–22. [CrossRef] [PubMed]
53. Costa, G. Long-term care Italian policies: A case of inertial institutional change. In *Reforms in Long-Term Care Policies in Europe*; Ranci, C., Pavolini, E., Eds.; Springer-Verlag: New York, NY, USA, 2013; pp. 221–242.
54. Bökberg, C.; Ahlström, G.; Karlsson, S. Utilisation of formal and informal care and services at home among persons with dementia: A cross-sectional study. *Scand. J. Car. Sci.* **2018**, *32*, 843–851. [CrossRef] [PubMed]
55. Chiatti, C.; Rimland, J.M.; Bonfranceschi, F.; Masera, F.; Bustacchini, S.; Cassetta, L.; Lattanzio, F. UP-TECH Research Group. The UP-TECH project, an intervention to support caregivers of Alzheimer's disease patients in Italy: preliminary findings on recruitment and caregiving burden in the baseline population. *Aging Ment. Health* **2015**, *19*, 517–525. [CrossRef] [PubMed]
56. Kerpershoek, L.; De Vugt, M.; Wolfs, C.; Jelley, H.; Orrell, M.; Woods, B.; Stephan, A.; Bieber, A.; Meyer, G.; Engedal, K.; et al. Access to timely formal dementia care in Europe: Protocol of the Actifcare (Access to Timely Formal Care) study. *BMC Health Serv. Res.* **2016**, *16*, 423.
57. Shutes, I.; Chiatti, C. Migrant labour and the marketisation of care for older people: The employment of migrant care workers by families and service providers. *J. Eur. Soc Policy* **2012**, *22*, 392–405. [CrossRef]

International Journal of
Environmental Research and Public Health

MDPI

Article

Relationships between Physical Activity and Quality of Life in Pregnant Women in the Second and Third Trimester

Justyna Krzepota [1,*], Dorota Sadowska [2] and Elżbieta Biernat [3]

[1] Department of Physical Culture and Health Promotion, University of Szczecin, al. Piastów 40b, blok 6, 71-065 Szczecin, Poland
[2] Department of Physiology, Institute of Sport—National Research Institute, ul. Trylogii 2/16, 01-982 Warsaw, Poland; sadowska.dorota@hotmail.com
[3] Department of Tourism, Collegium of World Economy, Warsaw School of Economics, al. Niepodległości 162, 02-554 Warsaw, Poland; elzbieta.biernat@sgh.waw.pl
* Correspondence: justyna.krzepota@usz.edu.pl

Received: 26 October 2018; Accepted: 3 December 2018; Published: 5 December 2018

Abstract: Background: The problem of an inadequate level of physical activity (PA) is important in the context of its relationship with the quality of life (QoL) of pregnant women. The aim of this study was to analyze the relationships between PA and QoL among pregnant women. Methods: The study analyzed 346 questionnaires filled in by pregnant women (157 in the second trimester and 189 in the third). The tool used for assessment of PA was the Pregnancy Physical Activity Questionnaire-Polish version (PPAQ-PL). The quality of life (QoL) was assessed by the World Health Organization Quality of Life Questionnaire-short form (WHOQoL-Bref). The results obtained from the PPAQ-PL and WHOQoL-Bref questionnaires for women in the second and third trimesters of pregnancy and intergroup differences were analyzed. Results: There was a significant correlation in the group of women in the second trimester of pregnancy between quality of life in the *physical health domain* and the intensity and type of physical activity. The women who rated their quality of life higher in this domain declared higher energy expenditures (EE) associated with *vigorous activity* ($R = 0.159$, $p \leq 0.05$), as well as with *occupational activity* ($R = 0.166$; $p \leq 0.05$) and *sport/exercise* activity ($R = 0.187$; $p \leq 0.05$). In women in the third trimester, higher EE related to *sport/exercise activity* coincided with higher assessments of the *overall quality of life* ($R = 0.149$, $p \leq 0.05$) and *general health* ($R = 0.170$, $p \leq 0.05$). In the case of the *psychological domain* ($R = 0.161$, $p \leq 0.05$) and *social relationship domain* ($R = 0.188$; $p \leq 0.05$) of QoL, positive correlations occurred with EE related to *vigorous activity*. In contrast, high assessment of *physical health domain* coincided with higher EE related to *occupational activity* ($R = 0.174$; $p \leq 0.05$). Conclusions: Our study makes an important contribution to knowledge concerning the correlations between PA and QoL in pregnancy. The results suggest the need for improvement in prenatal care and promotion of PA programs for pregnant women.

Keywords: physical activity in pregnancy; PPAQ; quality of life; WHOQoL-Bref; pregnant women

1. Introduction

The prevalence of an insufficient level of physical activity (PA) in pregnant women has been demonstrated in studies using representative samples in different countries [1–4]. Despite the well-documented benefits of involvement in PA in this period of life [5–7], it is emphasized that pregnancy continues to be one of the causes of a substantial reduction in PA [8–10].

Over the past two decades, most studies which have focused on these problems have estimated that most pregnant women do not participate in recommended PA. Clark and Gross [11] demonstrated that 39% of women who reported participating in some forms of weekly exercise before pregnancy did not report pursuing any similar activities during pregnancy. Similarly, Fell et al. [12], in a comparative study of women's levels of PA during early pregnancy and during the year before pregnancy, observed that most women reduced their PA levels during the first 20 weeks of pregnancy compared with their level of activity during the year prior to pregnancy.

According to the data presented by Evenson et al. [1], only 15.8% of pregnant women in the USA reported being active, in accordance with the recommendations. In a cohort study conducted in Brazil, Domingues and Barros [2] estimated that only 4.3% of women were active during the whole pregnancy. Furthermore, [13] found that only 14.6% of women in mid-pregnancy in Norway were involved in the exercise $\geq$3 times a week, >20 min at moderate intensity. According to Santo et al. [4], merely 9% of 1584 pregnant women met the American College of Obstetrics and Gynecology ACOG guidelines. Furthermore, nearly half of the women reported PA < 1 day/week during the third trimester. Similarly, Haakstad et al. [3] found that only 11% of pregnant women followed ACOG guidelines regarding PA.

The problem of inadequate levels of PA is particularly important in the context of the relationship of PA to the quality of life (QoL) of pregnant women. Recently published studies have provided important information on changes that occur in health-related QOL (HRQoL), both during pregnancy and [14] in the perinatal period [15]. A problem that remains to be discussed is the explanation of the relationships between PA and the quality of life of pregnant women [16–20]. No unambiguous findings have been published to date in this area, since interpretation is difficult due to the use of different measurement tools, both for the evaluation of PA during pregnancy and the quality of life. For example, some researchers have used the results obtained with the Global Physical Activity Questionnaire (GPAQ) and 36-Item Short-Form Health Survey (SF-36) [20], while others have used the Pregnancy Physical Activity Questionnaire (PPAQ) and the abridged World Health Organization Quality of Life (WHOQoL-Bref) [16,18]. In addition to these tools, other questionnaires have been popular and are frequently used in the assessment of the quality of life, with an extensive review and discussion of their use being presented by Mogos et al. [21].

Despite many analyses, the problem of correlations between PA and the quality of life of pregnant women remains to be solved and still raises many questions that need to be addressed. To the best of our knowledge, the few studies that have used the PPAQ and WHOOoL-Bref tools [16,18] have failed to analyze the Polish population. Therefore, the aim of this study was to obtain information about the relationships between PA and QoL among pregnant women in Poland. The QoL was evaluated by a reliable questionnaire used in surveys of pregnant women (WHOQoL-Bref) [16,22–24], whereas PA was measured using the increasingly popular PPAQ questionnaire [25–34]. Furthermore, we hope that the choice of the above measurement tools will allow for replication of the study results and comparison with future research conducted on a wider scale.

2. Material and Methods

The study analyzed 346 questionnaires filled in correctly by pregnant women (157 in the second trimester and 189 in the third). The data were directly distributed among pregnant women who participated in activities for pregnant women organized within antenatal classes or in fitness clubs in Szczecin and Warsaw, Poland. The survey was anonymous and conducted only in places where consent was obtained. The respondents consisted only of volunteers. In each case, the survey supervisor presented the purpose and scope of the research to the respondents and instructed them on how to complete the questionnaires. The method "pen and paper" was used in the study. The respondents were given unlimited time to fill in the questionnaires.

The analysis excluded three questionnaires in which the first trimester of pregnancy was declared. Furthermore, the questionnaires which were not filled in completely or were filled in incorrectly were rejected (contrary to the instructions). 400 questionnaires were distributed to women, 346 of which were fully and properly completed. 21 questionnaires were incompletely or incorrectly filled in, and 29 questionnaires were not returned. The project was approved by a local Bioethics Committee.

The age of the respondents was 30.4 ± 3.6 years. Over 90% of study participants were college graduates (91.9%), 8.1% graduated from high school, 81.5% were married, 17.3% were single, while 1.2% were divorced. Furthermore, 88.2% were childless, 10.1% had 1 child and 1.7% had 2 or more children.

The tool used for assessment of PA was the Pregnancy Physical Activity Questionnaire-Polish version (PPAQ-PL) [32–34]. The original version of PPAQ was designed by Chasan-Taber et al. [35]. The questionnaire is used by pregnant women to self-assess their PA in the current trimester. In PPAQ-PL, the respondents were asked to report the time spent on participation in 32 types of activities grouped under the following categories: *household/caregiving* (12 activities), *occupational* (5 activities), *sports/exercise* (9 activities = 7 questions + two open questions, allowing the respondent to add any activities not previously listed), *transportation* (3 activities), and *inactivity* (3 activities). The questionnaire measures energy expenditure related to *total activity* and *total activity of light intensity and above* expressed in Metabolic Equivalent of Task (MET) units (MET-h·week^{-1}). Based on the energy expenditure, each of these activities was additionally classified according to intensity: (a) *sedentary activity* [<1.5 METs], (b) *light intensity activity* [1.5–<3.0 METs], (c) *moderate intensity activity* [3.0–6.0 METs], (d) *vigorous-intensity activity* [>6.0 METs]. The MET values were assigned according to the values presented in the questionnaire instruction and the Compendium of Physical Activities [36].

The methodological basis for the assessment of the quality of life (QoL) was provided by the abridged World Health Organization Quality of Life Questionnaire (WHOQoL-Bref), a Polish version provided by Wołowicka and Jaracz [37]. The WHOQoL-Bref questionnaire assesses self-reported QoL and general health of respondents. The WHOQoL-Bref questionnaire consists of 26 questions. The first two questions were analyzed separately. They concerned self-assessed *overall quality of life* and *general health* of the respondents. The remaining 24 questions assessed four domains of the QoL (*physical health domain*: 7 questions, *psychological domain*: 6 questions, *social relationships domain*: 3 questions, and *environmental domain*: 8 questions). The respondents were asked to mark their answers using a five-level rating scale (from 1 to 5 points, in a positive direction: the higher the number of points, the better quality of life). The QoL in the domains was expressed as mean values, calculated according to the key and guidelines provided by the authors [37].

STATISTICA 12.5 software was used for statistical analysis. The significance of the analyzed variables in women in the second and third trimesters of pregnancy was evaluated by means of the Mann-Whitney U-test. Correlations between the variables were analyzed using Spearman's rank correlation test, with correlation coefficients calculated for each pair of variables. The level of statistical significance was set at $p \leq 0.05$.

3. Results

The results obtained from the PPAQ-PL and WHOQoL-Bref questionnaires for the women in the second and third trimesters of pregnancy and intergroup differences (between the second and third trimesters) are presented in Tables 1 and 2.

No statistically significant differences in the declared values of total energy expenditure (*total activity* and *total activity of light intensity and above*) were found among the women surveyed (Table 1). However, it was shown that the PA intensity differed significantly ($p \leq 0.05$) depending on the trimester of pregnancy. This concerns in particular *sedentary activity* (group of women in the second trimester: 30.4 ± 21.6 MET-h/week, group of women in the third trimester: 35.5 ± 23.1 MET-h/week) and *moderate activity* (42.7 ± 45.2 and 39.4 ± 52.8 MET-h/week, respectively).

Analysis of the type of activities showed that the MET-h/week values did not differ between the groups studied for *household/caregiving*, *occupational activity*, *transportation* and *sports/exercise*. However, it was noticeable that higher energy expenditure ($p \leq 0.01$) in women in the third trimester of pregnancy was observed for the activities related to *inactivity* (51.4 ± 28.87 MET-h/week) compared to those in the second trimester (42.7 ± 24.9 MET-h/week).

Table 1. Means (M), standard deviations (SD), medians, and 25th and 75th percentiles for the Pregnancy Physical Activity Questionnaire (PPAQ-PL) and intergroup comparisons (the second and third trimesters) using the Mann–Whitney U-test.

Factors	Trimester	PPAQ-PL (MET-h/week)				
		M ± SD	25th	Median	75th	*p*
Total Activity Scores:						
Total activity	2nd	183.3 ± 75.2	129.1	166.8	220.7	0.721
	3rd	192.1 ± 99.3	130.4	168.8	227.2	
Total activity of light intensity and above	2nd	152.9 ± 75.9	99.7	143.3	184.2	0.838
	3rd	156.6 ± 95.5	99.2	136.9	188.7	
by Intensity:						
Sedentary (<1.5 METs)	2nd	30.4 ± 21.6	15.4	29.4	43.4	0.025 *
	3rd	35.5 ± 23.1	17.9	29.4	46.2	
Light (1.5–<3.0 METs)	2nd	110.4 ± 547.4	73.3	104.7	140.7	0.355
	3rd	117.8 ± 55.3	73.5	109.8	151.6	
Moderate (3.0–6.0 METs)	2nd	42.7 ± 45.2	17.9	30.9	50.2	0.022 *
	3rd	39.4 ± 52.8	13.7	23.7	46.9	
Vigorous (>6.0 METs)	2nd	1.78 ± 4.2	0.0	0.0	0.8	0.937
	3rd	1.6 ± 4.3	0.0	0.0	0.8	
by Type:						
Household/Caregiving	2nd	56.4 ± 39.8	33.6	43.5	67.6	0.316
	3rd	59.1 ± 40.1	33.9	50.1	70.4	
Occupational activity	2nd	37.8 ± 60.9	0.00	0.00	74.9	0.337
	3rd	33.2 ± 61.4	0.0	0.0	67.2	
Sports/Exercise	2nd	12.8 ± 11.8	7.2	12.8	21.2	0.212
	3rd	14.5 ± 12.7	5.0	11.0	19.7	
Transportation	2nd	30.9 ± 31.1	10.7	21.4	36.8	0.505
	3rd	33.84 ± 34.5	10.7	22.6	42.0	
Inactivity	2nd	42.7 ± 24.9	24.2	38.2	57.4	0.005 **
	3rd	51.4 ± 28.87	29.6	44.9	70.0	

MET-Metabolic Equivalent of Task; * $p \leq 0.05$; ** $p \leq 0.01$.

The WHOQoL-Bref results indicated no differences in the self-rated quality of life of the women surveyed (Table 2). Pregnant women in both the second and third trimesters rated their quality of life in the *psychological domain* as the highest (16.48 ± 1.88 in the second trimester and 16.56 ± 1.64 in the third trimester), whereas the lowest ratings were recorded for the *environmental domain* (15.89 ± 1.96 and 15.78 ± 1.91, respectively).

The next stage of the statistical analysis focused on investigating whether there is a correlation between intensity and types of PA assessed using PPAQ-PL and domains of the QoL assessed using WHOQoL-Bref in women in the second (Table 3) and third (Table 4) trimesters of pregnancy.

Table 2. Medians, 25th and 75th percentiles, means (M) and standard deviations (SD) for the Quality of Life-Bref Questionnaire (WHOQoL-Bref) and intergroup comparisons (2nd and 3rd trimesters) using the Mann–Whitney U-test.

WHOQoL-Bref Factors	Trimester	M ± SD	25th	Median	75th	*p*
Overall quality of life	2nd	4.34 ± 0.61	4.00	4.00	5.00	0.300
	3rd	4.41 ± 0.65	4.00	4.00	5.00	
General health	2nd	4.13 ± 0.67	4.00	5.00	5.00	0.507
	3rd	4.19 ± 0.65	4.00	4.00	5.00	
WHO Domain						
Physical health	2nd	16.06 ± 2.16	15.00	16.00	18.00	0.187
	3rd	15.80 ± 01.99	14.00	16.00	17.00	
Psychological	2nd	16.48 ± 1.88	15.00	17.00	17.00	0.552
	3rd	16.56 ± 1.64	15.00	17.00	17.00	
Social relationships	2nd	16.47 ± 2.21	16.00	16.00	19.00	0.274
	3rd	16.14 ± 2.60	15.00	16.00	17.00	
Environmental	2nd	15.89 ± 1.96	15.00	16.00	17.00	0.781
	3rd	15.78 ± 1.91	15.00	16.00	17.00	

Table 3. Spearman's correlation coefficients between the Pregnancy Physical Activity questionnaire (PPAQ-PL) and Quality of Life-Bref Questionnaire (WHOQoL-Bref) in women in the second trimester of pregnancy.

WHOQoL-Bref / PPAQ-PL	Overall Quality of Life	General Health	WHOQoL-Bref Domain			
			Physical Health	psychological	Social Relationships	Environmental
Total Activity Scores:						
Total activity	−0.003	0.005	0.057	0.020	0.071	−0.061
Total activity of light intensity and above	0.002	0.007	0.089	0.037	0.129	−0.037
by Intensity						
Sedentary (<1.5 METs)	−0.040	−0.028	−0.041	−0.072	−0.104	−0.073
Light (1.5–<3.0 METs)	0.029	−0.025	0.129	0.061	0.091	0.011
Moderate (3.0–6.0 METs)	−0.035	0.043	0.029	−0.003	0.128	−0.073
Vigorous (>6.0 METs)	−0.065	0.043	0.159 *	0.072	0.122	0.079
by Type						
Household/Caregiving	0.0125	−0.059	−0.116	−0.143	−0.008	−0.035
Occupational activity	0.003	0.047	0.166 *	0.117	0.105	0.040
Sports/Exercise	0.035	0.081	0.187 *	0.103	0.153	0.103 *
Transportation	0.020	0.053	0.111	0.155	0.166 *	0.054
Inactivity	0.022	−0.017	−0.151 *	−0.075	−0.097	−0.004

MET-Metabolic Equivalent of Task; * $p \leq 0.05$.

Table 4. Spearman's correlation coefficients between the Pregnancy Physical Activity questionnaire (PPAQ-PL) and Quality of Life-Bref Questionnaire (WHOQoL-Bref) in women in the third trimester of pregnancy.

PPAQ-PL \ WHOQoL-Bref	Overall Quality of Life	General Health	WHOQoL-Bref Domain			
			Physical Health	Psychological	Social Relationships	Environmental
Total activity	−0.068	0.052	0.048	0.003	0.036	−0.094
Total activity of light intensity and above	−0.049	0.049	0.085	−0.029	−0.009	−0.073
Sedentary (<1.5 METs)	−0.090	−0.018	−0.125	0.020	0.0463	−0.068
Light (1.5–<3.0 METs)	−0.053	−0.013	0.065	−0.068	−0.019	−0.133
Moderate (3.0–6.0 METs)	0.010	0.140	0.117	0.054	0.040	0.099
Vigorous (>6.0 METs)	0.029	0.041	0.073	0.161 *	0.188 *	−0.072
Household/Caregiving	−0.112	0.056	−0.023	0.023	−0.037	−0.059
Occupational activity	0.050	0.038	0.174 *	−0.024	−0.013	0.034
Sports/Exercise	0.149 *	0.170 *	0.101	0.087	0.067	0.131
Transportation	−0.064	−0.0083	0.068	−0.060	0.033	−0.124
Inactivity	−0.077	−0.034	−0.128	−0.028	0.028	−0.096

MET-Metabolic Equivalent of Task; * $p \leq 0.05$.

There was a significant correlation in the group of women in the second trimester of pregnancy for *physical health domain* with intensity and the type of activities (Table 3). The women who rated their quality of life higher in this domain declared higher energy expenditures associated with *vigorous activity* (R = 0.159, $p \leq 0.05$), as well as with *occupational activity* (R = 0.166; $p \leq 0.05$) and *sport/exercise* activity (R = 0.187; $p \leq 0.05$).

Furthermore, a negative correlation was found between *physical health domain* and *inactivity* (R = −0.151, $p \leq 0.05$). This means that higher assessment of quality of life in this domain coincided with lower energy expenditure related to *inactivity*. Individual positive correlations were also documented between *social relationship domain* and *transportation activity* (R = 0.166, $p \leq 0.05$) and between *environmental domain* and *sport/exercise activity* (R = 0.103, $p \leq 0.05$).

In women in the third trimester, higher energy expenditures related to *sport/exercise activity* coincided with higher assessments of the *overall quality of life* (R = 0.149, $p \leq 0.05$) and *general health* (R = 0.170, $p \leq 0.05$). In the case of the *psychological domain* (R = 0.161, $p \leq 0.05$) and *social relationship domain* (R = 0.188; $p \leq 0.05$) of QoL, positive correlations occurred with energy expenditure related to *vigorous activity*. In contrast, high assessment of the *physical health domain* coincided with higher energy expenditure related to *occupational activity* (R = 0.174; $p \leq 0.05$).

4. Discussion

The aim of this study was to investigate relationships between PA and QoL in women in the second and third trimesters of pregnancy. The respondents' declarations from the PPAQ-PL and WHOQoL-Bref questionnaires were analyzed. Although studies have used these questionnaires in recent years [16,18], this is, to our knowledge, the first such study in Poland. Such research is also important because previous findings concerning PA in women in various stages of pregnancy have been ambiguous. Some authors have found increases in PA, for example, Huberty et al. [10] in the first and second trimesters, and Ko, Chen, Lin [38] from the second trimester, while others have documented a decline in overall PA [38] in the first trimester, Evenson and Wen [39] in the third trimester, and a decreasing percentage of physically active women in consecutive trimesters of pregnancy [2].

Borodulin et al. [8] argued that the overall physical activity level slightly decreased between 17–22 and 27–30 weeks of gestation, particularly in duration and volume of care, outdoor household, and recreational activity. Santos et al. [40] emphasized that a decline in PA from the first to the second trimester concerned total, light and moderate intensities, while Richardsen et al. [41], documented a decline in moderate and vigorous PA in the period between early pregnancy and mid-pregnancy.

Similar to findings published by Mourady et al. [16], our findings showed that the respondents in different trimesters of pregnancy did not differ in terms of *total PA*, and *total activity of light intensity and above*. However, they differed in the intensity of activities, especially in *energy expenditure* during *moderate activity* (in favor of those in the second trimester). The women in the third trimester of pregnancy reported significantly more energy expenditure on *sedentary activity* and *inactivity*, which is not an isolated phenomenon in the world [10]. According to Santos et al. [40], energy expenditure for particular types of PA (e.g., occupational activity, household activity, sports activity) changes significantly in individual trimesters of pregnancy. Pregnant women spent most of their weekly time on domestic, occupational and leisure time activities, except for sports activities. Similarly, the majority of the respondents' energy expenditure in the respondents surveyed in our study was spent on *household/caregiving activities*. This suggests that despite many campaigns to raise awareness of Polish women, such as: Pregnancy: Conscious Maternity, Find Out Whether You Are a Conscious Parent, or Different State, Different Treatment (Ciąża-Świadome Macierzyństwo, Sprawdź czy jesteś świadomym rodzicem, Odmienny stan, odmienne traktowanie), the role of the benefits of PA during pregnancy remains underestimated. In the public's opinion, healthy nutrition is more often perceived as more important for the health of mothers and children than involvement in physical activity [42]. Therefore, the low levels of energy expenditure related to *sport/exercise activity* found in the present study in both groups of women surveyed seem unsurprising. Perhaps, as argued by Clarke and Gross [11] and Guelfi et al. [43], women perceive relaxation as a safer behavior, which is more beneficial for ensuring full-term pregnancy rather than regular exercise and maintaining an active lifestyle. The reasons also include misconceptions about physical exercise [44], the inconveniences of late pregnancy, fatigue, poor moods, or being absorbed in numerous occupational duties [45,46]. There are also other determinants that represent barriers to physical exercise. For example, low physical activity during pregnancy occurs more often in mature and married women, as well as those financially less well-off and the less educated [47]. It seems, however, that regardless of the adversities, the role of physicians is also critical as they have the greatest effect on the beliefs of pregnant women, including their ideas on exercise during pregnancy [48]. Unfortunately, as Santos et al. [40] argued, medical staff often fail to recommend PA during pregnancy. Furthermore, according to Krans et al. [48], a low percentage of physicians help their patients to prepare physical exercise programs. Despite their knowledge, physicians do not always explain the need for physical exercise, both during pregnancy and in later decades of life [49]. They do not inform patients that it is necessary to consult both physicians and coaches before starting physical exercise in order to exclude medical contraindications and choose the right type of exercises and the load.

Knowledge about the quality of life plays a significant role both in diagnosis and patient care [50]. Despite being ambiguous, studies have widely documented the correlations between physical activity and quality of life. According to the literature review published by Poudevigne et al. [47], there is scientific evidence that inactivity during pregnancy is associated with poorer mood, whereas increasing participation in sports or physical activity from the period of pregnancy to that after birth leads to better overall well-being [51]. Mourady et al. [16] demonstrated that total and light intensity of PA are positively significantly correlated with the psychological domain of quality of life and social relationships; while sedentary PA is significantly correlated only with social relationships. Arizabaleta et al. [19] documented improvements in HRQoL in the physical component summary, physical function domain, the bodily pain domain and general health domain following a three-month program of aerobic exercise. However, there are also publications that showed no improvements in self-rated QoL caused by regular exercise such as water exercise [22].

Analysis of QoL of the women surveyed showed that there were no significant differences in self-rated domains of WHOOoL-Bref between pregnant women in the second and third trimesters of pregnancy. Similarly, Mourady et al. [16], who analyzed all the trimesters, also found no differences except in the environmental domain. In this case, the quality of life was significantly higher in the women in the third trimester compared to those in the first trimester.

In our study, we found higher QoL scores in the *environment domain* in women in the second trimester who declared higher *sport/exercise activity.* It should also be noted that *sport/exercise activity* of the respondents studied was also positively correlated with the *physical domain* in women in the second trimester and with *overall quality of life* and *general health* in women in the third.

Our findings are consistent with those presented by Mourady et al. [16], who showed that sports/exercise was significantly correlated with the majority of quality of life domains such as general quality of life, physical and psychological health, social relationships and the environmental domain. This is unsurprising since apart from its well-documented health benefits, sport [5–7] offers joy, relaxation and enhances psychological well-being [52]. Obviously, there have also been studies in the literature that have failed to support such findings. For example, Gustafsson et al. [53] indicated that a 12-week exercise program including aerobic and strength training during pregnancy is unlikely to influence the psychological and self-perceived well-being of healthy pregnant women. Kolu et al. [17] showed a decline in the overall HRQoL index during pregnancy, although they emphasized that this decrease was lower in women who were physically active during pregnancy. Nascimento et al. [23] argued that physical exercise does not significantly affect the perception of the quality of life of pregnant women because, regardless of their participation in the exercise program, the quality of life of women (in the physical and social domains) during pregnancy fell significantly.

An interesting finding of our study is that pregnant women who assessed QoL as higher in the field of *physical health* (both in the second and third trimester) were characterized by a higher energy expenditure during *occupational activity*. The explanation for this finding seems to be obvious; women with better self-rated physical well-being tend to work more. We are aware, however, that the explanation for this phenomenon may be more complex. The study published by Blum et al. [51] showed that women with older infants or no other children reported higher household/caregiving and lower occupation pre-pregnancy to postpartum activity. Physical activity in pregnancy may depend on the socio-economic status and support of a partner, friends or family [54,55]. These factors may, to a large extent, determine the quality of life [56]. Unfortunately, due to the lack of the above-mentioned information in our study, the impact of these factors was impossible to determine. Some limitations of this study should be mentioned and taken into consideration. First of all, a limited number of participants and the place where the women were selected (fitness clubs and antenatal classes) lead to a lack of representativeness of the total population with possible effects on the results. This in turn makes it impossible to draw general conclusions for the whole population of pregnant women. Furthermore, the lack of detailed information on socio-economic and psychological factors and data about pathologies makes the interpretation of the results difficult.

5. Conclusions

In conclusion, our study (the first study in Poland that has used reliable, internationally recognized questionnaires (PPAQ-PL and WHOQoL-Bref) makes an important contribution to the knowledge concerning the correlations between PA and QoL in women during different periods of pregnancy. The study showed that *total activity* and *total activity of light intensity and above* did not differentiate between women in the second and third trimesters of pregnancy. However, it indicated higher values of *moderate activity* in women in the second trimester of pregnancy and higher values of *sedentary activity* and *inactivity* in women in the third trimester. Our findings concerning the relationships between physical activity and quality of life should be approached with caution, due to the low values of correlation coefficients. The low MET values for sport/exercise recorded in both groups of women can indicate the need for improving the prenatal care, especially in terms of promotion of physical activity programs for pregnant women and encouraging women to participate in these programs. In terms of intervention activities, special attention should be given to barriers existing at the level of the provider, the patient, and practice [57]. Researchers [58] have specified concrete actions that should be taken by prenatal care providers in order to promote prenatal PA, suggesting, among other things, providing information by healthcare providers about both guidelines and contraindications for the involvement in physical activity during pregnancy. Most guidelines from around the world, gathered by Evenson et al. [59], promoted moderate-intensity physical activity during pregnancy and defined its frequency and duration/time. The latest physical activity guidelines published by the U.S. Department of Health and Human Services USDHHS in 2018 [60], indicate 150 min (2 h and 30 min) of moderate-intensity aerobic activity a week during pregnancy and the postpartum period. Recommending aerobic activity should be spread throughout the week.

Undoubtedly, our study does not exhaust the problems discussed but its findings emphasize the need for raising awareness of the importance of physical activity during pregnancy. We believe that further research on a larger sample with the consideration of socio-economic factors and a comprehensive inventory of pregnancy-related symptoms, along with a mechanism for assessing their effect on function [61] is needed to provide deeper understanding and identify correlations between PA and determinants of QoL in pregnant women.

Author Contributions: J.K. prepared the study design, searched the literature, wrote the background and discussion of the study, gathered the necessary data, prepared it for analysis, described the results of the study and prepared the manuscript. D.S. searched the literature, refined the data, performed the statistical analysis, and described its results. E.B. searched the literature, corrected and improved the manuscript of the study for the final version. All authors have read and approved the final manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors would like to thank all the women who participated in the study and Katarzyna Sempolska, who contributed to data acquisition.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Evenson, K.R.; Savitz, A.; Huston, S.L. Leisure-time physical activity among pregnant women in the US. *Paediatr. Perinat. Epidemiol.* **2004**, *18*, 400–407. [CrossRef] [PubMed]
2. Domingues, M.R.; Barros, A.J.D. Leisure-time physical activity during pregnancy in the 2004 Pelotas Birth Cohort Study. *Rev. Saude Publica* **2007**, *41*, 173–180. [CrossRef] [PubMed]
3. Haakstad, L.A.H.; Voldner, N.; Henriksen, T.; Bø, K. Why do pregnant women stop exercising in the third trimester? *Acta Obstet. Gynecol. Scand.* **2009**, *88*, 1267–1275. [CrossRef] [PubMed]
4. Santo, E.C.; Forbes, P.W.; Oken, E.; Belfort, M.B. Determinants of physical activity frequency and provider advice during pregnancy. *BMC Pregnancy Childbirth* **2017**, *17*, 286. [CrossRef] [PubMed]
5. Melzer, K.; Schutz, Y.; Soehnchen, N.; Othenin-Girard, V.; Martinez de Tejada, B.; Irion, O.; Boulvain, M.; Kayser, B. Effects of recommended levels of physical activity on pregnancy outcomes. *Am. J. Obstet. Gynecol.* **2010**, *202*, 266.e1-6. [CrossRef] [PubMed]
6. Hinman, S.K.; Smith, K.B.; Quillen, D.M.; Smith, M.S. Exercise in Pregnancy: A Clinical Review. *Sports Health* **2015**, *7*, 527–531. [CrossRef] [PubMed]
7. Brown, W. The benefits of physical activity during pregnancy. *J. Sci. Med. Sport* **2002**, *5*, 37–45. [CrossRef]
8. Borodulin, K.; Evenson, K.R.; Wen, F.; Herring, A.H.; Benson, A. Physical activity patterns during pregnancy. *Med. Sci. Sports Exerc.* **2008**, *40*, 1901–1908. [CrossRef]
9. Clarke, P.E.; Rousham, E.K.; Gross, H.; Halligan, A.W.F.; Bosio, P. Activity patterns and time allocation during pregnancy: A longitudinal study of British women. *Ann. Hum. Biol.* **2005**, *32*, 247–258. [CrossRef]
10. Huberty, J.L.; Buman, M.P.; Leiferman, J.A.; Bushar, J.; Adams, M.A. Trajectories of objectively-measured physical activity and sedentary time over the course of pregnancy in women self-identified as inactive. *Prev. Med. Rep.* **2016**, *3*, 353–360. [CrossRef]
11. Clarke, P.E.; Gross, H. Women's behaviour, beliefs and information sources about physical exercise in pregnancy. *Midwifery* **2004**, *20*, 133–141. [CrossRef] [PubMed]

12. Fell, D.B.; Joseph, K.S.; Armson, B.A.; Dodds, L. The impact of pregnancy on physical activity level. *Matern. Child Health J.* **2009**, *13*, 597–603. [CrossRef] [PubMed]
13. Gjestland, K.; Bø, K.; Owe, K.M.; Eberhard-Gran, M. Do pregnant women follow exercise guidelines? Prevalence data among 3482 women, and prediction of low-back pain, pelvic girdle pain and depression. *Br. J. Sports Med.* **2013**, *47*, 515–520. [CrossRef] [PubMed]
14. Bai, G.; Raat, H.; Jaddoe, V.W.V.; Mautner, E.; Korfage, I.J. Trajectories and predictors of women's health-related quality of life during pregnancy: A large longitudinal cohort study. *PLoS ONE* **2018**, *13*, e0194999. [CrossRef] [PubMed]
15. Emmanuel, E.N.; Sun, J. Health related quality of life across the perinatal period among Australian Women. *J. Clin. Nurs.* **2014**, *23*, 1611–1619. [CrossRef] [PubMed]
16. Mourady, D.; Richa, S.; Karam, R.; Papazian, T.; Moussa, F.H.; El Osta, N.; Kesrouani, A.; Azouri, J.; Jabbour, H.; Hajj, A.; et al. Associations between quality of life, physical activity, worry, depression and insomnia: A cross-sectional designed study in healthy pregnant women. *PLoS ONE* **2017**, *12*, e0178181. [CrossRef] [PubMed]
17. Kolu, P.; Raitanen, J.; Luoto, R. Physical activity and health-related quality of life during pregnancy: A secondary analysis of a cluster-randomised trial. *Matern. Child Health J.* **2014**, *18*, 2098–2105. [CrossRef] [PubMed]
18. Bahadoran, P.; Mohamadirizi, S. Relationship between physical activity and quality of life in pregnant women. *Iran. J. Nurs. Midwifery Res.* **2015**, *20*, 282–286. [PubMed]
19. Arizabaleta, A.V.M.; Buitrago, L.O.; de Plata, A.C.A.; Escudero, M.M.; Ramírez-Vélez, R. Aerobic exercise during pregnancy improves health-related quality of life: A randomised trial. *J. Physiother.* **2010**, *56*, 253–258. [CrossRef]
20. Tendais, I.; Figueiredo, B.; Mota, J.; Conde, A. Physical activity, health-related quality of life and depression during pregnancy. *Cad. Saúde Pública* **2011**, *27*, 219–228. [CrossRef] [PubMed]
21. Mogos, M.F.; August, E.M.; Salinas-Miranda, A.A.; Sultan, D.H.; Salihu, H.M. A systematic review of quality of life measures in pregnant and postpartum mothers. *Appl. Res. Qual. Life* **2013**, *8*, 219–250. [CrossRef] [PubMed]
22. Vallim, A.L.; Osis, M.J.; Cecatti, J.G.; Baciuk, É.P.; Silveira, C.; Cavalcante, S.R. Water exercises and quality of life during pregnancy. *Reprod. Health* **2011**, *8*, 14. [CrossRef] [PubMed]
23. Nascimento, S.L.; Surita, F.G.; Parpinelli, M.; Siani, S.; Pinto e Silva, J.L. The effect of an antenatal physical exercise programme on maternal/perinatal outcomes and quality of life in overweight and obese pregnant women: A randomised clinical trial. *BJOG* **2011**, *118*, 1455–1463. [CrossRef] [PubMed]
24. Vachkova, E.; Jezek, S.; Mares, J.; Moravcova, M. The evaluation of the psychometric properties of a specific quality of life questionnaire for physiological pregnancy. *Health Qual. Life Outcomes* **2013**, *11*, 214. [CrossRef] [PubMed]
25. Sattler, M.C.; Jaunig, J.; Watson, E.D.; van Poppel, M.N.M.; Mokkink, L.B.; Terwee, C.B.; Dietz, P. Physical Activity Questionnaires for Pregnancy: A Systematic Review of Measurement Properties. *Sport Med.* **2018**, *48*, 2317–2346. [CrossRef] [PubMed]
26. Matsuzaki, M.; Haruna, M.; Nakayama, K.; Shiraishi, M.; Ota, E.; Murayama, R.; Murashima, S.; Yeo, S. Adapting the Pregnancy Physical Activity Questionnaire for Japanese Pregnant Women. *J. Obstet. Gynecol. Neonatal Nurs.* **2014**, *43*, 107–116. [CrossRef] [PubMed]
27. Matsuzaki, M.; Haruna, M.; Ota, E.; Yeo, S.A.; Murayama, R.; Murashima, S. Translation and cross-cultural adaptation of the pregnancy physical activity questionnaire (PPAQ) to Japanese. *Biosci. Trends* **2010**, *4*, 170–177. [PubMed]
28. Xiang, M.; Konishi, M.; Hu, H.; Takahashi, M.; Fan, W.; Nishimaki, M.; Ando, K.; Kim, H.K.; Tabata, H.; Arao, T.; et al. Reliability and Validity of a Chinese-Translated Version of a Pregnancy Physical Activity Questionnaire. *Matern. Child Health J.* **2016**, *20*, 1940–1947. [CrossRef] [PubMed]
29. Cirak, Y.; Yilmaz, G.D.; Demir, Y.P.; Dalkilinc, M.; Yaman, S. Pregnancy physical activity questionnaire (PPAQ): Reliability and validity of Turkish version. *J. Phys. Ther. Sci.* **2015**, *27*, 3703–3709. [CrossRef] [PubMed]
30. Ota, E.; Haruna, M.; Yanai, H.; Suzuki, M.; Anh, D.D.; Matsuzaki, M.; Tho le, H.; Ariyoshi, K.; Yeo, S.A.; Murashima, S. Reliability and validity of the Vietnamese version of the pregnancy physical activity questionnaire (PPAQ). *Southeast Asian J. Trop. Med. Public Health* **2008**, *39*, 562–570. [PubMed]

31. Silva, F.T.; Araujo Júnior, E.; Santana, E.F.M.; Lima, J.W.; Cecchino, G.N.; Silva Costa, F.D. Translation and cross-cultural adaptation of the Pregnancy Physical Activity Questionnaire (PPAQ) to the Brazilian population. *Ceska Gynekol.* **2015**, *80*, 290–298. [PubMed]
32. Krzepota, J.; Sadowska, D.; Sempolska, K.; Pelczar, M. Measuring physical activity during pregnancy—Cultural adaptation of the pregnancy physical activity questionnaire (PPAQ) and assessment of its reliability in Polish conditions. *Ann. Agric. Environ. Med.* **2017**, *24*, 640–643. [CrossRef] [PubMed]
33. Krzepota, J.; Sadowska, D. Pregnant and active—Suitability of the pregnancy physical activity questionnaire for measuring the physical activity of pregnant women in Poland. *Fam. Med. Prim. Care Rev.* **2017**, *19*, 29–33. [CrossRef]
34. Krzepota, J.; Sadowska, D. Kwestionariusz Aktywności Fizycznej Kobiet w Ciąży—Wersja polska (PPAQ-PL). *Med. Ogólna I Nauk. O Zdrowiu* **2017**, *23*, 100–106. [CrossRef]
35. Chasan-Taber, L.; Schmidt, M.D.; Roberts, D.E.; Hosmer, D.; Markenson, G.; Freedson, P.S. Development and validation of a Pregnancy Physical Activity Questionnaire. *Med. Sci. Sports Exerc.* **2004**, *36*, 1750–1760. [CrossRef]
36. Ainsworth, B.E.; Haskell, W.L.; Whitt, M.C.; Irwin, M.L.; Swartz, A.M.; Strath, S.J.; O'Brien, W.L.; Bassett, D.R., Jr.; Schmitz, K.H.; Emplaincourt, P.O.; et al. Compendium of physical activities: An update of activity codes and MET intensities. *Med. Sci. Sports Exerc.* **2000**, *32*, S498–S504. [CrossRef]
37. Wołowicka, L.; Jaracz, K. Polska wersja WHOQOL-WHOQOL 100 i WHOQOL BREF. In *Jakość Życia w Naukach Medycznych*; Wydawnictwo Akademii Medycznej: Poznań, Poland, 2001; pp. 259–281.
38. Ko, Y.L.; Chen, C.P.; Lin, P.C. Physical activities during pregnancy and type of delivery in nulliparae. *Eur. J. Sport Sci.* **2016**, *16*, 347–380. [CrossRef]
39. Evenson, K.R.; Wen, F. Prevalence and correlates of objectively measured physical activity and sedentary behavior among US pregnant women. *Prev. Med.* **2011**, *53*, 39–43. [CrossRef]
40. Santos, P.C.; Abreu, S.; Moreira, C.; Santos, R.; Ferreira, M.; Alves, O.; Moreira, P.; Mota, J. Physical activity patterns during pregnancy in a sample of Portuguese women: A longitudinal prospective study. *Iran. Red Crescent Med. J.* **2016**, *18*, e22455. [CrossRef]
41. Richardsen, K.R.; Mdala, I.; Berntsen, S.; Ommundsen, Y.; Martinsen, E.W.; Sletner, L.; Jenum, A.K. Objectively recorded physical activity in pregnancy and postpartum in a multi-ethnic cohort: Association with access to recreational areas in the neighbourhood. *Int. J. Behav. Nutr. Phys. Act.* **2016**, *13*, 78. [CrossRef]
42. Weir, Z.; Bush, J.; Robson, S.C.; McParlin, C.; Rankin, J.; Bell, R. Physical activity in pregnancy: A qualitative study of the beliefs of overweight and obese pregnant women. *BMC Pregnancy Childbirth* **2010**, *10*, 18. [CrossRef] [PubMed]
43. Guelfi, K.J.; Wang, C.; Dimmock, J.A.; Jackson, B.; Newnham, J.P.; Yang, H. A comparison of beliefs about exercise during pregnancy between Chinese and Australian pregnant women. *BMC Pregnancy Childbirth* **2015**, *15*, 345. [CrossRef] [PubMed]
44. Evenson, K.R.; Bradley, C.B. Beliefs about exercise and physical activity among pregnant women. *Patient Educ. Couns.* **2010**, *79*, 124–129. [CrossRef] [PubMed]
45. Hegaard, H.K.; Kjaergaard, H.; Damm, P.P.; Petersson, K.; Dykes, A.K. Experiences of physical activity during pregnancy in Danish nulliparous women with a physically active life before pregnancy. A qualitative study. *BMC Pregnancy Childbirth* **2010**, *10*, 33. [CrossRef] [PubMed]
46. Duncombe, D.; Wertheim, E.H.; Skouteris, H.; Paxton, S.J.; Kelly, L. Factors related to exercise over the course of pregnancy including women's beliefs about the safety of exercise during pregnancy. *Midwifery* **2009**, *25*, 430–438. [CrossRef] [PubMed]
47. Poudevigne, S.; Connor, P.J.O. A Review of Physical Activity Patterns in Pregnant Women and Their Relationship to Psychological Health. *Sport Med.* **2006**, *36*, 19–38. [CrossRef]
48. Krans, E.E.; Gearhart, J.G.; Dubbert, P.M.; Klar, P.M.; Miller, A.L.; Replogle, W.H. Pregnant women's beliefs and influences regarding exercise during pregnancy. *J. Miss. State Med. Assoc.* **2005**, *46*, 67–73. [PubMed]
49. Biernat, E.; Buchholtz, S. Poor health and contraindications—The most common barriers to physical activity in poles aged 50+. *Health Probl. Civiliz.* **2017**, *11*, 135–141. [CrossRef]
50. Bossola, M.; Murri, R.; Onder, G.; Turriziani, A.; Fantoni, M.; Padua, L. Physicians' knowledge of health-related quality of life and perception of its importance in daily clinical practice. *Health Qual. Life Outcomes* **2010**, *8*, 43. [CrossRef] [PubMed]

51. Blum, J.W.; Beaudoin, C.M.; Caton-Lemos, L. Physical activity patterns and maternal well-being in postpartum women. *Matern. Child Health J.* **2004**, *8*, 163–169. [CrossRef] [PubMed]
52. Costa, D.D.; Rippen, N.; Dritsa, M.; Ring, A. Self-reported leisure-time physical activity during pregnancy and relationship to psychological well-being. *J. Psychosom. Obstet. Gynecol.* **2003**, *24*, 111–119. [CrossRef]
53. Gustafsson, M.K.; Stafne, S.N.; Romundstad, P.R.; Mørkved, S.; Salvesen, K.Å.; Helvik, A. The effects of an exercise programme during pregnancy on health-related quality of life in pregnant women: A Norwegian randomised controlled trial. *BJOG* **2016**, *123*, 1152–1160. [CrossRef] [PubMed]
54. Harrison, A.L.; Taylor, N.F.; Shields, N.; Frawley, H.C. Attitudes, barriers and enablers to physical activity in pregnant women: A systematic review. *J. Physiother.* **2018**, *64*, 24–32. [CrossRef] [PubMed]
55. Muzigaba, M.; Kolbe-Alexander, T.L.; Wong, F. The perceived role and influencers of physical activity among pregnant women from low socioeconomic status communities in South Africa. *J. Phys. Act. Health* **2014**, *11*, 1276–1283. [CrossRef] [PubMed]
56. Ramírez-Vélez, R. Pregnancy and health-related quality of life: A cross sectional study. *Colomb. Med.* **2011**, *42*, 476–481.
57. Jansink, R.; Braspenning, J.; van der Weijden, T.; Elwyn, G.; Grol, R. Primary care nurses struggle with lifestyle counseling in diabetes care: A qualitative analysis. *BMC Fam. Pract.* **2010**, *11*, 41. [CrossRef]
58. Downs, D.S.; Chasan-Taber, L.; Evenson, K.R.; Leiferman, J.; Yeo, S. Physical activity and pregnancy: Past and present evidence and future recommendations. *Res. Q. Exerc. Sport* **2012**, *83*, 485–502. [CrossRef] [PubMed]
59. Evenson, K.R.; Barakat, R.; Brown, W.J.; Dargent-Molina, P.; Haruna, M.; Mikkelsen, E.M.; Mottola, M.F.; Owe, K.M.; Rousham, E.K.; Yeo, S. Guidelines for Physical Activity during Pregnancy: Comparisons From Around the World. *Am. J. Lifestyle Med.* **2014**, *8*, 102–121. [CrossRef] [PubMed]
60. U.S. Department of Health and Human Services. *Physical Activity Guidelines for Americans*, 2nd ed.; U.S. Department of Health and Human Services: Washington, DC, USA, 2018. Available online: https://health.gov/paguidelines/second-edition/pdf/Physical_Activity_Guidelines_2nd_edition.pdf (accessed on 30 November 2018).
61. Foxcroft, K.F.; Callaway, L.K.; Byrne, N.M.; Webster, J. Development and validation of a pregnancy symptoms inventory. *BMC Pregnancy Childbirth* **2013**, *13*, 3. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

Relative Density of Away from Home Food Establishments and Food Spend for 24,047 Households in England: A Cross-Sectional Study

Tarra L. Penney [1,*], Thomas Burgoine [1] and Pablo Monsivais [1,2]

[1] UKCRC Centre for Diet and Activity Research (CEDAR), MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Box 285 Institute of Metabolic Science, Cambridge Biomedical Campus, Cambridge CB2 0QQ, UK; tb464@medschl.cam.ac.uk

[2] Department of Nutrition and Exercise Physiology, Elson S. Floyd College of Medicine, Washington State University, Spokane, WA 99164, USA; p.monsivais@wsu.edu

* Correspondence: tarra.penney@mrc-epid.cam.ac.uk

Received: 11 September 2018; Accepted: 26 November 2018; Published: 11 December 2018

Abstract: Eating away from home is a risk factor for poor diet quality and obesity. With an ever-increasing proportion of household food spend directed toward eating out, the proliferation of these food establishments may contribute to their use, a potential precursor to less healthy food choices and low overall diet quality. However few studies are conducted at the national level and across a range of away from home food sources. The purpose of this study was to examine the association between the density of away from home food establishments (e.g., restaurants, fast food outlets and cafés) and household spend on away from home food within a nationally representative sample for England, UK. A cross-sectional analysis of data from Wave 1 of the UK Household Longitudinal Survey (n = 24,047 adults aged ≥19y) was conducted. Exposure was characterised as the density of away from home food establishments to all other food sources within 1 mile of the home, divided into quintiles (Q1 as lowest exposure and Q5 as highest exposure). The primary outcome included households with a high away from home equivalised monthly food spend (≥25% of total food spend). Logistic regression was used to estimate associations between away from home food establishment exposure and high away from home food spend. Away from home food establishment density was significantly associated with a greater odds of high monthly food spend (Q3: OR = 1.18, 95% CI = 1.07, 1.30; Q4: OR = 1.30, 95% CI = 1.18, 1.43; and Q5: OR = 1.52, 95% CI = 1.37, 1.68) with attenuation after controlling for known socioeconomic confounders (Q4: OR = 1.13, 95% CI = 1.02, 1.25; and Q5: OR = 1.16, 95% CI = 1.04, 1.30) compared to those least exposed (Q1). Those most exposed to away from home food establishments had a 16% greater odds of allocating more than 25% of household food spend on away from home food sources. This study provides one of the first analyses at the national level to examine the role of the local food environment in relation to household food spend, a potential precursor to diet quality and health.

Keywords: food availability; eating away from home; household food spending

---

1. Background

Unhealthy diet and obesity are considered global epidemics that present a significant challenge for public health and policy action [1]. The 2013 global burden of disease study ranked unhealthy diet and high body weight as the first and third contributors, respectively, to morbidity and mortality [2]. Unhealthy food and nutrient intake include a low intake of fruits and vegetables, whole grains, low fat milk, fibre, nuts and seeds and high intake of red and processed meat, sugar and sodium [3]. Consumption of unhealthy foods, nutrients and poor overall diet quality are often observed among those who eat frequently at retail food establishments such as sit down restaurants, fast food outlets and coffee shops or cafés [4–6]. Away from home food sources can be appealing as they tend to serve large energy-dense portions, creating a low calorie-to-cost ratio [7] that can contribute to a greater risk of obesity, higher body weight and weight gain over time among frequent users [8,9].

Recently public health efforts to reduce the impact of eating away from home on population level diet quality and risk of obesity have focused on addressing retail 'food swamps', a concept characterised by a disproportionally high density of unhealthy to healthy food sources within a geographic area [10]. While there is significant debate regarding what food outlets are healthy or unhealthy, proliferation of a range of retail food establishments themselves (i.e., the sheer density) is suggested to influence food choice independent of individual level factors via easy and convenient access, thereby making them more likely to be chosen as a food source [11]. The dual process model of food choice supports this view suggesting that choice is the result of both an intentional process that favours convenience, preferences, tastes [12] and sensitivity to price [11] and an automatic process, where a flood of sensory cues can encourage people to eat even when they may not require food [13].

Empirical studies examining the local food environment and health often include the implicit assumption that a direct and independent effect on behaviour exists. However the influence of the density of away from home food establishments on food choice might be more accurately theorised as one distal environmental exposure on a causal chain that includes a range of intermediate exposures and outcomes. In combination these factors along the exposure chain provide the structural conditions needed for the frequent utilisation of away from home food establishments, followed by exposure to the food provision within outlets and subsequent food choice (see Figure 1) [14]. It is then through repeating these food choices over time that a dietary pattern emerges that can ultimately influence health.

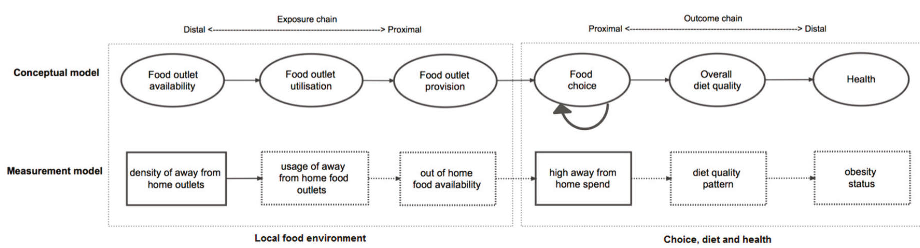


Figure 1. Conceptual model describing how the local food environment might relate to diet and health, highlighting measurement model for study hypothesis.

The evidence for the role of the local food environment on food choice and diet quality is plagued by conceptual and methodological challenges. The number of food outlets providing food for consumption away from home has been increasing [15] along with a rise in household spend on away from home sources [16]. However much of the research that examines the link between environmental exposures with diet and health remains equivocal. A proliferation of fast food outlets in one's neighbourhood has been shown to be associated with take away food intake and obesity [17–20]. One study from the US showed that for every standard deviation increase in fast-food exposure, the odds of consuming fast food near home increased 11%–61%, and the odds of a healthy diet decreased 3–17% (depending on adjustment) [21]. However, other studies from Canada [22] and the US have shown no association between density and proximity of exposure to fast food outlets and fast food purchasing after adjustment [23]. In addition to the mixed nature of the evidence base, several studies have been restricted to regional geographies [19,20], with a need to examine these relationships in nationally-representative samples [18–20].

The purpose of this cross-sectional study was to examine the association between exposure to away from home food outlets (e.g., restaurants, fast food outlets and cafés) and food spend on away from home sources within a nationally-representative sample for England.

2. Methods

2.1. Data Source and Participants

Data from the UK Household Longitudinal Study (UKHLS) were used [24]. UKHLS is a nationally-representative longitudinal panel survey of UK adults which began in 2009 and included over 40,000 households (57% household and 82% individual response rates) [25]. Details on the study and its sampling strategy are reported elsewhere [26]. In short, participants are surveyed annually to collect demographic, socioeconomic, behaviour and health related information using a computer assisted personal interview. The household questionnaire is answered by a reference person (see Supplementary Materials for sampling strategy, Figure S1) [25].

For this analysis the general population sample was restricted to English households with data reported by an adult (aged 19 years or older) reference person (n = 24,711) with complete household food spend for home and away from home sources (n = 24,047) (see Supplementary Materials for sample flow diagram, Figure S2.). The full household and analytic sample did not differ significantly on key demographic, socioeconomic, exposure or outcome variables; therefore only the analytic sample is presented. Ethical approval was not required for the analysis of secondary data presented here, but was obtained by UKHLS for data collection.

2.2. Exposure: Density of Away from Home Food Outlets to All Food Outlets

Data on the location of food establishments were obtained from Ordnance Survey's Points of Interest (POI), an administrative dataset for use by government and business [27]. The data are created and maintained by PointX, which sources the data from a list of over 150 suppliers, runs verification checks and classifies the features (see the reference for user guide) [28]. POI data has been found to be a viable alternative to accurate local council data in the UK [29]. Each food outlet is provided with geographic coordinates at a stated accuracy of within 1 m. The data is updated quarterly; this analysis used data from June 2014. The use of POI data for determining food environment exposure has been demonstrated in previous studies [30–32]. Away from home food establishments were comprised of three subcategories: 'sit-down restaurants', 'fast food outlets' and 'cafés'. Example sit-down restaurants include Bella Italia, Wetherspoons or Nando's; fast food outlets include McDonald's, Burger King or Kentucky Fried Chicken; and cafés include Café Nero, Starbucks and Costa. Food outlets primarily used as a food source for at home food preparation (e.g., supermarkets, convenience and green grocers etc.) were classified as 'Other' (see Supplementary Materials for food outlet frequency and classification, Table S1). Retail outlets in which food provision is not the primary service or food is not sold directly to the public were excluded (e.g., workplace cafeterias, cinemas and recreation facilities).

Using a geographic information system (ArcGIS 10, ESRI), relative away from home food establishment density was calculated for each household. Relative density is theorised as a spatial metric representing the intensity of exposure to features of the local food environment [33], and consistent with the focus on examining the role of density of away from home food establishments for this analysis [10]. Food outlet counts (away from home and other food outlet types) were made within a 1 mile Euclidean (straight line) radius buffer, centred on household addresses provided through UKHLS secure data access (see Figure 2). This distance is based upon previous work suggesting a behavioural relevance to food shopping among UK adults [34]. Relative density for each household was then calculated as the sum of the count of away from home establishments divided by the count of all food sources, divided into quintiles. Q1 represented those with the lowest proportion of away from home food outlets and Q5 representing those with the highest proportion.

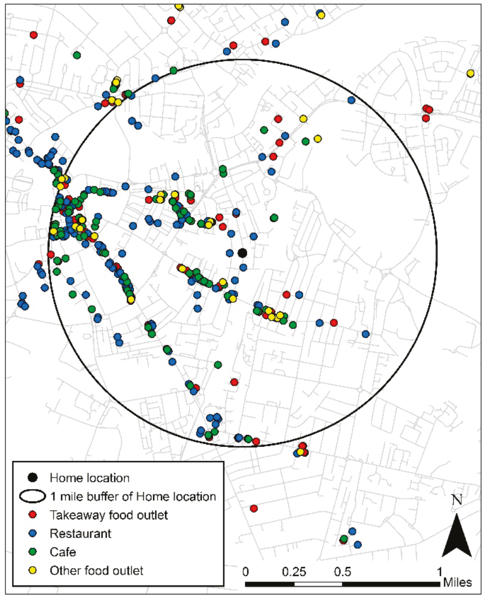


Figure 2. Depiction of how exposure was characterised as proportion of away from home food outlets to all food outlets around the home.

2.3. Outcome: Household with High Away from Home Food Spend

Household food expenditure was self-reported using two questions in the UKHLS survey: "About how much has your household spent in total on food and groceries in the last four weeks from a supermarket or other food shop or market? Please do not include alcohol" and "About how much have you and other members of your household spent in total on meals or snacks purchased outside the home in the last four weeks?" Monthly household food spend (£) was equivalised against household size using the OECD modified equivalence scale [35] and top-coded at the limit of the second highest decile (£500/month) that is, all values ≥£500 were recoded at that value. Sensitivity analyses informed the top coding through comparing mean home and away from home food spend to household spend estimates from the UK's Living Costs and Food survey for 2009 [36]. Total away-from-home food spend was divided by total household food spend then stratified into tertiles. The highest tertile (T3) was used to define households with a 'high' proportion of their total monthly household food spend on away from home food (approximately 25% or more). While no validation study of UKHLS data has been performed, previous work has found no significant differences between receipts and self-reported data on total food expenditures, expenditures at food stores, or eating out [37].

2.4. Additional Covariates

Demographic factors including age, categorised as young adult (18–35 years), middle-age (36–50 years) and older-age (>50 years), and sex of the household reference person were self-reported. Additionally, two indicators of socioeconomic status were patterned by both exposure and outcome: educational attainment categorised as 'None, GCSE or equivalent (≤11 years)', 'A-level (12–13 years)', 'Vocational qualifications (12–13 years)', 'Degree or higher (>13 years)' or 'Missing', and equivalised household income categorised as '£14,999 or below', '£15,000–£24,999', '£25,000–£34,999', '£35,000–£49,999', '£50,000 and above' or 'Missing'. Missing covariates were examined across all exposure variables, with no significant differences in percentages across exposure levels found. As a result, we included cases with missing covariate data in appropriate models to improve estimates and avoid case deletion (i.e., the missing indicator approach) [38].

2.5. Statistical Analyses

Descriptive statistics were used to summarise demographic, socioeconomic and household food spend variables, both overall and across away from home food outlet exposure levels. Study weights for Wave 1 cross-sectional analyses, prepared and provided by UKHLS, were used to account for participant non-response and clustered study design. Therefore, weighted mean percentages (with 95% CIs) are presented rather than raw frequencies.

Binary logistic regression was used to model high away from home food spend households by quintile of away from home food outlet density (Model 0). Crude models were adjusted for demographic variables (age and sex), proportion of restaurants, proportion of fast food outlets and proportion of cafés, total number of other food sources and total equivalised household food spend (Model 1). Lastly, Model 1 was additionally adjusted for socioeconomic variables (education and income) (Model 2). All statistical analyses were carried out in Stata version 14 (StataCorp LP., College Station, TX, USA) [39].

2.6. Sensitivity Analyses

Alternate multivariate models were used to examine sensitivity of our results to model specification. Additional covariate specifications were examined including with education as single indicator of socioeconomic status, ethnicity (White, all others) and rurality (Urban or Rural), and the removal of proportions of restaurants, fast food outlets and cafés.

3. Results

The weighted sample characteristics presented in Table 1 indicate that half the sample were female (50.6%), with representation across socioeconomic groups (30.8% having 'No or compulsory' education and 23.3% with 'Higher education'). Additionally, 40.7% of the sample had an equivalised income less than £14,999, with 5.7% having an income above £50,000. On average, away from home food outlet density was nearly 60% (of all outlets), composed of a mixture of 28.8% restaurants, 22.6% fast food outlets and 8.3% cafés. Average total food spend was £204 per month, with 17.6% of that spend being directed toward away from home food sources.

Table 2 shows selected weighted sample demographic and socioeconomic variables across level of away from home food outlet density. For the highest level of exposure to away from home food outlet density, a higher proportion of the sample are younger, belonged to a more socioeconomically-advantaged group (higher educational attainment and income), had higher total food related spend and allocated a greater proportion of household food spend toward away from home sources. Sex and mean age showed similar proportions of the sample across exposure levels.

Across quintile of away from home food outlet exposure, absolute density of food outlets range from just over 40% to more than 76% of total food outlet exposure (Figure 3a). The relative contribution of the three away from home food outlet types varied across quintiles, with an overall higher contribution of restaurants from lowest to highest quintile of exposure, where that outlet type provides the dominant source of away from home outlet exposure in the highest quintile (Figure 3b).

Table 1. Weighted sample characteristics for analytical sample (n = 24,047) from Wave 1 as percentage or mean and 95% CI where indicated.

	N (Unweighted)	Men	Women	All
		10,236	**13,811**	**24,047**
Demographic				
	Age in years [a]	50.5 (50.2–50.8)	50.6 (50.2–51.0)	50.3 (50.0–50.7)
	18–35 years	22.2	23.5	22.9
	36–50 years	29.7	30.8	30.3
	Over 50 years	48.1	45.7	46.8
Socioeconomic				
Education	None	14.7	20.5	17.9
	GCSE or equivalent (≤11 years)	28.2	32.88	30.8
	A-level (12–13 years)	19.2	13.7	16.2
	Vocational qualifications (12–13 years)	10.3	12.8	11.7
	Degree (>13 years)	27.4	20.0	23.3
	Missing	0.40	0.28	0.35
Equivalised household Income	<£14,999	44.4	36.2	40.7
	£15,000–£24,999	28.2	28.9	29.1
	£25,000–£34,999	13.9	16.3	15.0
	£35,000–£49,999	7.81	10.9	9.26
	£50,000+	4.49	7.23	5.74
	Missing	0.17	0.27	0.22
Density of AFH food establishments				
	Restaurant to all other	29.1 (28.6–29.5)	28.5 (28.1–29.0)	28.8 (28.4–29.1)
	Fast food to all other	22.3 (22.0–22.7)	22.7 (22.4–23.1)	22.6 (22.3–22.9)
	Cafe to all other	8.39 (8.16–8.61)	8.14 (7.93–8.36)	8.25 (8.07–8.44)
	Total AFH establishments to all other	59.8 (59.4–60.3)	59.5 (59.1–60.0)	59.6 (59.3–60.0)
Household food Spend				
	Total Equivalised food spend/mo. [a]	£209 (207–211)	£200 (198–202)	£204 (203–206)
	Percent AFH food spend/mo. [a]	19.4 (19.1–19.8)	16.1 (15.8–16.4)	17.6 (17.4–17.9)

[a] weighted mean (95% CI); AFH = Away from home.

Table 2. Weighted sample characteristics (n = 24,047) as column percentages (unless otherwise stated) by quintile of proportion of away from home food outlet density.

		Quintile of % Away from Home Food Outlet Exposure					
		Q1 = Lowest	**Q2**	**Q3**	**Q4**	**Q5 = Highest**	**All**
	N (unweighted)	5177	4785	4871	4697	4517	24,047
	Density of AFH establishments (min–max)	0–0.50	0.50–0.57	0.57–0.63	0.63–0.69	0.69–1	0–1
Demographic							
	Age in years [a]	52.1 (51.4–52.7)	50.2 (49.5–50.9)	50.2 (49.6–50.9)	50.2 (49.5–50.9)	49.8 (49.0–50.5)	50.5 (50.2–50.8)
	18–35 years	19.8	22.5	22.9	24.7	24.9	22.9
	36–50 years	29.6	32.1	31.2	29.4	29.4	30.3
	Over 50 years	50.7	45.5	45.9	45.9	45.7	46.8
	Sex (% Male)	44.1	45.2	44.5	46.3	47.3	45.5
Socioeconomic							
Education	None	19.4	19.6	19.6	17.0	14.1	17.9
	GCSE or equivalent (≤11 yrs)	33.8	31.8	31.2	30.9	25.9	30.8
	A-level (12–13 yrs)	16.2	16.8	16.2	15.7	16.4	16.2
	Vocational qualifications (12–13 yrs)	12.3	11.4	11.2	11.0	12.5	11.7
	Degree (>13 yrs)	18.1	20.3	21.8	25.3	31.1	23.3
Income	<£14,999	42.5	44.8	41.1	38.3	37.2	40.7
	£15,000–£24,999	30.2	28.5	29.7	29.3	27.6	29.1
	£25,000–£34,999	14.2	14.3	15.0	15.9	15.7	15.0
	£35,000–£49,999	8.28	8.30	8.80	10.2	10.6	9.26
	£50,000+	4.61	4.11	5.07	5.14	8.63	5.74
Household food spend							
	Total Equivalised food spend/mo. [a]	£207 (198–204)	£196 (193–199)	£201 (198–204)	£206 (202–209)	£217 (213–221)	£204 (203–206)
	Percent AFH food spend/mo. [a]	15.9 (15.4–16.5)	16.5 (15.9–17.7)	17.8 (17.2–18.4)	18.3 (17.7–18.9)	19.5 (18.8–20.1)	17.6 (17.4–17.8)

[a] weighted mean (95% Confidence Interval); AFH = Away from home.

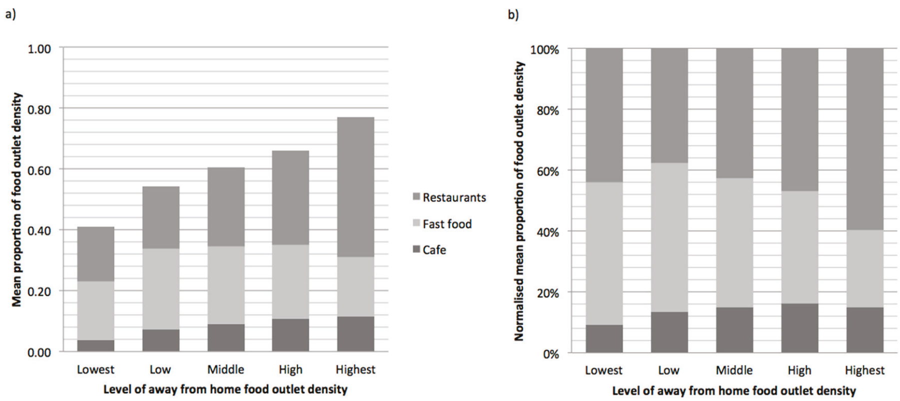


Figure 3. (**a**) Stacked weighted mean % of food outlet density by quintile of away from home exposure for each food outlet type; (**b**) Normalised contribution of weighted mean % food outlet density for type of food outlet by proportion of away from home exposure quintile.

Regression analyses (Table 3) showed that compared to the lowest level of away from home food outlet density (Q1) higher levels of density were associated with a greater odds of high monthly away from home food spend (Q3: OR = 1.18, 95% CI = 1.07, 1.30; Q4: OR = 1.30, 95% CI = 1.18, 1.43; and Q5: OR = 1.52, 95% CI = 1.37, 1.68) (Model 0).

Table 3. Odds ratios and 95% confidence intervals for high % (top tertile) of away from home food spend (N = 24,047) by quintile of proportion of away from home food outlets exposure.

Proportion of Food Outlet Density	Odds of High Monthly % Away from Home Food Spend [1]		
	Model 0	Model 1 [2]	Model 2 [3]
Away from home			
Q1	1.00 (-)	1.00 (-)	1.00 (-)
Q2	0.99 [0.90, 1.10]	0.97 [0.87, 1.07]	0.96 [0.86, 1.06]
Q3	1.18 *** [1.07, 1.30]	1.12 * [1.02, 1.24]	1.09 [0.99, 1.21]
Q4	1.30 *** [1.18, 1.43]	1.20 *** [1.08, 1.33]	1.13 * [1.02, 1.25]
Q5	1.52 *** [1.37, 1.68]	1.26 *** [1.13, 1.41]	1.16 ** [1.04, 1.30]

[1] % of away from home food spend was divided into tertiles, with the highest tertile being 'high' in % of away from home food spending; [2] Adjusted for age, sex, total number of food outlets for restaurant, fast food, cafe, other and equivalised total food spend; [3] Additionally adjusted for education and equivalised income. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

This relationship was attenuated after adjustment for known confounders with Q3 becoming non-significant (Q3: OR = 1.09, 95% CI = 0.99, 1.21) but with Q4 and Q5 remaining significantly associated with a greater odds of high away from home food spend (Q4: OR = 1.13, 95% CI = 1.02, 1.25 and Q5: OR = 1.16, 95% CI = 1.04, 1.30, respectively) (Model 2).

Sensitivity analyses with alternative model specifications were performed as described but results did not differ meaningfully from those described here.

4. Discussion

Using a nationally-representative sample of UK households, the purpose of this work was to contribute to our understanding to whether exposure to away from home food establishments in residential neighbourhoods is a precursor to population-level food choice. We observed that exposure to the greatest density of away from home food establishments, relative to all food sources, was associated with greater odds of households directing a high proportion of their monthly food spend on away from home food. Specifically, after adjustment, those in the most-exposed group had 16% higher odds of being high-spending households with respect to away from home food sources.

Our findings, although based on economic outcomes, provide evidence of an important but neglected theoretical link between the local food environment and health. Previous studies have observed associations between residentially-based estimates of fast food access and dietary and/or weight outcomes [18–20]. This work demonstrates an association with a more proximal outcome of food spending in particular and across a wider range of away from home food establishments. While our study had no direct measures of the health-related characteristics of food spend, other research has found that eating out-of-home is associated with lower diet quality and poor health outcomes [5,40].

The demonstration of a consistent relationship between the local food environment, diet and health remains somewhat elusive [41–43]. Previous research has suggested much of this heterogeneity may be due to methodological differences in exposure estimation, including the various decisions made surrounding the selection of spatial measures, geographic units, buffers around individual addresses [41,42] and the quality of dietary outcome measurement [42,43]. While the findings here should be interpreted within the context of a mixed evidence base that is yet unexplained, the findings do support previous work that has reported a significant relationship between exposure to away from home food establishments, diet and weight, and the amplification of inequalities [19,20,44]. However, a degree of caution may be warranted when using the observed association found here to *infer how a change* in the local food environment might influence diet and health. There is a paucity of studies available that examine the impact of a change in away from home food outlet exposure on diet or health, or address issues that could undermine causal inference including the role of neighbourhood self-selection (i.e., individual preferences drive selection of a neighbourhood that provides a local food environment that supports their preferences) [45].

In addition to structural factors, personal food tastes and preferences might play an important role in understanding differences in the relationship between exposure to away from home food sources and diet [44]. For example, consumers report preferences for fast food outlets that are convenient, easy to access and provide tasty foods, with availability of nutritious foods being the least important factor [46]. Additionally, individual preferences may be socioeconomically patterned, with more advantaged individuals possessing the material, psychosocial and time related resources needed to select food outlets and food items of their choosing regardless of what is easily accessible [47,48]. In our findings, higher income and more-educated individuals tended to have higher away from home food spend, suggesting a preference for the service and convenience of restaurants, cafes and fast food outlets among these population groups.

4.1. Policy Implications

As discussed, these research findings contribute to a growing scientific evidence base, which suggests that greater access to away from home food establishments contributes to unhealthy dietary behaviour, excess body weight and obesity [20,49,50] Although some questions remain, this growing literature has direct links to public health policy through informing 'healthy' neighbourhood design [51], which is increasingly understood by planners as a low-agency, population-level public health intervention [52]. For example, planners in English local government are actively encouraged to implement planning laws that limit growth in the fast food sector [53], focussing on areas of perceived need, where levels of obesity are currently high, or where existing access to fast food outlets is sufficient. Internationally, there are examples of similar practice [54] and anecdotal evidence to suggest that implementation of such policies is increasingly commonplace [55]. The effectiveness of many of these policies is yet to be determined.

4.2. Methodological Considerations and Limitations

For this study, food outlet exposure was estimated within residential neighbourhoods, defined as within 1 mile of each participant's home address. This home-based characterisation may overestimate some forms of outlet exposure, particularly if there are physical barriers in the environment not accounted for using the buffer method, and underestimate other forms of exposure, particularly for food outlets accessed from the workplace or while commuting between locations [20]. Additionally, the OS POI data on food outlet locations is not a routinely validated data set. Approximately 60% of the data is reported as 'ground truthed', however given the size of the data, it is not feasible to check each location. We used 2014 data with the assumption that secular food outlet change is relatively slow, and therefore unlikely to result in a significant difference in quintile of exposure between the year the food outlet data was collected and the year the survey data was collected, however this assumption is not based on a validation study for the UK. Additionally, the use of self-report data for estimates of household food spending is an important consideration. Although validation of self-report expenditure has been done in previous work [37], there is no UKLHS validation study for household food expenditure that we are aware of. Household food spend was benchmarked against Living Cost and Food data, however, while these adjusted data are likely appropriate for the type of analysis presented here, they may not provide perfectly accurate population level estimates of household food spend. Also, while we have adjusted for known confounders that were available to us and used routinely in previous research, the role of other unmeasured confounders cannot be ruled out, including car ownership or access to or use of public transportation.

Major strengths of this study include the use of a nationally-representative geographically diverse sample of UK households, the use of objectively measured food outlet data and coherence between the characterisation of the food environment and outcome. Using data from this national survey increases the generalizability of our findings to the UK population. Previous studies have typically used data from geographically-circumscribed samples from study cohorts. Although such samples can be representative of their underlying regional populations, they may be more limited in terms of generalizability to a national context. For the first time in the published literature, we exploited novel household food spend data within this national social and economic panel survey, including information on the amount spent purchasing foods specifically for consumption away from home.

5. Conclusions

Our findings suggest that those most exposed to away from home food establishments had a greater odds of allocating a high proportion of household food spend on away from home food sources. This study provides one of the first analyses at the national level to examine the role of the local food environment in relation to household food spend, a potential precursor to diet quality and health. Further research is required to better understand how and why different populations interact with their local food environment over time to inform the most effective policies to support healthy food choice while eating away from the home, diet and health.

Supplementary Materials: The following are available online at http://www.mdpi.com/1660-4601/15/12/2821/s1, Figure S1: Sample strategy for UKHLS, Figure S2: Analytic sample flow diagram, Figure S3: Food outlet classification and away from home food outlet groups for England, UK.

Author Contributions: T.L.P. designed the study with guidance from P.M., T.L.P. and T.B. prepared the data, T.L.P. conduced the data analysis and drafted the manuscript. All authors contributed to the interpretation of results and critically reviewed the manuscript. All authors read and approved the final manuscript.

Funding: The work was undertaken by the Centre for Diet and Activity Research (CEDAR), a UKCRC Public Health Research Centre of Excellence. Funding from the British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council, the National Institute for Health Research and the Wellcome Trust, under the auspices of the UK Clinical Research Collaboration, is gratefully acknowledged. T.L.P. additionally received support from the Cambridge Commonwealth Trust. P.M. also received support from the Health Equity Research Center at Washington State University.

Acknowledgments: Understanding Society is an initiative funded by the Economic and Social Research Council and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex and survey delivery by NatCen Social Research and Kantar Public. The research data are distributed by the UK Data Service.

Conflicts of Interest: The authors declare that they have no competing interests.

References

1. Geneva: World Health Organization. Global Action Plan for the Prevention And Control of Noncommunicable Diseases 2013–2020. 2013. Available online: http://apps.who.int/iris/bitstream/10665/94384/1/9789241506236_eng.pdf (accessed on 3 December 2018).
2. Newton, J.N.; Briggs, A.D.M.; Murray, C.J.L.; Dicker, D.; Foreman, K.J.; Wang, H.; Naghavi, M.; Forouzanfar, M.H.; Ohno, S.L.; Barber, R.M.; et al. Changes in health in England, with analysis by English regions and areas of deprivation, 1990–2013: A systematic analysis for the Global Burden of Disease Study 2013. *Lancet* **2015**, *386*, 2257–2274. [CrossRef]
3. Bates, B.; Lennox, A.; Prentice, A.; Bates, C.; Page, P.; Nicholson, S.; Swan, G. *National Diet and Nutrition Survey Results from Years 1, 2, 3 and 4 (combined) of the Rolling Programme (2008/2009–2011/2012)*; Food Standards Agency and Public Health England: London, England, 2014. Available online: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/310995/NDNS_Y1_to_4_UK_report.pdf (accessed on 12 February 2018).
4. Lachat, C.; Nago, E.; Verstraeten, R.; Roberfroid, D.; Van Camp, J.; Kolsteren, P. Eating out of home and its association with dietary intake: A systematic review of the evidence. *Obes. Rev.* **2012**, *13*, 329–346. [CrossRef] [PubMed]
5. Penney, T.L.; Jones, N.R.V.; Adams, J.; Maguire, E.R.; Burgoine, T.; Monsivais, P. Utilization of Away-From-Home Food Establishments, Dietary Approaches to Stop Hypertension Dietary Pattern, and Obesity. *Am. J. Prev. Med.* **2017**, *53*, e155–e163. [CrossRef] [PubMed]
6. Kant, A.K.; Whitley, M.I.; Graubard, B.I. Away from home meals: Associations with biomarkers of chronic disease and dietary intake in American adults, NHANES 2005-2010. *Int. J. Obes.* **2015**, *39*, 820–827. [CrossRef] [PubMed]
7. Jaworowska, A.M.; Blackham, T.; Long, R.; Taylor, C. Nutritional composition of takeaway food in the UK. *Nutr. Food Sci.* **2014**, *44*, 414–430. [CrossRef]
8. Nago, E.; Lachat, C.; Dossa, R.; Kolsteren, P. Association of out-of-home eating with anthropometric change: A systematic review of prospective studies. *Crit. Rev. Food Sci. Nutr.* **2014**, *54*, 1103–1116. [CrossRef]
9. Bezerra, I.N.; Curioni, C.; Sichieri, R. Association between eating out of home and body weight. *Nutr. Rev.* **2012**, *70*, 65–79. [CrossRef]
10. Luan, H.; Law, J.; Quick, M. Identifying food deserts and swamps based on relative healthy food access: A spatio-temporal Bayesian approach. *Int. J. Health Geogr.* **2015**, *14*, 1–11. [CrossRef]
11. Larson, N.; Story, M. A review of environmental influences on food choices. *Ann. Behav. Med.* **2009**, *38*, 56–73. [CrossRef]
12. Drewnowski, A. Taste Preferences and Food Intake. *Annu. Rev. Nutr.* **1997**, *17*, 237–253. [CrossRef]
13. Cohen, D.; Farley, T.A. Eating as an automatic behavior. *Prev. Chronic Dis.* **2008**, *5*, A23. [PubMed]
14. Sturm, R. Affordability and Obesity: Issues in the Multifunctionality of Agricultural/Food Systems. *J. Hunger. Environ. Nutr.* **2009**, *4*, 454–465. [CrossRef] [PubMed]
15. Wang, M.C.; Cubbin, C.; Ahn, D.; Winkleby, M.A. Changes in neighbourhood food store environment, food behaviour and body mass index, 1981–1990. *Public Health Nutr.* **2008**, *11*, 963–970. [CrossRef] [PubMed]
16. Office for National Statistics. Family Spending in the UK—Office for National Statistics. Available online: https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/expenditure/bulletins/familyspendingintheuk/financialyearendingmarch2016 (accessed on 21 September 2017).
17. Dubowitz, T.; Ghosh-Dastidar, M.; Eibner, C.; Slaughter, M.E.; Fernandes, M.; Whitsel, E.A.; Bird, C.E.; Jewell, A.; Margolis, K.L.; Li, W.; et al. The Women's Health Initiative: The Food Environment, Neighborhood Socioeconomic Status, BMI, and Blood Pressure. *Obesity* **2012**, *20*, 862–871. [CrossRef]
18. Dunn, R.A.; Sharkey, J.R.; Horel, S. The effect of fast-food availability on fast-food consumption and obesity among rural residents: An analysis by race/ethnicity. *Econ. Hum. Biol.* **2012**, *10*, 1–13. [CrossRef]
19. Kruger, D.J.; Greenberg, E.; Murphy, J.B.; DiFazio, L.A.; Youra, K.R. Local concentration of fast-food outlets is associated with poor nutrition and obesity. *Am. J. Health Promot.* **2013**, *28*, 340–343. [CrossRef]
20. Burgoine, T.; Forouhi, N.G.; Griffin, S.J.; Wareham, N.J.; Monsivais, P. Associations between exposure to takeaway food outlets, takeaway food consumption, and body weight in Cambridgeshire, UK: Population based, cross sectional study. *BMJ* **2014**, *348*, g1464. [CrossRef]

21. Moore, L.V.; Diez Roux, A.V.; Nettleton, J.A.; Jacobs, D.R.; Franco, M. Fast-Food Consumption, Diet Quality, and Neighborhood Exposure to Fast Food: The Multi-Ethnic Study of Atherosclerosis. *Am. J. Epidemiol.* **2009**, *170*, 29–36. [CrossRef]
22. Paquet, C.; Daniel, M.; Knäuper, B.; Gauvin, L.; Kestens, Y.; Dubé, L. Interactive effects of reward sensitivity and residential fast-food restaurant exposure on fast-food consumption. *Am. J. Clin. Nutr.* **2010**, *91*, 771–776. [CrossRef]
23. Thornton, L.E.; Bentley, R.J.; Kavanagh, A.M. Fast food purchasing and access to fast food restaurants: A multilevel analysis of VicLANES. *Int. J. Behav. Nutr. Phys. Act.* **2009**, *6*, 28. [CrossRef]
24. NatCen Social Research University of Essex, Institute for Social and Economic Research. *Understanding Society: Waves 1-7, 2009-2016 and Harmonised BHPS: Waves 1-18, 1991-2009: Secure Access*, 8th ed.; UK Data Service; University of Essex: Essex, UK, 2018. [CrossRef]
25. Knies, G. *Understanding Society The UK Household Longitudinal Study Waves 1-5, User Manual*; University of Essex: Essex, UK, 2015.
26. Lynn, P. Sample Design for Understanding Society. 2009. Available online: https://www.iser.essex.ac.uk/research/publications/working-papers/understanding-society/2009-01 (accessed on 3 December 2018).
27. Ordnance Survey. Points of Interest. 2015. Available online: https://www.ordnancesurvey.co.uk/business-and-government/products/points-of-interest.html (accessed on 12 February 2018).
28. Ordnance Survey Limited. Available online: https://www.ordnancesurvey.co.uk/business-and-government/help-and-support/products/points-of-interest.html (accessed on 6 November 2018).
29. Burgoine, T.; Harrison, F. Comparing the accuracy of two secondary food environment data sources in the UK across socio-economic and urban/rural divides. *Int. J. Health Geogr.* **2013**, *12*, 2. [CrossRef] [PubMed]
30. Lake, A.A.; Burgoine, T.; Stamp, E.; Grieve, R. The foodscape: Classification and field validation of secondary data sources across urban/rural and socio-economic classifications in England. *Int. J. Behav. Nutr. Phys. Act.* **2012**, *9*, 37. [CrossRef] [PubMed]
31. Powell, L.M.; Han, E.; Zenk, S.N.; Khan, T.; Quinn, C.M.; Gibbs, K.P.; Pugach, O.; Barker, D.C.; Resnick, E.A.; Myllyluoma, J.; et al. Field validation of secondary commercial data sources on the retail food outlet environment in the U.S. *Health Place.* **2011**, *17*, 1122–1131. [CrossRef] [PubMed]
32. Cummins, S.; Macintyre, S. Are secondary data sources on the neighbourhood food environment accurate? Case-study in Glasgow, UK. *Prev. Med.* **2009**, *49*, 527–528. [CrossRef]
33. Thornton, L.E.; Pearce, J.R.; Kavanagh, A.M. Using Geographic Information Systems (GIS) to assess the role of the built environment in influencing obesity: A glossary. *Int. J. Behav. Nutr. Phys. Act.* **2011**, *8*, 71. [CrossRef] [PubMed]
34. Smith, G.; Gidlow, C.; Davey, R.; Foster, C. What is my walking neighbourhood? A pilot study of English adults' definitions of their local walking neighbourhoods. *Int. J. Behav. Nutr. Phys. Act.* **2010**, *7*, 34. [CrossRef] [PubMed]
35. Anyaegbu, G. Using the OECD equivalence scale in taxes and benefits analysis. *Econ. Labour Mark. Rev.* **2010**, *4*, 49–54. [CrossRef]
36. Office for National Statistics. A Report on the Living Costs and Food Survey 2009. *Fam. Spend.* **2010**. Available online: http://www.ons.gov.uk/ons/rel/family-spending/family-spending/2010-edition/index.html (accessed on 17 September 2017).
37. Tang, W.; Aggarwal, A.; Liu, Z.; Acheson, M.; Rehm, C.D.; Moudon, A.V.; Drewnowski, A. Validating self-reported food expenditures against food store and eating-out receipts. *Eur. J. Clin. Nutr.* **2016**, *70*, 352–357. [CrossRef]
38. Kim, S.; Egerter, S.; Cubbin, C.; Takahashi, E.R.; Braveman, P. Potential Implications of Missing Income Data in Population-Based Surveys: An Example from a Postpartum Survey in California. *Public Health Rep.* **2007**, *122*, 753–763. [CrossRef] [PubMed]
39. StataCorp. *Stata Statistical Software*, StataCorp LP.: College Station, TX, USA, 2015.
40. Beydoun, M.A.; Powell, L.M.; Wang, Y. Reduced away-from-home food expenditure and better nutrition knowledge and belief can improve quality of dietary intake among US adults. *Public Health Nutr.* **2008**, *12*, 1. [CrossRef]
41. Cobb, L.K.; Appel, L.J.; Franco, M.; Jones-Smith, J.C.; Nur, A.; Anderson, C.A.M. The relationship of the local food environment with obesity: A systematic review of methods, study quality, and results. *Obesity* **2015**, *23*, 1331–1344. [CrossRef] [PubMed]

42. Caspi, C.E.; Sorensen, G.; Subramanian, S.V.; Kawachi, I. The local food environment and diet: A systematic review. *Health Place* **2012**, *18*, 1172–1187. [CrossRef] [PubMed]
43. Giskes, K.; van Lenthe, F.; Avendano-Pabon, M.; Brug, J. A systematic review of environmental factors and obesogenic dietary intakes among adults: Are we getting closer to understanding obesogenic environments? *Obes. Rev.* **2011**, *12*, e95–e106. [CrossRef] [PubMed]
44. Burgoine, T.; Forouhi, N.G.; Griffin, S.J.; Brage, S.; Wareham, N.J.; Monsivais, P. Does neighborhood fast-food outlet exposure amplify inequalities in diet and obesity? A cross sectional study. *Am. J. Clin. Nutr.* **2016**, 1540–1547. [CrossRef] [PubMed]
45. Boone-Heinonen, J.; Gordon-Larsen, P.; Guilkey, D.K.; Jacobs, D.R.; Popkin, B.M. Environment and physical activity dynamics: The role of residential self-selection. *Psychol. Sport Exerc.* **2011**, *12*, 54–60. [CrossRef] [PubMed]
46. Rydell, S.A.; Harnack, L.J.; Oakes, J.M.; Story, M.; Jeffery, R.W.; French, S.A. Why eat at fast-food restaurants: Reported reasons among frequent consumers. *J. Am. Diet. Assoc.* **2008**, *108*, 2066–2070. [CrossRef]
47. Nestle, M.; Wing, R.; Birch, L.; DiSogra, L.; Drewnowski, A.; Middleton, S.; Sigman-Grant, M.; Sobal, J.; Winston, M.; Economos, C. Behavioral and Social Influences on Food Choice. *Nutr. Rev.* **2009**, *56*, 50–64. [CrossRef]
48. Jabs, J.; Devine, C.M. Time scarcity and food choices: An overview. *Appetite* **2006**, *47*, 196–204. [CrossRef]
49. Fleischhacker, S.E.; Evenson, K.R.; Rodriguez, D.A.; Ammerman, A.S. A systematic review of fast food access studies. *Obes. Rev.* **2011**, *12*, e460–e471. [CrossRef]
50. Burgoine, T.; Sarkar, C.; Webster, C.J.; Monsivais, P. Examining the interaction of fast-food outlet exposure and income on diet and obesity: Evidence from 51,361 UK Biobank participants. *Int. J. Behav. Nutr. Phys. Act.* **2018**, *15*, 1–12. [CrossRef]
51. Public Health England. Encouraging Healthier "Out of Home" Food Provision—A Toolkit. 2017. Available online: https://www.gov.uk/government/publications/encouraging-healthier-out-of-home-food-provision (accessed on 19 November 2018).
52. Adams, J.; Mytton, O.; White, M.; Monsivais, P. Why Are Some Population Interventions for Diet and Obesity More Equitable and Effective Than Others? The Role of Individual Agency. *PLOS Med.* **2016**, *13*, e1001990. [CrossRef]
53. Local Government Association. Tipping the Scales: Case Studies on the Use of Planning Powers to Limit Hot Food Takeaway | Local Government Association. 2017. Available online: https://www.local.gov.uk/tipping-scales-case-studies-use-planning-powers-limit-hot-food-takeaway (accessed on 19 November 2018).
54. Los Angeles Fast Food Interim Control, Ordinance No. 180103 | Growing Food Connections. 2008. Available online: http://growingfoodconnections.org/gfc-policy/los-angeles-fast-food-interim-control-ordinance-no-180103/ (accessed on 19 November 2018).
55. Lake, A.A.; Henderson, E.J.; Townshend, T.G. Exploring planners' and public health practitioners' views on addressing obesity: Lessons from local government in England. *Cities Heal.* **2018**, *1*, 185–193. [CrossRef]

International Journal of
Environmental Research and Public Health

MDPI

Article

Concentrations, Distribution, Sources and Ecological Risk Assessment of Trace Elements in Soils from Wuhan, Central China

Ababo Workineh Tadesse [1,2,3], Tekleweini Gereslassie [1,2,3], Qiang Xu [4], Xiaojun Tang [5] and Jun Wang [1,2,*]

1 Key Laboratory of Aquatic Botany and Watershed Ecology, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China; abiyework@gmail.com (A.W.T.); tekle206@gmail.com (T.G.)
2 Sino-Africa Joint Research Center, Chinese Academy of Sciences, Wuhan 430074, China
3 Wuhan Botanical Garden, University of Chinese Academy of Sciences, Beijing 100049, China
4 Department of management engineering, Shangrao Vocational and Technical College, Shangrao 334109, China; xq3217@126.com
5 Wuhan Britain-China School, Wuhan Foreign Languages School, Wuhan, 430022 China; Tangxj2002@163.com
* Correspondence: wangjun@wbgcas.cn; Tel.: +86-027-8770-0859

Received: 20 November 2018; Accepted: 11 December 2018; Published: 14 December 2018

Abstract: This study aimed to determine the concentration levels, potential sources and ecological risks of eleven trace elements, namely Cr, Fe, Co, Ni, Cu, As, Sb, Cd, Zn, Hg and Pb, in the soil from Huangpi district, Wuhan, Central China. Soil samples were collected from eighteen sites at soil depths of 1–10 and 10–20 cm and analyzed using Inductively Coupled Plasma-Mass Spectrometer ICP-MS (Thermo X SERIES 2, Scientific and Innovative Technology Co. Ltd., Beijing, China). The recorded mean concentration of the elements were in a decreasing order of Fe > Co > Cr > Ni > Pb > Cu > As > Cd > Sb > Zn > Hg. The mean concentration of trace elements, soil pH and total organic carbon (TOC) were higher at a soil depth of 1–10 cm. The obtained mean concentration of Cr, Co, As, Cd, Ni, Cu, Hg and Pb were above the soil background values of Wuhan and Hubei Province. The mean concentration values of Co, Ni and Cd, exceeded the recommended FAO (Food and Agriculture Organization)/ISRIC (International Soil Reference and Information Centre) (2004) and WHO/FAO (2001) values. Pearson's correlation analysis illustrated that there was a strong and significant correlation between trace elements, whereas, a weak positive and negative correlation between elements and soil properties (pH and TOC). The principal component analysis (PCA) and cluster analysis (CA) result indicated that the concentration of trace elements in Huangpi soil were originated from anthropogenic sources. Potential ecological risk index (RI) of this study revealed that there is a high ecological risk of trace elements in the soil. Enrichment factor (EF) and geo-accumulation index (I_{geo}) of trace elements for this study indicated that the study area is strongly contaminated with Cd and Co. Generally, the finding of this research showed that Huangpi soil is contaminated.

Keywords: trace elements; concentration; microwave digester; Wuhan; soil contamination; ecological risk

1. Introduction

Trace elements are identified as a dangerous group of environmental pollutants, due to their persistence, non-degradability and toxicity to living organisms [1,2]. Unlike many organic pollutants, which eventually degrade to carbon dioxide and water, trace elements tend to accumulate in the environment, especially in soil and sediments [3]. Trace elements have specific gravity greater than 5 g/cm^3 and they include elements such as Cd, Zn, Pb, Hg, Sb, Cr, Co, As, Ni Cu, Mo, and Mg [4,5].

Trace elements like Cd, Cr, Pb, Hg, Ni and As have been listed as the most dangerous elements and priority control pollutants by the United States Environmental Protection Agency (USEPA) [6–8]. Trace elements can originate from natural sources (e.g., parent materials, weathering of rocks, volcanic eruptions and soil erosion) and anthropogenic sources (e.g., industrialization, urbanization, vehicular emissions, mining activities, smelting, burning of fossil fuels and agricultural inputs such as fertilizers, pesticides, herbicides and fungicides [9,10]. However, currently anthropogenic activities significantly accelerate the accumulation of trace elements in the environment [11,12].

Soil is one of the environmental compartments which plays an important role in plant growth, development and other ecosystem services [8,13]. On the other hand, it also serves as a sink for different environmental pollutants like toxic trace elements [14–16]. Several studies from different countries point out the toxic effects of trace elements on soil, plants and animals [1,3,12]. Excessive accumulation of trace elements in soil can reduce soil microorganism levels which results in soil quality degradation [16,17], and reduction of agricultural productivity [3,15,18]. It has been reported that foodstuffs and domestic feeds in Asian countries are severely contaminated with trace metals [19]. According to China national census of pollution report, more than 1.5 million sites in China have been exposed to toxic trace elements, 20 million hectares of agricultural lands have been polluted [20] and over 12 million tons of grains are contaminated by toxic elements per year [7]. About 1.5×10^5 km^2 of cultivated lands is polluted by Cd [21]. In addition to the above reports, a survey conducted in 2014 indicated that 16.1% of the sampled lands were contaminated by Hg, As and Pb [6].

Plants not only absorb essential nutrients, they also absorb toxic elements from the environmental compartments [22], which results in a bio-accumulation effect [14,23]. Consequently, consumption of contaminated plants and animals can cause health problems [24]. Particularly, Cd is amongst trace metals that increased international concern due its carcinogenic effect [21]. The adverse effects of toxic trace elements on human beings have been reported from different countries of the world [24,25]. For example: As causes dermal lesions, skin cancer, peripheral vascular diseases and peripheral neuropathy [26,27], Cd can result in kidney dysfunction, hypertension, lung cancer, bone fractures, prostatic hyperplasia and adenocarcinomas, Cu can cause Alzheimer's, prion disease [28,29], and Pb can affect the endocrine system, immune system, skeletal, circulatory system and nervous system [30].

Due to the rapid economic growth in China, there are heterogeneous anthropogenic activities which increase the accumulation of trace elements in the environment [31]. Agricultural inputs (fertilizers, herbicides, pesticides and fungicides), the use of wastewater for irrigation, urbanization, industrialization and construction of road networks are the major anthropogenic activities in Wuhan, Hubei Province [32–34]. Huangpi is one of the districts in Wuhan, in which the above anthropogenic activities are commonly observed [20]. However, there is lack of data and research works on the status of trace elements in Huangpi district. Previously conducted studies in Wuhan were focused only on few trace elements such as Cd, Cr, Pb Ni, As, Cu and Hg. Therefore, it's important to investigate the status of the above listed and other additional trace elements like Co, Sb, Fe, Zn in the soil of Huangpi district. Thus, the objectives of this study were: (1) to determine the concentration level of trace elements, (2) to identify the potential sources of trace elements, (3) to evaluate pollution status and ecological risks of trace elements in the soil and (4) to determine the effect of soil pH and total organic matter on the concentration of trace elements.

2. Materials and Methods

2.1. Study Area

Soil samples were collected from Huangpi district, Wuhan, Central China. Huangpi is one of thirteen districts of Wuhan, located on northern outskirts of Wuhan between 30°52′30″ N and 114°22′30″ E (Figure 1). The area of the district is about 2261 km^2, with a population of 1,107,565 [31]. From the total area of the land 56.12%, 18.29%, 0.22%, 7.09%, 0.21%, 23% and 4.08% are covered by cultivated land, forest land, grassland, settlements and industrial sites, land for transport, water area and unused land, respectively [35].

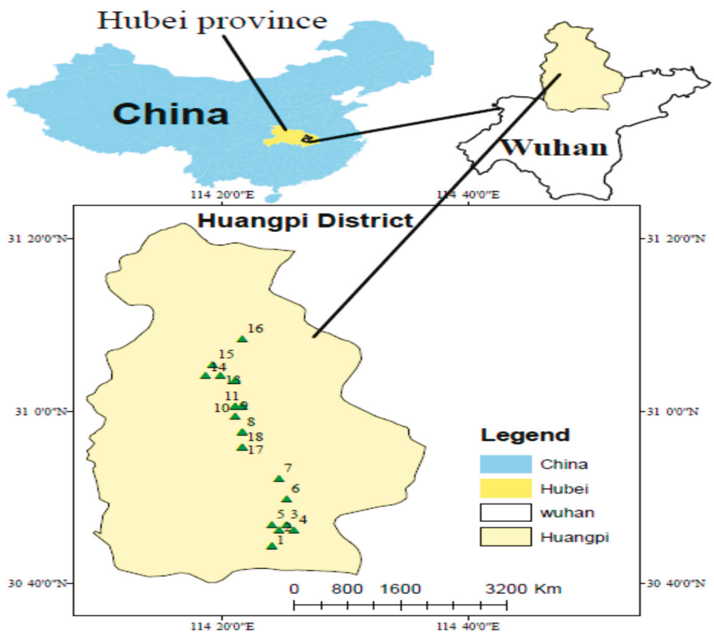


Figure 1. Study area. (The number (1–18) represents sampling sites: S1 (Tangjiawan), S2 (Fengdouhu) S3 (Erpaiqu), S4 (Changdi), S5 (Zhujiashan), S6 (Tujiadun), S7 (Zhulinyuan). S8 (Zhoujiawan), S9 (Lishuwan), S10 (Xinyang), S11 (Leqianwan), S12 (Bomogang), S13 (Hanjiafan), S14 (Wanjiatian), S15 (Hongguanshanxiawan), S16 (Dujiatian), S17 (Tianjiaxiaowan), S18 (Tianjiaxiaowan).

2.2. Sample Collection and Pretreatment

Soil samples were collected from eighteen sampling sites of different land use types (barren land, farmland, paddy field, plastic greenhouse) at soil depths of 1–10 and 10–20 cm. The sample sites were located using a global positioning system (GPS). For each site, three replicates were taken to make up a composite soil. The collected samples were packed in polyethylene containers, labeled and transported to the laboratory for analysis. All samples were dried using a benchtop lab vacuum freeze dryer (Xinzhi Biotechnology Co., Ltd., Ningbo, China) at −40 °C for 24 h, the samples were ground, sieved through 0.6 mm mesh nylon sieve and stored in a refrigerator at −20 °C for further analysis [12,36].

2.3. Sample Extraction and Analysis

About 0.1–0.15 g of soil samples were measured using a digital analytical balance (METTLER TOLEDO, Columbus, OH, USA), placed into Teflon vessels, digested with 4 mL of nitric acid (HNO_3, 63%), 2 mL of hydrogen fluoride (HF $\geq$ 40%) and perchloric acid ($HClO_4$ 70–72%) using a microwave digester (ETHOS ONE, Milestone, Leutkirch im Allgau, Germany). The digester was operated for 2 h at a controlled pressure, temperature and output power according to [37]. The digested samples were heated on the heating plate at a temperature of 135 °C for 2 h and cooled to room temperature. Then the final volume was topped up to 50 mL using double distilled water and filtered using 0.22 um membrane filter paper for analysis [22,36] and the filtered samples were stored in plastic bottles at −4 °C to minimize volatilization and biodegradation [38]. Finally, the samples were analyzed using ICP-MS (Thermo XSERIES2, Beijing, Scientific and Innovative Technology Co. Ltd., Beijing, China). In addition, selected soil properties (soil pH and TOC) were measured. Soil pH was measured according to [22], using a pH meter (METTLER TOLEDO) and the total organic carbon (TOC) of the soil was measured by [39] method using a TOC analyzer (Elementar GmbH, Langenselbold, Germany).

2.4. Standard Preparations and Calibration Curves

All chemicals and reagents used were analytical grades with purity of 99%. Five working solutions with concentrations of 0.1, 0.5, 1, 5, 10 and 20 µg/L were prepared from the stock solutions of 1000 µg/mL using double distilled water with 10% nitric acid (HNO_3, 63%). The working solutions were analyzed before running samples to check the accuracy and reliability of the instrument. Calibration curves produced for all elements indicated that the obtained curves had r^2 values greater than 0.994.

2.5. Statistical Analysis

Statistical Package for Social Sciences (SPSS) version 20 (IBM, Armonk, NY, USA) and Microsoft Offices Excel 2013 (Microsoft Corporation, Albuquerque, NM, United States) were used to analyze the data. Pearson's correlation coefficient was used to analyze the relationships between elements, soil pH and TOC. The potential sources of trace elements were identified using factor analysis (PCA and cluster analyses).

2.6. Quality Analysis and Quality Control

To ensure the quality of the experiment all reagents and chemicals used were analytical-reagent grades. All plastic containers were soaked with 10% HNO_3 (63%) for 12 h and washed three times with double distilled water to remove other contaminants, dried in oven at 60–65 °C for 24 h. The relative standard deviations (RSD) of three consecutive measurements of the standard solutions were used to determine the precision of ICP-MS. The obtained percentage of relative standard deviation (% RSD) was less than 10% indicating a good precision of the instrument. The recovery value of elements ranged between (90–104%).

2.7. Methods of Evaluating Contamination Level and Ecological Risk of Trace Elements

2.7.1. Potential Ecological Risk Index (RI)

Potential ecological risk index helps to evaluate the pollution of trace elements in the soil [40]. It is the sum of the ecological risk factor of a single element in the sample [41]. The following formulas given by [41] were used to calculate potential ecological risk indices of the elements.

$$C_f^i = C_S^i / C_R^i \quad (1)$$

$$Cd = \Sigma_f^i C_f^i \quad (2)$$

$$E_f^i = C_f^i * T_f^i \quad (3)$$

$$RI = \Sigma_f^i / E_f^i \quad (4)$$

where $C^i{}_f$ is the contamination factor for trace elements; $C^i{}_S$ is the measured concentrations value of elements in the soil; $C^i{}_R$ is the background reference values of trace elements in the soil; C_d is a degree of contamination; $E^i{}_f$ is the ecological risk index of a single element; $T^i{}_f$ is the toxicity coefficient of an element and *RI* represents the total potential risk index of elements [40] (Table 1).

Due to the absence of soil background reference values of trace elements in Huangpi district, Wuhan soil background values were used as a reference: Cr (90), Ni (40), Cu (35), Zn (100), As (15), Cd (0.20), Hg (0.15), Pb (35) mg/kg [42] and Hubei Province soil background values were used for

Co (15.4) [32], Sb (1.65 mg/kg) [43] and Fe (29,400 mg/kg) [40]. $T^i{}_f$ values given by [41] were Cr = 2, As = 10, Cd = 30, Ni = Cu = Pb =5, Zn =1, Hg = 40 and Co = 2, Fe = 0 and Sb = 15.

Table 1. The standard given for the contamination factor, degree of contamination, ecological risk factor and risk index of elements.

$C^i{}_f$	Contamination Level	C_d Class	Degree Contamination Level	$E^i{}_f$	Pollution Degree	RI	Risk Degree
$C^i{}_f < 1$	Low contamination factor	$C_d < 8$	Low	$E^i{}_f < 40$	Low risk	$RI < 150$	Low ecological risk
$1 \leq C^i{}_f < 3$	Moderate contamination	$8 \leq C_d < 16$	Moderate	$40 \leq E^i{}_f \leq 80$	Moderate risk	$150 \leq RI < 300$	Moderate ecological risk
$3 \leq C^i{}_f < 6$	Considerable contamination factor	$16 \leq C_d < 32$	Considerable	$80 \leq E^i{}_f < 160$	Considerable risk	$300 \leq RI < 600$	Considerable ecological risk
$C^i{}_f \geq 6$	Very high contamination	$C_d \geq 32$	Very high	$160 \leq E^i{}_f < 320$	high risk	$RI > 600$	Very strong
				$E^i{}_f \geq 320$	Extremely high		

Source: [41].

2.7.2. Enrichment Factor (EF)

Enrichment factor (*EF*) was evaluated to determine the degree of anthropogenic factors on trace elements accumulation in the soil. It reflects the disturbance degree of human activities on the natural environment [44]. *EF* is calculated as follows:

$$EF = \frac{\left(\frac{C_i}{C_{ref}}\right) sample}{\left(\frac{B_i}{B_{ref}}\right) background} \tag{5}$$

where C_i is the concentration of trace elements in the sample; C_{ref} is the concentration of reference element in the sample; B_i is the background value of interest element and B_{ref} is the background value of reference elements in the study area. Fe was selected as the reference element for this study because Fe is a major sorbent and it is a quasi-conservative tracer of natural elements in fluvial and coastal sediments [12]. Wuhan soil background was used as the reference. The following EF class were given for elements; $EF < 2$, $2 \leq EF < 5$, $5 \leq EF < 20$, $20 \leq EF < 40$ and $EF \geq 40$ indicated low, moderate, significant, very high and extremely high enrichment factor, respectively.

2.7.3. Geo-accumulation Index (I_{geo})

The geo-accumulation indexes (I_{geo}) of elements were calculated to identify the degree of contamination and to compare it with the pre-industrial level [45]. This method classifies the pollution level of elements in terms of seven enrichment classes. It is calculated using the following formula:

$$I_{geo} = Log_2 \frac{C_n}{1.5\ B_n} \tag{6}$$

where I_{geo} is the geo-accumulation index, C_n is the measured concentration of an element (mg/kg) in the sample, B_n is the geochemical background value of an element (mg/kg) and 1.5 is the factor used for lithological variations of elements.

Seven classes and contamination intensity are given for I_{geo}: $I_{geo} \leq 0$ (uncontaminated), $0 < I_{geo} \leq 1$ (uncontaminated to moderately contaminated), $1 < I_{geo} \leq 2$ (moderately contaminated), $2 < I_{geo} \leq 3$ (moderately to strongly contaminated), $3 < I_{geo} \leq 4$ (strongly contaminated), $4 < I_{geo} \leq 5$ (strongly to extremely contaminated) and $I_{geo} > 5$ (extremely contaminated) [45].

3. Result and Discussion

3.1. Spatial Distribution of Selected Soil Properties (pH and TOC)

It has been reported that soil properties such as soil pH and total organic carbon (TOC) are the most important factors that influence cation mobility and regulate the solubility of trace elements in the soil. The obtained mean result of soil pH and TOC for this study were presented in Table 2. As illustrated, the mean soil pH results ranged from 4.20–6.87 with a mean value of 5.71. The comparison between individual samples (Table 3) indicated the soil pH ranged from 4.12–8.07 with an average of 5.87 at the soil depth of 1–10 cm and 4.28–6.52 with an average of 5.55 at the soil depth of 10–20 cm. The highest soil pH value (8.07) was recorded in the greenhouse soil from Changdi site at a soil depth of 1–10 cm. The main reason for this might be connected with relatively low precipitation amount and less leaching of base-forming cation in the green house. The mean comparison between the two depths indicated that the soil pH at 1–10 cm (5.87) is higher than the soil depth of 10–20 cm (5.55) (Figure 2). The result is in line with the findings of [46,47], whereas the result is different from the findings of [48]. According to the classification of soil pH grade, the pH value of (<5) indicated slightly acidic, (5–6.5) mildly acidic, (6.5–7.5) neutral, (7.5–8.5) mildly alkaline and (>8.5) indicated strongly alkaline [49]. According to these classifications, Huangpi soil is classified as slightly acidic, mildly acidic, neutral and mildly alkaline.

The average total organic carbon of soil ranged from 0.65–2.41 with the mean value of 1.71. The highest mean value of TOC (2.41) was recorded in Dujiatian site (farmland), whereas the lowest TOC (0.65) was recorded from barren land at the Zhoujiawan site. TOC ranged from 0.63–2.60 with a mean value of 1.74 at soil depth of 1–10 cm and 0.66–2.71 with a mean value of 1.71 at the soil depth of 10–20 cm. The highest mean was recorded at the soil depth of 1–10 cm (Figure 2). A similar result was reported by [47,50]. Numerous studies indicated that TOC of soil decreases with soil depth [50,51]. The main reason for this might be biological activity in the top layer of soil [51].

3.2. Concentrations and Distribution of Trace Elements

Eleven trace elements in Huangpi district soil were studied. The results of descriptive statistics; mean, maximum, minimum, standard deviation and Skewness were presented in Table 2 and individual results of all samples and mean for both depths were presented in Table 3 and Figure 2, respectively. The results indicated that the concentrations of the elements showed variation between samples and within samples at different soil depths. All elements except As, Sb, Hg and Cd were detected in all samples at soil depths of 1–10 and 10–20 cm with detection frequency of 100% for Cr, Fe, Co, Ni, Cu, Zn and Pb; and 72%, 89%, 78% and 42% for As, Sb, Cd, Hg respectively. The mean concentration of trace elements at soil depth of 1–10 cm were in decreasing order of Fe > Co > Cr > Ni > Pb > Cu > Cd > As > Zn > Sb > Hg, whereas, Fe > Co > Cr > Ni> Pb > Cu > As > Cd > Zn > Sb > Hg at soil depth 10–20 cm (Figure 2).The highest mean concentration values for all elements were recorded at a soil depth of 1–10 cm. The results of this study was in line with those of [52–54]. According to Camobreco [55] the highest accumulation of trace elements in the surface layer of the soil might be due to a high sorption capacity of trace elements which results from soil chemical reactions between soil solid phases, including silicate clays, hydroxides and oxides of elements. Another study by Rahaman [53] also indicated that trace elements were found abundantly in the surface layer and the value decreased with an increase in soil depth with few exceptions. Converse to the results obtained in this study, the report by [56] from Kenya, indicated that subsurface soil accumulates high concentration of trace elements than surface soil due to soil leaching.

Table 2. Descriptive statistics of trace elements (mg/kg) and selected soil properties (pH and TOC) of soil from Huangpi district and background value of Wuhan, Hubei and China.

Elements	Min	Max	Mean	SD	Skewness	(a)	(b)	(c)	(d)	(e)	(f)
Cr	66.56	321.73	140.1	58.84	1.795	90	86	200	250	-	1000
Fe	13,583.04	55,398.01	27,304.9	10,705.1	1.148	-	29,400	29,400	-	-	-
Co	7,244.46	5,4621.91	22,656.94	10,317.8	1.578	-	15.4	40	50	50	-
Ni	51.18	210.63	117.8	50.93	0.701	40	37.3	50	100	100	500
Cu	26.09	139.98	60.73	30.06	1.197	35	30.7	100	100	100	100
As	ND	47.58	15.58	17.68	0.826	15	12.3	30	-	-	75
Sb	ND	1.54	0.58	0.42	0.662	-	1.65	10	-	-	-
Cd	0.07	77.62	15.44	23.84	1.69	0.2	0.172	0.5	5	3	0.7
Zn	1.53	4.81	3.32	0.94	−0.544	100	83.6	250	500	300	300
Hg	ND	1.13	0.15	0.26	3.458	0.15	0.08	0.7	-	2	1
Pb	38.81	117.9	74.16	20.22	0.395	35	26.7	80	150	50	200
pH	4.2	6.87	5.71	0.73	−0.543	-	6.5	-	-	-	-
TOC	0.65	2.41	1.72	0.47	−0.815	-	-	-	-	-	-

ND: Not detected; Max = Maximum; Min=minimum; SD = Standard deviation. (a), Wuhan soil background value [42], (b), Hubei province soil background [37], (c), china soil background value [27], (d), Food and Agriculture Organization (FAO)/International Soil Reference and Information Centre (ISRIC) (2004) [57] (e), World Health Organization (WHO)/Food and Agricultural Organization (FAO) (2001) [12], (f), United States and Environmental Protection Agency (USEPA) (1983) [58].

Table 3. The concentration of trace elements (mg/kg), soil pH and TOC for each sample at soil depths of (1–10 and 10–20 cm).

S	Cr	Fe	Co	Ni	Cu	As	Sb	Cd	Zn	Hg	Pb	PH	TOC
Depth of 1–10 cm													
1	94.20	14,491.1	10,133.9	73.64	39	ND	0.58	ND	2.85	0.15	53.42	5.21	1.59
2	169.20	28,732.1	22,910.7	163.3	109	31.78	1.09	34.61	4.35	0.36	70.71	5.91	1.75
3	163.13	28,633.9	23,232.1	160.8	61	22.11	0.63	13.82	4.38	0.51	77.95	7.86	1.71
4	251.54	40,557.7	33,375	293.75	123	42.93	1.75	75.13	5.65	0.19	106.44	8.07	2.6
5	266.25	46,817.3	36,740.4	253.08	138	63.83	1.95	124.59	6.57	0.58	151.06	5.39	2.14
6	112.69	20,019.2	16,653.9	97.6	62	ND	0.2	ND	1.97	0.19	61.88	5.73	1.51
7	224.50	39,650	31,433.3	181	165	60.82	1.78	108.46	5.4	0.31	141.33	6.29	1.03
8	137.77	23,946.4	35,053.6	120.27	33	43.64	0.65	28.45	4.28	ND	75.77	6.57	0.63
9	124.04	21,750	22,846.2	96.92	93	1.01	0.7	0.71	4.13	ND	94.17	6.14	2.58
10	130.17	20,175	20,925	120.08	53	3.24	0.52	1.67	3.67	0.08	106.75	6.58	2.31
11	138.56	28,903.9	24,884.6	100.19	43	10.02	0.52	5.2	2.58	ND	74.23	5.86	1.14
12	149.02	29,151.8	29,098.2	122.86	61	18.28	0.71	12.89	4.03	0.08	101.07	4.79	2.23
13	88.51	15,875	10,846.2	57.9	52	ND	0.04	ND	2.75	ND	65.77	6.48	2.25
14	172.50	29,033.3	28,258.3	125.42	60	ND	0.71	ND	3.47	ND	83.05	5.76	1.71
15	110.36	23,071.4	15,116.1	70.31	36	1.3	ND	ND	2.65	ND	65.43	4.6	1.46
16	328.66	51,901.8	55,580.4	203.75	131	23.03	0.13	3.08	3.73	1.41	57.6	5.05	1.4
17	138.17	20,908.3	25,366.7	109.33	47	5.52	0.87	4.78	4.93	0.33	87.17	4.12	1.55
18	107.05	16,642.9	19,991.1	81	35	ND	0.56	ND	3.48	ND	74.8	5.33	1.68
Depth of 10–20 cm													
1	38.93	12,675	4355	28.71	13	0.69	0.22	0.15	1.13	ND	24.2	6.01	2
2	165.17	46,416.7	23,508.3	207.67	69	29.74	1.03	30.73	3.84	ND	77.34	5.89	1.85
3	95	26,758.3	13,491.7	87.08	51	15.43	0.36	5.53	2.6	ND	40.59	5.46	1.81
4	125.42	33,083.3	17,733.3	127.5	42	30.59	0.39	11.98	2.27	ND	55.28	5.66	1.5
5	145.09	36,517.9	21,232.1	138.93	57	27.24	1.13	30.65	3.06	ND	65.51	5.47	2
6	61.47	16,575	10,041.7	48.78	18	5.36	0.08	0.4	1.09	ND	30.76	6.07	1.42
7	130.63	34,928.6	20,723.2	107.23	30	34.34	0.53	18.09	2.82	0.05	94.46	5.9	0.93
8	118.75	31,000	28,875	100.8	42	44.49	0.58	25.82	3.2	0.18	79.41	5.66	0.66
9	99.13	20,884.6	17,682.7	82.69	46	ND	0.45	ND	2.54	ND	67.79	6.3	2.24
10	113.3	20,232.1	17,598.2	101.34	42	ND	0.29	ND	3.16	ND	78.18	6.42	2.17
11	70.08	17,044.6	11,857.1	65.87	12	ND	ND	0.74	1.01	ND	39.4	6.52	1.08
12	106.16	25,500	21,000	90.36	34	13.6	0.39	5.34	2.47	ND	62.59	4.4	1.94
13	65.62	13,182.7	8,005.77	49.54	57	ND	ND	2.52	2.38	ND	49.65	5.36	2.04
14	135.45	23,392.9	22,187.5	106.52	50	2.73	0.14	0.39	2.64	ND	58.49	5.35	1.99
15	93.71	19,682.7	12,990.4	59.91	23	ND	ND	0.29	1.93	ND	52.13	4.65	1.71
16	314.81	58,894.2	53,663.5	207.02	149	16.05	0.14	2.31	4.08	0.86	70.4	5.15	2.71

Table 3. *Cont.*

S	Cr	Fe	Co	Ni	Cu	As	Sb	Cd	Zn	Hg	Pb	PH	TOC
17	112.23	20,446.4	21,982.1	93.39	43	12.54	0.56	7.06	3.71	ND	78.78	4.28	1.46
18	146.52	25,500	26,276.8	106.52	66	0.54	1.08	0.59	4.79	0.26	96.25	5.38	1.28

S: Sample, ND: Not detected.

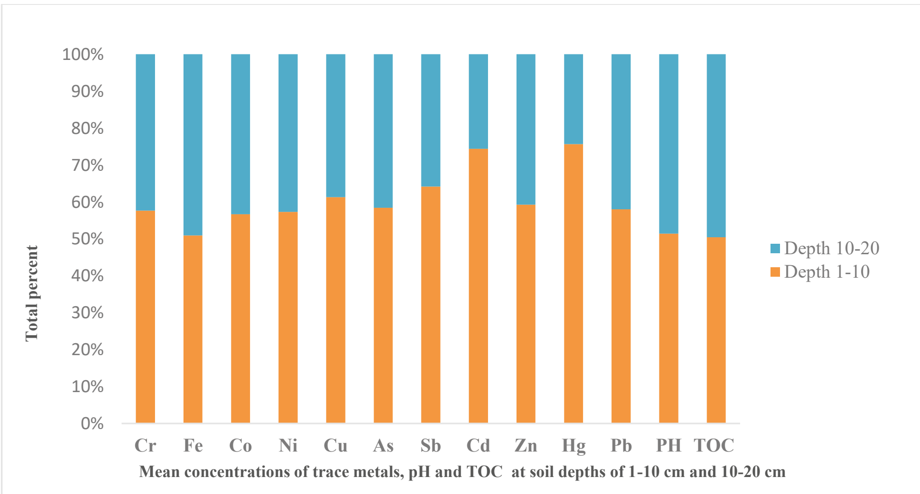


Figure 2. The mean concentrations of trace elements and selected soil properties (pH and TOC) at soil depth of (1–10 and 10–20 cm).

The highest mean of Cr was recorded in the sample from Dujiatian (farmland), whereas the lowest was recorded at Tangjiawan (barren land). The obtained mean result of Cr (140.10 mg/kg) was higher than the soil background value of Wuhan (90 mg/kg) and Hubei Province (86 mg/kg), but less than soil background of China (200 mg/kg). In comparison to the permissible limit of FAO (Food and Agriculture Organization)/ISRIC (International Soil Reference and Information Centre) (2004) (250 mg/kg) and USEPA (1983) (1000 mg/kg) the mean concentration value of Cr (140.10 mg/kg) was lower. However, the individual mean result of single site indicated that the sample from Dujiatian tea farmland (321.73 mg/kg) was above the permissible limit of FAO/ISRIC (2004). As compared to the other findings in Wuhan, the obtained mean value of Cr for this study was higher than the finding of [32] (85 mg/kg), but less than the finding by [33] (152.78 mg/kg). In comparison to the other studies from other places, the mean concentration of Cr for this study was higher than the findings from Cuba (85.9 mg/kg) [59], Bangladesh (53.7, 34.2 mg/kg) [60], India (8.01 mg/kg) [61], Tanzania (7.68 mg/kg) [10], Brazil (20.61 mg/kg) [62], Pakistan (5.86 mg/kg) [63], Iran (48.08, 53.21 mg/kg) [64], along Chao River in China (118 mg/kg) [65] and Northern Pakistan (29.94 mg/kg) [66].

The maximum mean value of Fe (55,398.01 mg/kg) was recorded at Dujiatian farmland site, whereas the lowest value was recorded at Tangjiawan site in the soil from a barren land. The recorded mean value of Fe was below the soil background value of Hubei Province and China (29,400 mg/kg). As compared to the other studies, the mean concentration of Fe (27,304.9 mg/kg) from this study was higher than a study from Bangladesh [11] (1800 mg/kg).The highest concentration of Co (55,580.36 mg/kg was obtained at the Dujiatian site (farmland), whereas, the lowest value (4355 mg/kg) was recorded at Tangjiawan site (barren land). The obtained mean Co was higher than the soil background value of Wuhan, Hubei Province and China (Table 2). The mean concentration value of Co (22,656.94 mg/kg) was above the permissible limit of FAO/ISRIC (2004) and WHO/FAO (2001)

(Table 2). Compared to other findings the mean concentration of Co was higher than the study in Cuba (9.16 mg/kg) [59], Brazil (7.44 mg/kg) [62], Pakistan (7.56 mg/kg) [63], Wuhan (China) (16 mg/kg) [32] and (16.37 mg/kg) [33], Iran (38.5, 16.51 mg/kg) [64], Chao River China (17.5 mg/kg) [65] and Northern Pakistan (36.76 mg/kg) [66].

The recorded mean concentration of Ni (117.80 mg/kg) was higher than the soil background value of Wuhan, Hubei province and China. In addition, the mean value of Ni (117.80 mg/kg) was above the permissible limit of FAO/ISRIC (2004) (100 mg/kg), WHO /FAO (2001), whereas below the permissible limit of USEPA (1983) (500 mg/kg). The highest value of Cu was recorded at Dujiatian (farmland), whereas the lowest was obtained from the soil of Tangjiawan site (grassland soil). According to the obtained result, the mean concentration of Cu recorded was below the permissible limit of FAO/ISRIC (2004), WHO/FAO (2001), USEPA (1983) and China (Table 2). However, the individual result of each site indicated that the concentration of Cu at Dujiatian site exceeded the permissible limit of FAO/ISRIC (2004), WHO/FAO (2001) and USEPA (1983). As compared to the other studies the mean concentration value of Cu (60.73 mg/kg)was less than the finding in Brazil (111.54 mg/kg) [62], Wuhan, China (60.85 mg/kg) [33], Hubei Province, China (386mg/kg), Democratic Republic of Congo (10,320 mg/kg) [67], Iran (100.84 mg/kg) [64], but the result was higher than the findings in soil from Cuba (43.10 mg/kg) [59], Bangladesh (20.6 mg/kg) [11], India (52.72 mg/kg) [61], Tanzania (5.62 mg/kg) [10], Pakistan (18.12 mg/kg) [63], Northern Pakistan (35.28 mg/kg) [66] and in soil along the Chao River in China (46.5 mg/kg) [65].

The highest value of As was recorded at Zhulinyuan site from a barren land, whereas the lowest was recorded at Hanjiafan from a paddy field. The mean concentration result of As was higher than soil background values of Wuhan and Hubei province (Table 2). The average concentration of As recorded was below the permissible limit of USEPA (1983). However, the individual result of each site indicated that the concentrations of As in samples from Fengdouhu, Changdi, Zhulinyuan, Zhoujiawan and Zhujiashan sites were above the soil background value of China (30 mg/kg). The mean comparison of As concentration of this study with the other studies are presented in Table 4.

Table 4. Comparison of the median concentrations (mg/kg) result of trace elements for this study with the other studies.

	Trace Elements											
Countries	**Cr**	**Fe**	**Co**	**Ni**	**Cu**	**As**	**Sb**	**Cd**	**Zn**	**Hg**	**Pb**	**Reference**
China (Wuhan)	140.10	27,304.9	22,656.94	119.12	60.73	15.58	0.58	15.44	3.32	0.15	74.16	This study
Bangladesh	-	18,000	-	24.48	20.06	8.34	-	-	-	-	0.85	[11]
Cuba	85.9	-	9.16	69.57	43.1	20.32	-	0.52	100.2	95.4	14.22	[59]
Bangladesh	53.7	30,404	-	48.1	60	4073.1	-	0.0072	209	486.6	49.66	[60]
China (Xihu district)	53.3	-	7.32	22.9	38.7	-	-	0.387	139	-	70	[8]
China (Hubei)	-	-	-	-	386	35.4	-	2.59	-	-	120	[34]
India	8.01	32.12	-	10.86	52.72	-	-	-	44.72	-	-	[61]
Tanzania	7.68	-	-	-	5.62	-	-	0.22	33.18	-	14.32	[10]
Brazil	20.61	20,273.75	7.44	12.86	111.54	-	13.81	38.31	224.29	954.88	-	[62]
Pakistan	5.86	-	7.56	22.16	18.12	-	-	0.59	-	-	16.18	[64]
Wuhan (China)	85	-	16	34	34	-	-	0.2	85	0.11	33	[32]
Wuhan (China)	152.78	-	16.37	52.87	60.85	-	-	3.98	86.4	-	30.17	[33]
Democratic republic of Congo	-	4.5	990	20	10,320	29	-	49	726	-	135	[67]
Iran	48.08	101,588.89	38.5	74.69	100.84	16	0.22	0.16	72.96	22.30	10.80	[64]
Iran	53.21	33,428.13	16.51	71.56	38.95	11.77	0.21	0.22	71.91	31.45	15.07	[64]
Soil along Chao River (China)	118	-	17.5	20	46.5	6.07	-	0.16	113	0.360	20.3	[65]
Northern Pakistan	29.94	-	36.76	26.61	35.28	-	-	2.04	101.76		4.69	[66]
Miyun Reservoir (China)	-	-	10.7–38.74	-	-	-	-	-	-	-	-	[68]

The obtained mean concentrations of Sb (0.58 mg/kg) was lower than the soil background value of Hubei (1.65 mg/kg) and China (10 mg/kg). The mean result of Sb in this study was lower than the finding from Brazil (13.81 mg/kg) [62], whereas higher than the finding from Iran (0.22, 0.21 mg/kg) [64]. The maximum value of Cd was recorded in the sample from Zhujiashan from a greenhouse, whereas the lowest was recorded at Tangjiawan from barren land. The mean concentration of Cd (15.44 mg/kg) was above the soil background values of Wuhan, Hubei Province and China. In addition the concentration of Cd in this study passes the permissible limit of FAO/ISRIC (2004), WHO/FAO (2001) USEPA (1983). In comparison to the other studies the concentration of Cd for this study was lower than the finding of [62] (13.81 mg/kg) and [67] (49 mg/kg) (Table 4).

The mean concentration of Zn for this was lower than the soil background value of Wuhan, Hubei and China. As compared to the permissible limit of FAO/ISRIC (2004) (500 mg/kg), WHO/FAO (2001) (300 mg/kg), USEPA (1983) (500 mg/kg) and China 250 mg/kg the mean concentration of Zn in this study was lower.

The mean result of Hg was higher than the soil background value of Wuhan, Hubei Province and China (Table 2), but below, however, the permissible limits of FAO/ISRIC (2004), WHO/FAO (2001) and USEPA (1983). However, the result of the individual site indicated that sample from Dujiatian farm (1.13 mg/kg) exceeded the permissible limit of WHO/FAO (2001) (1 mg/kg). In comparison to the other studies, the obtained mean concentration of Hg was lower than the finding of in soil from Cuba [59], Bangladesh [60], Brazil [62], Iran [64] and soil along the Chao River in China [65].

The mean concentration of Pb was below the recommended value of FAO/ISRIC (2004) and USEPA (1983), however, above the permissible limit of WHO/FAO (2001) (Table 2). The concentration of Pb in samples from Changdi, Zhujiashan, Zhulinyuan, Lishuwan, Xinyang, Bomogang and Tianjiaxiaowan site exceeded the soil background value of China (80 mg/kg) (Table 2).

3.3. Relationships of Trace Elements and Selected Soil Properties (pH and TOC)

Pearson's correlation was applied to analyse the relationships between trace elements, soil pH and TOC. The obtained result indicated that there was a significant positive correlation between trace elements at a significant level of 0.01 and 0.05 (2-tailed) whereas, there was a negative and weak correlation between trace elements and selected soil propertis (pH and TOC). Strong positive correlations of Cr with Fe, Co, Ni, Cu, and Hg, Fe with Co, Ni, Cu and Hg, Co with Ni, Cu, and Hg, Cu with Hg, As with Cd, Sb with Cd, Zn and Pb and Zn with Pb were observed at a significance level of 0.01 and 0.05 (Table 5). Moderate and weak correlations between trace elements were also obtained (Table 5). The relationships between trace elements and soil properties indicated that soil pH has a weak positive correlation with Fe, Ni, Cu, As, Sb, Cd and Pb, whereas a negative correlation with Cr, Co, Zn and Hg. The same finding was reported by [47], for Cr, As, Hg and Pb from Guangzhou (China). TOC had a negative correlation with Co, As, Sb, Cd and Hg, and a weak correlation with Cr, Fe, Ni, Cu, Zn and Hg. The same result was reported on Cd by [69]. Weak and negative correlations were observed between soil pH and TOC. A significant positive correlation among elements indicates a common origin. Moderate correlations among elements show those elements share a common source but they are not totally from the same sources, and a weak correlation among those elements indicates they have different origins. A negative correlation among elements and soil properties indicates no influence of soil properties (pH and TOC) on the distribution and concentration of trace elements in the soil.

Table 5. Pearson's correlation coefficient between individual trace elements and soil properties.

Elements	Cr	Fe	Co	Ni	Cu	As	Sb	Cd	Zn	Hg	Pb	pH	TOC
Cr	1												
Fe	0.961 **	1											
Co	0.920 **	0.861 **	1										
Ni	0.874 **	0.917 **	0.730 **	1									
Cu	0.904 **	0.895 **	0.750 **	0.838 **	1								
As	0.535 *	0.692 **	0.489 *	0.725 **	0.544 *	1							
Sb	0.34	0.445	0.235	0.634 **	0.426	0.733 **	1						
Cd	0.422	0.576 *	0.273	0.662 **	0.514 *	0.902 **	0.854 **	1					
Zn	0.618 **	0.609 **	0.599 **	0.719 **	0.631 **	0.635 **	0.770 **	0.609 **	1				
Hg	0.849 **	0.772 **	0.809 **	0.588 *	0.763 **	0.269	-0.003	0.1	0.363	1			
Pb	0.388	0.412	0.375	0.495 *	0.448	0.592 **	0.759**	0.694 **	0.786 **	0.022	1		
pH	−0.07	0.008	−0.165	0.173	0.07	0.222	0.139	0.202	−0.072	−0.182	0.017	1	
TOC	0.148	0.059	−0.024	0.175	0.292	−0.318	−0.048	−0.162	0.087	0.118	−0.026	0.007	1

** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed).

3.4. Factor Analysis

Principal Component Analysis (PCA) is a dimensional reduction tool that is used to reduce large and complex data to a small set of variables which makes it easy for interpretation [70]. It is used to identify correlated variables having common sources [71]. PCA result for trace elements is presented in Table 6. The significant principal component is selected based on the basis of Varimax rotation of Kaiser Criteria with an eigenvalue of greater than 1 [9]. The result indicated that eigenvalues greater than 1 gave a total cumulative value of 86.029%. The variables were correlated with two principal components.

Table 6. Rotated component matrix of trace elements in soil.

Elements	Component	
	1	2
Cr	0.949	0.282
Fe	0.886	0.408
Co	0.904	0.198
Ni	0.737	0.579
Cu	0.846	0.377
As	0.357	0.805
Sb	0.077	0.949
Cd	0.178	0.911
Zn	0.438	0.748
Hg	0.948	−0.103
Pb	0.130	0.848
Eigenvalues % of	7.167	2.296
Variance	65.152	20.877
Cumulative %	65.152	86.029

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

Two-component factors were extracted and the result indicated that the first component with an eigenvalue of 7.167 and with a variance of 65.125% was highly correlated with the high loadings of Cr, Fe, Co, Ni, Cu and Hg (Table 6). According to PCA l values, high and strong positive loading of the elements Cr, Fe, Co, Ni and Cu was connected to anthropogenic factors. The samples were collected from different agricultural fields (paddy field, farm land, vegetable field, plastic greenhouse and barren land). Comparison between fields indicated that the highest value of Cr, Fe, Co, Ni, Cu and Hg was recorded in farmland field. This indicated that the potential sources for these elements might be associated with agricultural inputs (fertilizers, pesticides, herbicides and fungicides) and use of wastewater for irrigation. Different studies have indicated that phosphate fertilizers are highly used in China [6]. A previous study [6] reported that phosphate fertilizers are the main sources of trace elements in the environment. PCA 2 gave eigenvalue of 2.296 and % variance of 20.877 with high loading of As, Sb, Cd Zn and Pb. The PCA 2 might be connected with both anthropogenic and natural factors (parent minerals, weathering processes) and different point and nonpoint sources (application fertilizer, mining, industrial discharge, using wastewater for irrigation). A study by [19] indicated that As might have originated from parent minerals.

3.5. Cluster Analysis

Cluster Analysis (CA) is a method used to group data according to their similarity [19]. Hierarchical Cluster Analysis (HCA) was applied based on the Wards method and Euclidean distance methods [9]. The HCA result (Figure 3) indicated that elements were clustered into two major clusters. The first cluster includes Cr, Fe, Cu, Ni, Co and Hg. The second cluster includes As, Cd, Zn, Pb and Sb. HCA indicated that there is a close cluster between Cr and Fe, Ni and Cu, Co and Hg, As and Cd and Zn and Pb. This implied that those elements were from the same sources. Pearson's correlation analysis also pointed out that those elements had a strong positive correlation which suggested that the elements have the same common origin. The moderate cluster between elements (Figure 3) indicated that these elements share similar sources.

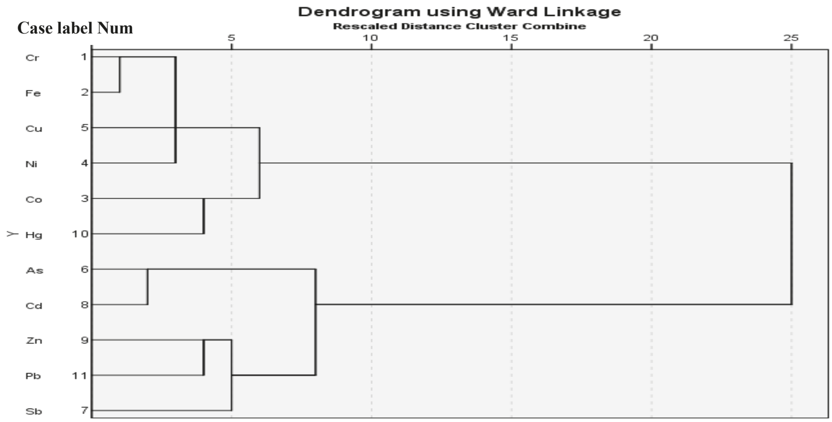


Figure 3. Dendrogram of hierarchal cluster analysis for trace elements in soil using Ward methods.

3.6. Contamination Level and Ecological Risk of Trace Elements

Contamination factor, degree of contamination, ecological risk factor, potential ecological risk index and enrichment factor and the geo-accumulation index of trace elements were evaluated and the results presented in Table 7.

Table 7. Contamination Factor ($C^i{}_f$), Degree of Contamination (C_d), Ecological Risk Factor ($E^i{}_f$), Risk Index (*RI*), Enrichment Factor (*EF*) and Geo-accumulation Index (I_{geo}) of trace elements.

Elements	$C^i{}_f$	$E^i{}_f$	*EF*	Degree of *EF*	I_{geo}	Contamination Level
Cr	1.56	3.11	1.75	Low enrichment	0.05	Uncontaminated to moderately contaminated
Fe	0.93	0	1	Low enrichment	−0.69	Uncontaminated
Co	1471.23	2942.46	1584.12	Extremely high enrichment	9.94	Extremely contaminated
Ni	2.98	14.89	3.44	Moderate enrichment	0.99	Uncontaminated to moderately contaminated
Cu	1.74	8.68	2.13	Moderate enrichment	0.21	Uncontaminated to moderately contaminated
As	1.04	10.39	1.36	Low enrichment	−0.53	Uncontaminated
Sb	0.35	5.27	0.38	Low enrichment	−2.09	Uncontaminated
Cd	77.22	2316.61	90.4	Extremely high enrichment	5.69	Extremely contaminated
Zn	0.03	0.03	0.04	Low enrichment	−5.5	Uncontaminated
Hg	1	40	2.02	Moderate enrichment	−0.58	Uncontaminated
Pb	2.12	10.59	2.99	Moderate enrichment	0.5	Uncontaminated to moderately contaminated
	$C_d = \Sigma C^i{}_f = 1560.2$, $RI = \Sigma E^i{}_f = 5352.03$					

The obtained result of ($C^i{}_f$) indicated that Co > Cd > Ni > Pb > Cu > Cr > As > Hg > Fe > Sb > Zn. According to the obtained value of ($C^i{}_f$), the contamination level of trace elements are grouped as low contamination from Zn (0.03), Fe (0.93) and Sb (0.35), moderate contamination from Cr, Ni, Cu, As, Hg and Pb and a high contamination of soil from Co and Cd. The degree of contamination (C_d) for elements for this the study was 1560.2, which indicates a high degree of soil contamination. Ecological risk factors ($E^i{}_f$) for single elements were in the decreasing order of Co > Cd > Hg > Ni > Pb > As > Cu > Sb > Cr > Zn > Fe. According to the classification given for ecological risk factor for single elements in Table 1, Cr, Ni, Cu, As, Zn, Pb, Fe and Sb had low-risk factors, Hg had a moderate risk factor and Cd and Co had an extremely high risk factors. Thus, Cd, Hg and Co can cause high risk to human and environment in Huangpi district. The same result was reported on Cd and Hg from East Dongting and Honghu Lake in Hunan Province, China [37]. The potential risk index (*RI*) for this study was 5352.03, which indicated a high ecological risk due to these trace elements. The obtained enrichment factor of elements revealed that soil is enriched with Co and Cd, moderately enriched with Ni, Cu, Hg and Pb; and less enriched with Cr, Fe As and Sb (Table 7). This indicated that there is a high rate of anthropogenic disturbance in Huangpi soil. Extremely high enrichment for Cd was also reported by [72] and [73]. The obtained I_{geo} values of the elements in soil were in decreasing order of Co > Cd > Ni >Cu > Pb > Cr > As > Hg > Fe > Sb > Zn. According to the seven classes proposed by [45], the obtained I_{geo} result revealed that the soils under study were extremely contaminated with Cd and Co, moderately contaminated with Cr, Ni, Cu and Pb and less contaminated with As, Zn, Hg, Fe and Sb. The obtained I_{geo} result for Cd was in line with those of [72,74,75].

4. Conclusions

Eleven trace elements in soil from Huangpi district were studied. All trace elements except As, Sb, Hg and Cd were detected in all samples at both soil depths. The obtained mean concentration of Cr, Ni, Cu, As, Cd, Pb were above the soil background value of Wuhan and Hubei Province. The mean concentration values of Cd and Co exceed FAO/ISRIC (2004), WHO/FAO (2001) and USEPA (1983) recommended values of trace elements in soil. The mean concentration of Ni was above the permissible limit of FAO/ISRIC (2004), WHO/FAO (2001), but less than the permissible limit of USEPA (1983). Pearson's correlation result indicated that there was a significant positive correlation among trace elements, whereas, weak and negative correlations between trace elements and soil properties (pH and TOC). PCA, HCA and EF of the soil indicated that anthropogenic factors are the major sources of trace elements in Huangpi soil. The result of contamination factor ($C^i{}_f$) for trace elements were in decreasing order of Co > Cd > Ni > Pb > Cu > Cr > As > Hg > Fe > Sb > Zn. The obtained C_d values indicated that there is high degree of soil contamination. The result of ecological risk factor ($E^i{}_f$) of elements were in the decreasing order of Co > Cd > Hg > Ni > Pb > As > Cu > Sb > Cr > Zn > Fe. Potential ecological risk index (*RI*) result for the studied soil (*RI* > 600) indicated that there is a high ecological risk of elements in the study area. Enrichment factor (EF) showed that there is extremely high enrichment from Cd and Co, moderate from Ni, Cu, Hg and Pb, and low from Cr, Fe, As, Sb and Zn. The geo accumulation index (I_{geo}) results point out that the study area is highly contaminated with Cd and Co. From the obtained result it's concluded that Huangpi soil is contaminated with trace elements. Therefore, more attention should be given and remediation action should be set to minimize the concentration and ecological impacts of trace elements in the study area. Further research should be conducted on human health impact of trace element in Huangpi soil.

Author Contributions: Author Contributions: Sampling, Q.X.; and X.J.; Statistical Analysis, T.G.; Writing—Original Draft Preparation, A.W.T.; Writing-Review and Editing, J.W.

Funding: This research was funded by the Sino-Africa Joint Research Center, Chinese Academy of Sciences, grant number [Y623321K01].

Acknowledgments: The authors would like to thank all members of Pollution Ecology Laboratory for their support during laboratory analysis. In addition, the authors are grateful to all friends that helped to edit final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hashmi, M.Z.; Yu, C.; Shen, H.; Duan, D.; Shen, C.; Lou, L. Risk Assessment of Heavy Metals Pollution in Agricultural Soils of Siling Reservoir Watershed in Zhejiang Province, China. *Bio. Med. Res. Int.* **2013**. [CrossRef]
2. Lu, C.; Wu, Y.; Hu, S.; Zhang, X.; Fu, Y. Distribution and Transport of Residual Lead and Copper Along Soil Profiles in a Mining Region of North China. *Pedosphere* **2016**, *26*, 848–860. [CrossRef]
3. Asati, A.; Pichhode, M.; Nikhil, K. Effect of Heavy Metals on Plants: An Overview. *Int. J. App. Innov. Eng. Manag.* **2016**, *5*, 56–66.
4. Arao, T.; Ishikawa, S.; Murakami, M.; Abe, K.; Maejima, Y.; Makino, T. Heavy metal contamination of agricultural soil and countermeasures in Japan. *Paddy Water Environ.* **2010**, *8*, 247–257. [CrossRef]
5. Masson, P.; Dalix, T.; Bussière, S. Determination of major and trace elements in plant samples by inductively coupled plasma-mass spectrometry. *Commun. Soil Sci. Plant Anal.* **2010**, *41*, 231–243. [CrossRef]
6. Delang, C.O. Environmental & Socio-economic Studies Causes and distribution of soil pollution in China. *Environ. Socio-Econ. Studies.* **2017**, *5*, 1–17.
7. Xu, Y.; Liang, X.; Xu, Y.; Qin, X.; Huang, Q.; Wang, L. Remediation of Heavy Metal-Polluted Agricultural Soils Using Clay Minerals: A. Review. *Pedosphere* **2017**, *27*, 193–204. [CrossRef]
8. Wang, M.; Zhang, H. Accumulation of Heavy Metals in Roadside Soil in Urban Area and the Related Impacting Factors. *Int. J. Environ. Res. Public Health.* **2018**, *15*, 1064. [CrossRef]
9. Ma, L.; Yang, Z.; Li, L.; Wang, L. Source identification and risk assessment of heavy metal contaminations in urban soils of Changsha, a mine-impacted city in Southern China. *Environ. Sci. Pollut. Res.* **2016**, *23*, 17058–17066. [CrossRef] [PubMed]
10. Kibassa, D.; Kimaro, A.A.; Shemdoe, R.S. Heavy metals concentrations in selected areas used for urban agriculture in Dar es Salaam, Tanzania. *Sci. Res. Essays Acad. J.* **2013**, *8*, 1296–1303.
11. Rahman, M.S.; Biswas, P.K.; Al Hasan, S.M.; Rahman, M.M.; Lee, S.H.; Kim, K.H.; Rahman, S.M.; Islam, M.R. The occurrences of heavy metals in farmland soils and their propagation into paddy plants. *Environ. Monit. Assess.* **2018**. [CrossRef] [PubMed]
12. Fosu-Mensah, B.Y.; Addae, E.; Yirenya-Tawiah, D.; Nyame, F. Heavy metals concentration and distribution in soils and vegetation at Korle Lagoon area in Accra, Ghana. *Cogent. Environ. Sci.* **2017**, *3*, 1–14. [CrossRef]
13. Marrugo-negrete, J.; Pinedo-hernández, J.; Díez, S. Assessment of heavy metal pollution, spatial distribution and origin in agricultural soils along the Sinú River Basin, Colombia. *Environ. Res.* **2017**, *154*, 380–388. [CrossRef] [PubMed]
14. Fernandes Azevedo, B.; Barros Furieri, L.; Peçanha, F.M.; Wiggers, G.A.; Frizera Vassallo, P.; Ronacher Simões, M. Toxic Effects of Mercury on the Cardiovascular and Central Nervous Systems. *J. Biomed. Biotechnol.* **2012**, *2012*, 1–11. [CrossRef] [PubMed]
15. Chauhan, G.; Chauhan, P.U.K. Risk Assessment of Heavy Metal Toxicity Through Contaminated Vegetables From Waste. *Int. J. Adv. Technol. Eng. Sci.* **2014**, *51*, 444–460.
16. Olatunji, O.S.; Opeolu, B.O.; Fatoki, O.S.; Ximba, B.J. Heavy metals concentration levels in selected arable agricultural soils in South Western Nigeria. *Int. J. Phys. Sci.* **2013**, *8*, 421–427.
17. Chang, C.Y.; Yu, H.Y.; Chen, J.J.; Li, F.B.; Zhang, H.H.; Liu, C.P. Accumulation of heavy metals in leaf vegetables from agricultural soils and associated potential health risks in the Pearl River Delta, South China. *Environ. Monit. Assess.* **2014**, *186*, 1547–1560. [CrossRef]
18. Tangahu, B.V.; Sheikh Abdullah, S.R.; Basri, H.; Idris, M.; Anuar, N.; Mukhlisin, M. A review on heavy metals (As, Pb, and Hg) uptake by plants through phytoremediation. *Int. J. Chem. Eng.* **2011**. [CrossRef]
19. Yuan, G.L.; Sun, T.H.; Han, P.; Li, J.; Lang, X.X. Source identification and ecological risk assessment of heavy metals in topsoil using environmental geochemical mapping: Typical urban renewal area in Beijing, China. *J. Geochem. Explor.* **2014**, *136*, 40–47. [CrossRef]
20. Hu, H.; Jin, Q.; Kavan, P. A Study of Heavy Metal Pollution in China: Current Status, Pollution-control Policies and Countermeasures. *Sustainability* **2014**, *6*, 5820–5838. [CrossRef]
21. Yu, H.; Wang, J.; Fang, W.; Yuan, J.; Yang, Z. Cadmium accumulation in different rice cultivars and screening for pollution-safe cultivars of rice. *Sci. Total Environ.* **2006**, *370*, 302–309. [CrossRef] [PubMed]

22. Boke, A.; Megersa, N.; Teju, E. Quantitative Determination of the Heavy Metal Levels in the Wild Edible Plant Parts and their Corresponding Soils of the Central and Western Regions of the Oromia State, Ethiopia. *J. Environ. Anal. Toxicol.* **2015**. [CrossRef]
23. Kim, R.Y.; Yoon, J.K.; Kim, T.S.; Yang, J.E.; Owens, G.; Kim, K.R. Bioavailability of heavy metals in soils: Definitions and practical implementation—A critical review. *Environ. Geochem. Health* **2015**, *37*, 1041–1061. [CrossRef] [PubMed]
24. Ogunlaja, O.O.O.; Moodley, R.; Baijnath, H.; Jonnalagadda, S.B. Nutritional evaluation, bioaccumulation and toxicological assessment of heavy metals in edible fruits of FicussurForssk (Moraceae). *J. Environ. Sci. Health* **2017**, *52*, 84–91. [CrossRef] [PubMed]
25. Hussain, A.; Abbas, N.; Arshad, F.; Akram, M.; Khan, Z.I.; Ahmad, K. Effects of diverse doses of Lead (Pb) on different growth attributes of *Zea-Mays* L. *Agric. Sci.* **2013**, *4*, 262–265.
26. Barrachina, A.C.; Carbonell, F.B.; Beneyto, J.M. Arsenic uptake, distribution, and accumulation in tomato plants: Effect of arsenite on plant growth and yield. *J. Plant Nutr.* **1995**, *18*, 1237–1250. [CrossRef]
27. Kamunda, C.; Mathuthu, M.; Madhuku, M. Health risk assessment of heavy metals in soils from witwatersrand gold mining basin, South Africa. *Int. J. Environ. Res. Public Health* **2016**. [CrossRef]
28. Wang, M.Y.; Chen, A.K.; Wong, M.H.; Qiu, R.L.; Cheng, H.; Ye, Z.H. Cadmium accumulation in and tolerance of rice (*Oryza sativa* L.) varieties with different rates of radial oxygen loss. *Environ. Pollut.* **2011**, *159*, 1730–1736. [CrossRef]
29. Lin, Y.S.; Caffrey, J.L.; Lin, J.W.; Bayliss, D.; Faramawi, M.F.; Bateson, T.F. Increased risk of cancer mortality associated with cadmium exposures in Older Americans with low Zinc Intake. *J. Toxicol. Environ. Health* **2013**, *76*, 1–15. [CrossRef]
30. Nieboer, E.; Tsuji, L.J.S.; Martin, I.D.; Liberda, E.N. Human biomonitoring issues related to lead exposure. *Environ. Sci. Process. Impacts* **2013**, *15*, 1824. [CrossRef]
31. Hagberg, A. *Industrial Wastewater Treatment and other Environmental Problems in Wuhan—Is Swedish Technology a Solution?* Uppsala University: Uppsala, Sweden, 2007.
32. Gong, M.; Wu, L.; Bi, X.Y.; Ren, L.M.; Wang, L.; Ma, Z.D.; Bao, Z.Y.; Li, Z.G. Assessing heavy-metal contamination and sources by GIS-based approach and multivariate analysis of urban-rural topsoils in Wuhan, central China. *Environ. Geochem. Health* **2010**, *32*, 59–72. [CrossRef] [PubMed]
33. Qu, M.K.; Li, W.D.; Zhang, C.R.; Wang, S.Q.; Yang, Y.; He, L.Y. Source Apportionment of Heavy Metals in Soils Using Multivariate Statistics and Geostatistics. *Pedosphere* **2013**, *23*, 437–444. [CrossRef]
34. Li, Z.; Zhang, G.; Liu, Y.; Wan, K.Y.; Zhang, R.; Chen, F. Soil Nutrient Assessment for Urban Ecosystems in Hubei, China. *PLoS ONE* **2013**, *8*, e75856. [CrossRef] [PubMed]
35. Lv, J.; Ma, T.; Dong, Z.; Yao, Y. Temporal and Spatial Analyses of the Landscape Pattern of Wuhan City Based on Remote Sensing Images. *Int. J. Geo-Inf.* **2018**, *7*, 340. [CrossRef]
36. Sekabira, K.; Oryem Origa, H.; Basamba, T.A.; Mutumba, G.; Kakudidi, E. Assessment of heavy metal pollution in the urban stream sediments and its tributaries. *Int. J. Environ. Sci. Tech.* **2010**, *7*, 435–446. [CrossRef]
37. Makokha, V.A.; Qi, Y.; Shen, Y.; Wang, J. Concentrations, Distribution, and Ecological Risk Assessment of Heavy Metals in the East Dongting and Honghu Lake, China. *Expo. Health* **2016**, *8*, 31–41. [CrossRef]
38. Weldegebriel, Y.; Chandravanshi, B.S.; Wondimu, T. Concentration levels of metals in vegetables grown in soils irrigated with river water in Addis Ababa, Ethiopia. *Ecotoxicol. Environ. Saf.* **2012**, *77*, 57–63. [CrossRef] [PubMed]
39. Wlakey, A. A critical examination of a rapid method for determination of organic carbon in soils—Effect of variations in digestion conditions and of inorganic soil constituents. *Soil Sci.* **1947**, *63*, 251–257.
40. Liu, H.; Liu, G.; Zhou, C.; Yuan, Z.; Da, C. Geochemical speciation and ecological risk assessment of heavy metals in surface soils collected from the Yellow River Delta National Nature Reserve, China. *Hum. Ecol. Risk Assess.* **2017**, *23*, 1585–1600. [CrossRef]
41. Hakanson, L. An ecological risk index for aquatic pollution control.a sedimentological approach. *Water Res.* **1980**, *14*, 975–1001. [CrossRef]
42. Yang, Z.; Wang, Y.; Shen, Z.; Niu, J.; Tang, Z. Distribution and speciation of heavy metals in sediments from the mainstream, tributaries, and lakes of the Yangtze River catchment of Wuhan, China. *J. Hazard Mater.* **2009**, *166*, 1186–1194. [CrossRef]

43. He, M.; Wang, X.; Wu, F.; Fu, Z. Antimony pollution in China. *Sci. Total Environ.* **2012**, *421–422*, 41–50. [CrossRef] [PubMed]
44. Jiao, X.; Teng, Y.; Zhan, Y.; Wu, J.; Lin, X. Soil heavy metal pollution and risk assessment in Shenyang industrial district, Northeast China. *PLoS ONE* **2015**, *10*, e0127736. [CrossRef] [PubMed]
45. Muller, H.G. Mechanical Properties, Rheology, and Haptaesthesis of Food. *J. Texture Stud.* **1969**, *1*, 38–42. [CrossRef] [PubMed]
46. Ogundiran, M.B.; Osibanjo, O. Mobility and speciation of heavy metals in soils impacted by hazardous waste. *Chem. Speciat. Bioavailab.* **2009**, *21*, 59–69. [CrossRef]
47. Liu, L.; Lu, J.; Zhang, Z.; Zheng, H.; Gao, X.; Zhang, W. Heavy Metals Contamination in Greenhouse Soils and Vegetables in Guanzhong, China. *J. Encapsulation Adsorption Sci.* **2014**, *4*, 80–88. [CrossRef]
48. Bai, Y.; Wang, M.; Peng, C.; Alatalo, J.M. Impacts of urbanization on the distribution of heavy metals in soils along the Huangpu River, the drinking water source for Shanghai. *Environ. Sci. Pollut. Res.* **2016**, *23*, 5222–5231. [CrossRef] [PubMed]
49. Jiang, X.; Lu, W.X.; Zhao, H.Q.; Yang, Q.C.; Yang, Z.P. Potential ecological risk assessment and prediction of soil heavy-metal pollution around coal gangue dump. *Nat. Hazards Earth Syst. Sci.* **2014**, *14*, 1599–1610. [CrossRef]
50. Fang, C.; Moncrieff, J.B. The variation of soil microbial respiration with depth in relation to soil carbon composition. *Plant Soil* **2005**, *268*, 243–253. [CrossRef]
51. Celik, I. Land-use effects on organic matter and physical properties of soil in a southern Mediterranean highland of Turkey. *Soil Tillage Res.* **2005**, *83*, 270–277. [CrossRef]
52. Behbahaninia, A.; Mirbagheri, S.A.; Khorasani, N.; Nouri, J.; Javid, A.H. Heavy metal contamination of municipal effluent in soil and plants. *J. Food Agric. Environ.* **2009**, *7*, 851–856.
53. Rahaman, A.; Sadia Afroze, J.; Bashar, K.; Farhad Ali, M.; Razib Hosen, M. A Comparative Study of Heavy Metal Concentration in Different Layers of Tannery Vicinity Soil and Near Agricultural Soil. *Am. J. Anal. Chem.* **2016**, *7*, 880–889. [CrossRef]
54. Li, J.; Lu, Y.; Yin, W.; Gan, H.; Zhang, C.; Deng, X.; Lian, J. Distribution of heavy metals in agricultural soils near a petrochemical complex in Guangzhou, China. *Environ. Monit. Assess.* **2009**, *153*, 365–375. [CrossRef] [PubMed]
55. Camobreco, V.J.; Richards, B.K.; Steenhuis, T.S.; Peverly, J.H.; McBride, M.B. Movement of heavy metals through undisturbed and homogenized soil columns. *Soil Sci.* **1996**, *161*, 740–750. [CrossRef]
56. Akenga, T.; Sudoi, V.; Machuka, W.; Kerich, E. Heavy Metal Concentrations in Agricultural Farms in Homa Hills Homa Bay County, Kenya. *Int. J. Sci. Res.* **2016**, *5*, 1664–1669.
57. Yabe, J.; Ishizuka, M.; Umemura, T. Current Levels of Heavy Metal Pollution in Africa. *J. Vet. Med. Sci.* **2010**, *72*, 1257–1263. [CrossRef]
58. Mungai, T.M.; Owino, A.A.; Makokha, V.A.; Gao, Y.; Yan, X.; Wang, J. Occurrences and toxicological risk assessment of eight heavy metals in agricultural soils from Kenya, Eastern Africa. *Environ. Sci. Pollut. Res.* **2016**, *23*, 18533–18541. [CrossRef]
59. Accioly, A.; Montero, A.; Ugarte, O.M.; do Nascimento, C.W.; de Aguiar Accioly, A.M.; Biondi, C.M.; da Silva, Y.J. Background concentrations and reference values for heavy metals in soils of Cuba. *Environ. Monit. Assess.* **2015**, *187*, 4198.
60. Rahman, S.H.; Khanam, D.; Adyel, T.M.; Islam, M.S.; Ahsan, M.A.; Akbor, M.A. Assessment of Heavy Metal Contamination of Agricultural Soil around Dhaka Export Processing Zone (DEPZ), Bangladesh: Implication of Seasonal Variation and Indices. *Appl. Sci.* **2012**, *2*, 584–601. [CrossRef]
61. Rakesh Sharma, M.S.; Raju, N.S. Correlation of Heavy Metal contamination with Soil properties of Industrial areas of Mysore, Karnataka, India by Cluster analysis. *Int. Res. J. Environ. Sci.* **2013**, *2*, 22–27.
62. De Almeida Júnior, A.B.; do Nascimento, C.W.A.; Biondi, C.M.; de Souza, A.P.; Barros, F.M.; do Rêgo Barros, F.M. Background and reference values of metals in soils from Paraíba state, Brazil. *Rev. Bras. Cienc. Sol.* **2016**, *40*, 1–13. [CrossRef]
63. Tariq, S.R.; Shafiq, M.; Chotana, G.A. Distribution of Heavy Metals in the Soils Associated with the Commonly Used Pesticides in Cotton Fields. *Scientifica* **2016**. [CrossRef] [PubMed]
64. Soltani, N.; Keshavarzi, B.; Moore, F.; Sorooshian, A.; Ahmadi, M.R. Distribution of potentially toxic elements (PTEs) in tailings, soils, and plants around Gol-E.-Gohar iron mine, a case study in Iran. *Environ. Sci. Pollut. Res.* **2017**, *24*, 18798–18816. [CrossRef] [PubMed]

65. Qin, F.; Ji, H.; Li, Q.; Guo, X.; Tang, L.; Feng, J. Evaluation of trace elements and identification of pollution sources in particle size fractions of soil from iron ore areas along the Chao River. *J. Geochem. Explor.* **2014**, *138*, 33–49. [CrossRef]
66. Rehman, I.U.; Ishaq, M.; Ali, L.; Khan, S.; Ahmad, I.; Din, I.U. Enrichment, spatial distribution of potential ecological and human health risk assessment via toxic metals in soil and surface water ingestion in the vicinity of Sewakht mines, district Chitral, Northern Pakistan. *Ecotoxicol. Environ. Saf.* **2018**, *154*, 127–136. [CrossRef] [PubMed]
67. Pourret, O.; Lange, B.; Bonhoure, J.; Colinet, G.; Decrée, S.; Mahy, G.; Séleck, M.; Shutcha, M.; Faucon, M.-P. Assessment of soil metal distribution and environmental impact of mining in Katanga (Democratic Republic of Congo). *Appl. Geochem.* **2016**, *64*, 43–55. [CrossRef]
68. Gao, L.; Gao, B.; Zhou, Y.; Xu, D.; Sun, K. Predicting remobilization characteristics of cobalt in riparian soils in the Miyun Reservoir prior to water retention. *Ecol. Indic.* **2017**, *80*, 196–203. [CrossRef]
69. Khaledian, Y.; Pereira, P.; Brevik, E.C.; Pundyte, N.; Paliulis, D. The Influence of Organic Carbon and pH on Heavy Metals, Potassium, and Magnesium Levels in Lithuanian Podzols. *Land Degrad. Dev.* **2017**, *28*, 345–354. [CrossRef]
70. Constantin, C. Principal Component Analysis—A Powerful Tool in Computing Marketing Information. *Bull. Transilv. Univ. Brasov Ser. V Econ. Sci.* **2014**, *7*, 25–30.
71. Yang, J.; Teng, Y.; Song, L.; Zuo, R. Tracing sources and contamination assessments of heavy metals in road and foliar dusts in a typical mining city, China. *PLoS ONE* **2016**, *11*, e0168528. [CrossRef]
72. Abdulqaderismaeel, W.; Kusag, A. Enrichment Factor and Geo-accumulation Index for Heavy Metals at Industrial Zone in Iraq. *IOSR J. Appl. Geol. Geophys.* **2015**, *3*, 2321–2990.
73. Zhang, Z.; Juying, L.; Mamat, Z. Sources identification and pollution evaluation of heavy metals in the surface sediments of Bortala River, Northwest China. *Ecotoxicol. Environ. Saf.* **2016**, *126*, 94–101. [CrossRef] [PubMed]
74. Gupta, S.; Jena, V.; Matic, N.; Kapralova, V.; Solanki, J.S. Assessment of Geo-Accumulation Index of Heavy Metal and Source of Contamination By Multivariate Factor Analysis. *Int. J. Hazard Mater.* **2014**, *2*, 18–22.
75. Alshahri, F.; El-Taher, A. Assessment of Heavy and Trace Metals in Surface Soil Nearby an Oil Refinery, Assessment of Heavy and Trace Metals in Surface Soil Nearby an Oil Refinery, Saudi Arabia, Using Geoaccumulation and Pollution Indices. *Arch. Environ. Contam. Toxicol.* **2018**, *75*, 390–401. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

Compliance with Tuberculosis Screening in Irregular Immigrants

Francesca Bonvicini [1], Silvia Cilloni [2], Rossano Fornaciari [1], Carmen Casoni [1], Cristina Marchesi [1], Marina Greci [1], Lucia Monici [1], Fausto Nicolini [1] and Marco Vinceti [2,*]

1. Azienda Unità Sanitaria Locale di Reggio Emilia, Reggio Emilia 42122, Italy; francesca.bonvicini2@ausl.re.it (F.B.); rossano.fornaciari@ausl.re.it (R.F.); carmen.casoni@ausl.re.it (C.C.); cristina.marchesi@ausl.re.it (C.M.); marina.greci@ausl.re.it (M.G.); lucia.monici@ausl.re.it (L.M.); fausto.nicolini@ausl.re.it (F.N.)
2. Università degli Studi di Modena e Reggio Emilia, Dipartimento di Scienze Biomediche, Metaboliche e Neuroscienze, Modena 41125, Italy; silvia.cilloni@unimore.it

* Correspondence: marco.vinceti@unimore.it; Tel.: +39-059-205-5481

Received: 23 November 2018; Accepted: 21 December 2018; Published: 23 December 2018

Abstract: Tuberculosis (TB) is a serious public health problem in many regions of the world, especially in the poorest areas. For this reason, screening for active and latent forms must be considered when dealing with high-risk groups such as irregular immigrants in Western countries. We conducted a retrospective cohort study by recruiting subjects aged ≥15 years who underwent a tuberculin skin test at a dedicated National Health Service Centre in a northern Italian province between 1 January 2012 and 31 December 2013. These participants were followed up until 31 December 2016. We aimed at evaluating an experimental protocol for active and latent tuberculosis screening, focusing on patient compliance, feasibility, and capability to detect clinical forms of the disease. We enrolled 368 irregular immigrants, i.e., immigrants not having a valid residence permit and who were therefore not entitled to choose a general practitioner. In total, 90.22% of these completed all the steps for the screening of active TB, while 87.33% also undertook screening for the latent form of the disease. Homelessness, self-reported prostitution, female sex, and employment status adversely affected compliance. Chronic alcohol consumption was associated with increased risk of no beginning or interruption of the procedures. All of the five patients with active TB successfully completed the treatment. Overall, adherence to the screening program was high compared to other studies in immigrants, possibly owing to organizational factors such as the availability of cultural mediators, the network between the different health services, the presence of dedicated nursing staff and a free-of-charge service. In addition, selected vulnerable subgroups should be targeted using tailored screening and follow-up programs.

Keywords: tuberculosis; screening for tuberculosis; public health; immigrants

1. Introduction

Tuberculosis (TB) is the main cause of death from infectious disease globally. Drug-resistant forms of the disease are a major risk to global health security [1]. The WHO estimates that 10.4 million individuals became ill with TB and 1.7 million died in 2016. Despite a fall in mortality rate by 3% per year, TB remains the ninth leading cause of death worldwide, resulting in 1.3 million deaths in HIV-negative people and almost 400,000 deaths in HIV-positive people [2,3]. TB epidemiology in most low-incidence countries is characterized by a low rate of transmission in the general population, occasional outbreaks and a majority of TB cases generated from progression of latent tuberculosis infection (LTBI) rather than local transmission [4]. Migration flows have changed drastically since the beginning of the 21st century. Because most immigrants come from countries with a high incidence of tuberculosis, the contribution of the migrant population to new cases of tuberculosis is stronger in relative terms than for its weight in the total population [5]. For this reason, it is necessary to both diagnose TB early by including universal drug susceptibility testing, and to implement systematic screening for TB in selected high-risk groups. Early detection helps to reduce the risk of further TB transmission, poor treatment outcomes and undesirable health sequelae, thus also decreasing adverse social and economic consequences of the disease [6,7].

In the last 50 years, the annual incidence of TB in Italy decreased by about 70%, from around 25 to 7 cases per 100,000 individuals, and it now seems to be quite stable [8]. In this country, however, only limited data are available about TB epidemiology in potentially high-risk groups such as undocumented immigrants. Consequently, TB prevalence and risk factors in illegal immigrants are unknown. Furthermore, they have not been monitored over time. In spite of this, while national guidelines [9] recommend LTBI screening in high-risk subjects such as pulmonary TB contacts, HIV-infected patients, and patients undergoing immunosuppressive therapy and health care workers, less attention is devoted to the immigrant population, whereby early detection of active and LTBI cases should be pursued using the following strategies [10]: screening for both active TB and LTBI and therapy of LTBI in children, chest X-ray (CXR) screening for active TB in symptomatic subjects especially recent irregular immigrants. In addition, screening for active and latent TB in asymptomatic adults with recent immigration (<5 years) or living in social and health conditions of marginalization has no strong evidence. This also applies to undocumented immigrants, although these may represent an important source of LTBI and active TB. At the Centre for Health of Foreign Family (CFF), systematic screening for both active TB and LTBI had been performed since January 2012.

2. Materials and Methods

2.1. Setting and Clinical Procedures

The study protocol was approved by the Ethics Committee of the Reggio Emilia Province on 19 April 2017 and by the Reggio Emilia Local Health Unit on 5 June 2017 (document code 2017/DS/0038).

This study took place in the CFF of Reggio Emilia, an area in northern Italy with approximately 550,000 inhabitants. The CFF is an outpatient clinic located in the city centre. It is devoted to immigrants without a valid residence permit, who are not entitled to choose a general practitioner.

In accordance with national recommendations [10], the reference test for the diagnosis of LTBI is the tuberculin skin test (TST) based on the Mantoux method. TST is carried out by the nursing staff in accordance with the Mantoux method, by inoculating 0.1 mL of purified protein [11]. The nursing or medical staff read the induration in millimetres after 48–72 h. If positive, the TST is always evaluated by medical personnel. The cut-off for TST positivity is usually 10 mm. In case of immunosuppression, it drops to 5 mm. Screening is performed according to a local protocol based on the combined use of TST and interferon gamma release assay (IGRA). Specifically we have used Quantiferon-TB Gold (QFT) for evaluating the possible presence of active TB. In Figure 1 it is possible to see the algorithm used at CFF for the screening of active TB and LTBI; in immigrants with any symptom of TB and/or peripheral lymphadenopathy, TST is performed as a first step. As far as TB symptoms are concerned, we consider at least one of the following: cough, haemoptysis, fever, chest pain, weight loss, fatigue, night sweats, chills or loss of appetite. The clinician visiting the patient may prescribe CXR at the time of TST and at any time during the screening protocol if a high risk for active TB is estimated. If the CXR shows any abnormality, the patient is referred for the pneumological examination (PE) or hospitalization. If CXR is negative, the patient's LBTI diagnosis can be proceeded with. TST-positive patients undergo CXR (if not yet performed) and possibly QFT. QFT is generally prescribed for confirmation for the purpose of offering LTBI treatment. The indication for LTBI treatment is decided by the pneumologist, after patient examination and evaluation of CXR and QFT results. For LTBI treatment, we generally prescribe isoniazid for 6 months at a dose of 5 mg/kg/day, combined with vitamin B6, in order to prevent neurological adverse effects. Isoniazid preventive therapy (IPT) is recommended in people aged ≤35 years and in people of any age with HIV co-infection. In case of contraindication, failure to accept or need to discontinue IPT, patients with LTBI undergo health surveillance (HS) through clinical examination every six months for two years, for the purpose of early detection of the onset of active TB symptoms.

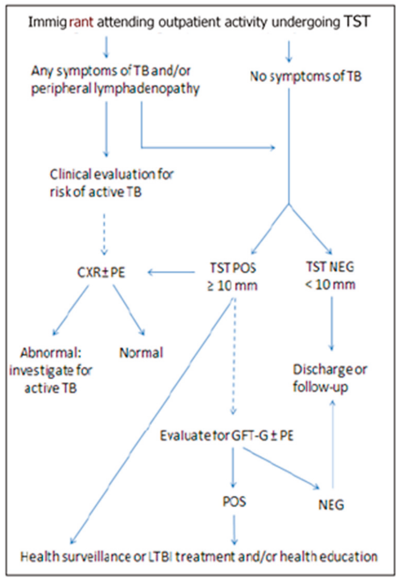


Figure 1. Procedures for active tuberculosis (**TB**) and latent tuberculosis infection (**LTBI**) screening at the Centre for Health of Foreign Family (**CFF**) of Reggio Emilia. **TST**: Tuberculin Skin Test; **CXR:** chest X-ray; **PE:** Pneumological Examination; **POS:** positive; **NEG:** negative.

All examinations and visits are free of charge. Cultural mediators belonging to the most common ethnic groups of patients are constantly present to assist health personnel and immigrants during outpatient activity.

2.2. Study Design and Definitions

In this retrospective cohort study, all foreign-born patients aged ≥15 years who underwent TST at the CFF between 1 January 2012 and 31 December 2013 were enrolled, provided that there was no evidence of previous contact with a case of pulmonary TB or positive TST. The period of active follow-up in patients undergoing HS ended on 31 December 2016. We defined patients with LTBI as immigrants with both positive TST and QFT and patients with positive TST but QFT not requested or not performed. We considered patients without LTBI as subjects with negative TST and with positive TST but negative QFT. According to the World Health Organization [12], a bacteriologically confirmed TB case was one from whom a biological specimen was positive by smear microscopy, culture or rapid diagnostic tests (such as Xpert MTB/RIF assay that is a nucleic acid amplification test). A clinically diagnosed TB case was one diagnosed with active TB by a clinician or other medical practitioner who decided to start a full course of TB treatment; this definition included cases diagnosed on the basis of X-ray abnormalities or suggestive histology and extra pulmonary cases without laboratory confirmation.

2.3. Data Analysis

We analysed compliance with screening for active TB, LTBI, and the whole protocol, also taking patient characteristics into account. We also assessed the capacity of the adopted procedures to detect LTBI and active TB. In order to do this, we computed odds ratios (ORs) along with their 95% confidence interval (CI) using crude and multivariable logistic regression models. The potential confounders included in multivariable analysis were adjusted for sex, age, presence of cough, education, knowledge of Italian, employment status (employed versus unemployed), years from arrival in Italy and homelessness. We included knowledge of Italian language because it could affect compliance, influencing patient's ability to understand the motivations of the clinical examinations and therapies proposed. When the use of multivariable analysis led to exceedingly high statistical instability, we limited our assessment to crude estimates only (unadjusted for other variables). Risk analysis was performed by calculating odds ratios estimated from conditional logistic regression with crude and multivariate models. When the odds ratios could not be calculated, we applied the chi2 test and probability according to Fisher's exact Test. For the evaluation of the differences between continuous variables, we applied the *t*-test. For PE, HS, and IPT, we performed only the crude analysis because of the smallness of the sample, causing too much instability in the statistical analysis.

3. Results

From 2012 to 2016, TST was administered to a total number of 404 immigrants. Of these, 36 patients were excluded from our analysis: 8 were born in Italy from undocumented parents and 28 were younger than 15 years. The final study population consisted of 368 immigrants: 186 (50.5%) and 182 (49.5%) were tested in 2012 and 2013, respectively. The demographic characteristics of the study population are reported in Table S1. Adherence to the various stages of the screening procedures and treatment outcomes are shown in Table 1.

Table 1. Compliance with the different steps of the algorithm and treatment outcomes.

Steps of the Algorithm	*N* Performed/Prescribed	%
TST read	349/368	94.8
CXR performed	134/151	88.7
QFT performed	89/110	80.9
PE attended	78/84	92.9
Screening TB (TST, CXR, PE) completed	332/368	90.2
negative TST with CXR indication	12/12	100
positive TST	122/139	87.8
Screening LTBI (TST, CXR, QFT, PE) completed	317/363	87.3
negative TST with CXR indication	12/12	100
positive TST	107/134	79.9
LTBI treatment		
refusal or voluntary interruption	8/28	28.6
medical contraindication or interruption	6/28	21.4
treatment completed	14/28	50.0
LTBI health surveillance never started	17/41	41.5
voluntary interruption	18/41	43.9
completed	6/41	14.6
TB treatment completed	5/5	100
Whole algorithm completed	281/368	76.4
negative TST with CXR indication	12/12	100
positive TST	71/139	51.1

TST: tuberculin skin test; CXR: chest-X-ray; QFT: quantiferon test; PE: pneumological examination; TB: tuberculosis; LTBI: latent tuberculosis infection.

Patients who did not return for the reading of TST amounted to 15.2%. Patients who did not undergo CXR, QFT and PE despite the prescriptions amounted to 11.3%, 19.1% and 7.1%, respectively. Finally, 90.2% of subjects completed all steps expected for the screening of active TB, considering TST reading and CXR/PE, when prescribed. In immigrants with a positive TST result, adherence was lower, since their path is usually longer. A total of 5 cases of active TB were diagnosed during the study period. After discharge from hospital, they were followed monthly at TBOA (Outpatient Activity dedicated to management of LTBI and TB at CFF) and they all successfully completed the treatment.

The overall compliance with LTBI screening, which involved TST, CXR, QFT, and PE was 87.3%. Also in this case, adherence of patients with positive TST was observed to be lower (79.9%). LTBI treatment was prescribed to 28 patients, while only 20 out of 28 patients started prophylaxis with Isoniazid, because eight patients had increased levels of transaminases for chronic alcohol abuse or HBsAg, or they refused therapy. Eventually, 14 of the 20 patients who started IPT were able to complete the treatment, whereas 6 patients discontinued therapy voluntarily or after medical indication. HS for 2 years was indicated in 41 patients yet only 6 completed all follow-up. Finally, 76.4% of immigrants completed the whole diagnostic and therapeutic protocol for active TB and LTBI. Considering only TST-positive ones, adherence decreased to 51%.

Tables 2 and 3 deal with the full path for the screening of active TB and LTBI. In multivariable analysis regarding the screening of active TB (Table 2), it is possible to appreciate that female sex, homelessness and age between 25–34 years were associated with a considerably lower rate of completion of the screening. On the contrary, a higher level of education, time of arrival in Italy for at least 5 years and the presence of cough were all associated with better compliance. With regard to LTBI (Table 3), being a native of the Americas, age between 25 and 34 years, employment, and homelessness were associated with considerably lower adherence. Refugee status, time of arrival in Italy for at least 5 years and the presence of cough were associated with better compliance. It is interesting to note that being from the Americas was associated with very low compliance for both active TB and LTBI screening (OR 0.27, 95% CI 0.05–1.44; OR 0.32, 95% CI 0.06–1.64, respectively). Table 4 outlines compliance with the full algorithm. As can be seen, age had a negative impact on adherence. In addition, being employed, homelessness and prostitution were associated with considerably lower compliance. Being a native of South-East Asia and the Western Pacific (OR 4.48, 95% CI 1.57–12.73), presence of cough, having been in Italy for a longer time, in particular for more than 5 years, and the attainment of higher education levels were observed to be related to a much higher probability of completion. Female Chinese patients' compliance with the protocol was substantially higher than in the rest of the population (OR 55.63, 95% CI 5.35–578), while Nigerians, Georgians, and Tunisians had lower levels of adhesion (Table S2).

Considering compliance with TST reading, both bivariate and multivariate analysis show that American and European origin, age between 25–34 years old, prostitution and homelessness were associated with considerably lower rates of return for the reading of the test. On the other hand, higher levels of education were associated with better compliance. In multivariable analysis, the presence of cough was the only symptom associated with greater compliance (OR 7.61, 95% CI 0.92–63.2) (Table S3). Compliance with CXR showed a positive correlation to the presence in Italy for at least 5 years, while a negative correlation was associated to the female sex (Table S4). Compliance with QFT was inversely correlated with education and positively associated with the age group 25–34 years (OR 3.80, 95% CI 0.87–16.53) (Table S5). Female sex, being a native of the Americas, and sufficient or good knowledge of Italian were associated with a 100%-rate of IPT completion. Conversely, refugees showed a tendency to refuse or to voluntarily stop the treatment, and higher education levels inversely correlated with the probability of completion (Table S6).

Table 2. Association between selected characteristics and compliance with screening for active tuberculosis (TB).

Sample Characteristics	TB Screening Started/Completed		Bivariate OR (95% CI)	p-Value	Multivariate OR (95% CI)	p-Value
	N	%				
Sex:						
Male	222/243	91.36	1.00 (reference)		1.00 (reference)	
Female	110/125	88.00	0.69 (0.34–1.40)	0.306	0.47 (0.22–1.03)	0.061
Age at TST in Years:						
15–24	90/97	92.78	1.00 (reference)		1.00 (reference)	
25–34	113/130	86.92	0.52 (0.21–1.30)	0.161	0.44 (0.16–1.18)	0.103
35–44	87/95	91.58	0.85 (0.29–2.43)	0.756	0.57 (0.18–1.84)	0.344
≥45	42/46	91.30	0.82 (0.23–2.94)	0.757	0.90 (0.20–4.03)	0.891
Continuous OR			0.99 (0.70–1.41)	0.972	0.96 (0.64–1.43)	0.827
Region of Origin:						
Africa	103/111	92.79	1.00 (reference)		1.00 (reference)	
Eastern Mediterranean	75/83	90.36	0.73 (0.26–2.03)	0.544	0.75 (0.25–2.27)	0.607
Europe	60/72	83.33	0.39 (0.15–1.00)	0.051	0.44 (0.15–1.30)	0.138
SE Asia/West Pacific	85/90	94.44	1.32 (0.42–4.19)	0.637	2.49 (.57–10.87)	0.226
Americas	9/12	75.00	0.23 (0.05–1.04)	0.056	0.27 (0.05–1.44)	0.125
TB Incidence °:			1.26 (0.89–1.79)	0.187	1.15 (0.79–1.69)	0.458
Any Symptoms *:						
No	211/236	89.41	1.00 (reference)			
Yes	121/132	91.67	1.30 (0.62–2.74)	0.485		
Cough:						
No	249/280	88.93	1.00 (reference)		1.00 (reference)	
Yes	83/88	94.32	2.07 (0.78–5.49)	0.145	3.23 (1.04–10.05)	0.043
Education **:			1.36 (0.88–2.12)	0.168	1.49 (0.93–2.40)	0.101
Italian Language **°:			0.89 (0.52–1.50)	0.652	0.90 (0.51–1.58)	0.721
Employment Status:						
No	234/258	90.70	1.00 (reference)		1.00 (reference)	
Yes	90/102	88.24	0.77 (0.37–1.60)	0.484	0.79 (0.32–1.96)	0.612
Years in Italy:						
<5	223/252	88.49	1.00 (reference)		1.00 (reference)	
≥5	102/109	93.58	1.89 (0.80–4.47)	0.144	3.29 (1.18–9.22)	0.023
Continuous OR			1.06 (0.96–1.18)	0.248	1.18 (0.72–1.92)	0.513
Homelessness:						
No	313/345	90.72	1.00 (reference)		1.00 (reference)	
Yes	19/23	82.61	0.49 (0.16–1.52)	0.213	0.35 (0.10–1.25)	0.107
Refugees:						
No	304/340	89.41	-			
Yes	28/28	100.00	-	0.070 ***		
Prostitution:						
No	323/357	90.48	1.00 (reference)			
Yes	9/11	81.82	0.47 (0.10–2.28)	0.352		
Pregnancy:						
No	323/359	89.97	-			
Yes	9/9	100.00	-	0.319 ***		
TST Result:						
Negative	210/210	100.00	-			
Positive	122/139	87.77	-	0.000 ***		

°: TB incidence in the country of origin is sub-divided into 4 categories: 0–49, 50–99, 100–199, ≥200 per 100,000 population/year. *: Meaning at least one among cough, hemoptysis, fever, chest pain, weight loss, fatigue, night sweats, chills or loss of appetite. **: Education is sub-divided into three categories: illiterate/primary school, secondary school, high school/degree. **°: Knowledge of Italian is sub-divided into three categories: none, sufficient, good. ***: Two-sample Mann–Whitney test. TB: tuberculosis; LTBI: latent tuberculosis infection; TST: tuberculin skin test; SE Asia: South-East Asia; OR: odds ratio; CI: confidence interval.

Table 3. Association between selected characteristics and compliance with screening for latent TB infection (LTBI).

Sample Characteristics	LTBI Screening Started/Completed		Bivariate OR (95% CI)	*p*-Value	Multivariate OR (95% CI)	*p*-Value
	N	%				
Sex:						
Male	211/241	87.55	1.00 (reference)		1.00 (reference)	
Female	106/122	86.89	0.94 (0.49–1.80)	0.857	0.76 (0.38–1.54)	0.451
Age at TST in Years:						
15–24	88/97	90.72	1.00 (reference)		1.00 (reference)	
25–34	107/127	84.25	0.55 (0.24–1.26)	0.157	0.54 (0.22–1.31)	0.173
35–44	80/93	86.02	0.63 (0.26–1.55)	0.314	0.57 (0.21–1.56)	0.273
≥45	42/46	91.30	1.07 (0.31–3.69)	0.910	1.39 (0.33–5.86)	0.658
Continuous OR			0.98 (0.72–1.34)	0.892	1.01 (0.70 –1.45)	0.956
Region of Origin:						
Africa	97/109	88.99	1.00 (reference)		1.00 (reference)	
Eastern Mediter.	73/83	87.95	0.90 (0.37–2.20)	0.823	0.93 (0.35–2.47)	0.892
Europe	60/72	83.33	0.62 (0.26–1.47)	0.275	0.72 (0.27–1.91)	0.509
SE Asia/West Pacific	78/87	89.66	1.07 (0.43–2.67)	0.881	2.39 (0.72–7.93)	0.155
Americas	9/12	75.00	0.37 (0.09–1.56)	0.177	0.32 (0.06–1.64)	0.174
TB Incidence °:			1.11 (0.82–1.52)	0.489	0.99 (0.70–1.40)	0.975
Any Symptoms *:						
No	203/235	86.38	1.00 (reference)			
Yes	114/128	89.06	1.28 (0.66–2.51)	0.464		
Cough:						
No	240/279	86.02	1.00 (reference)		1.00 (reference)	
Yes	77/84	91.67	1.79 (0.77–4.16)	0.178	2.38 (0.92–6.17)	0.073
Education **:			1.09 (0.73–1.63)	0.660	1.09 (0.71–1.68)	0.688
Italian Language **°:			1.13 (0.68–1.87)	0.636	1.15 (0.67–1.95)	0.616
Employment Status:						
No	228/257	88.72	1.00 (reference)		1.00 (reference)	
Yes	81/98	82.65	0.61 (0.32–1.16)	0.131	0.56 (0.25–1.26)	0.162
Years in Italy:						
<5	215/249	86.35	1.00 (reference)		1.00 (reference)	
≥5	95/107	88.79	1.25 (0.62–2.52)	0.530	1.88 (0.80–4.40)	0.145
Continuous OR			1.04 (0.95–1.13)	0.376	0.99 (0.64–1.54)	0.972
Homelessness:						
No	299/340	87.94	1.00 (reference)		1.00 (reference)	
Yes	18/23	78.26	0.49 (0.17–1.40)	0.185	0.35 (0.11–1.12)	0.076
Refugees:						
No	290/335	86.57	1.00 (reference)			
Yes	27/28	96.43	4.19 (0.56–31.60)	0.165		
Prostitution:						
No	309/352	87.78	1.00 (reference)			
Yes	8/11	72.73	0.37 (0.09–1.45)	0.155		
Pregnancy:						
No	309/355	87.04	-			
Yes	8/8	100.00	-	0.277 ***		
TST result:						
Negative	210/210	100.00	-			
Positive	107/134	79.85	-	0.000 ***		

°: TB incidence in the country of origin is sub-divided into four categories: 0–49, 50–99, 100–199, ≥200 per 100,000 population/year. *: Meaning at least one among cough, hemoptysis, fever, chest pain, weight loss, fatigue, night sweats, chills or loss of appetite. **: Education is sub-divided into three categories: illiterate/primary school, secondary school, high school/degree. **°: Knowledge of Italian is sub-divided into three categories: none, sufficient, good. ***: Two-sample Mann–Whitney test. TB: tuberculosis; LTBI: latent tuberculosis infection; TST: tuberculin skin test; SE Asia: South-East Asia.

Table 4. Association between selected characteristics and compliance with the full protocol (bivariate and multivariate analysis).

Protocol Started/Completed			Bivariate OR (95% CI)	p-Value	Multivariate OR (95% CI)	p-Value
Sample characteristics	***N***	**%**				
Sex:						
Male	183/243	75.31	1.00 (reference)		1.00 (reference)	
Female	98/125	78.40	1.19 (0.71–1.99)	0.509	1.18 (0.65–2.12)	0.588
Age at TST in Years:						
15–24	83/97	85.57	1.00 (reference)		1.00 (reference)	
25–34	97/130	74.62	0.50 (0.25–0.99)	0.046	0.44 (0.21–0.93)	0.031
35–44	70/95	73.68	0.47 (0.23–0.98)	0.043	0.39 (0.17–0.88)	0.024
≥45	31/46	67.39	0.35 (0.15–0.81)	0.014	0.26 (0.10–0.68)	0.006
Continuous OR			0.74 (0.58–0.94)	0.014	0.67 (0.50–0.89)	0.007
Region of Origin:						
Africa	85/111	76.58	1.00 (reference)		1.00 (reference)	
Eastern Mediterranean	62/83	74.70	0.90 (0.47–1.75)	0.763	0.87 (0.41–1.85)	0.723
Europe	49/72	68.06	0.65 (0.34–1.26)	0.205	0.90 (0.42–1.97)	0.801
SE Asia/West Pacific	76/90	84.44	1.66 (0.81–3.41)	0.167	4.48 (1.57–12.73)	0.005
Americas	9/12	75.00	0.92 (0.23–3.64)	0.903	0.67 (0.14–3.16)	0.617
TB Incidence °:			0.91 (0.72–1.16)	0.461	0.79 (0.59–1.05)	0.107
Any Symptoms *:						
No	179/236	75.85	1.00 (reference)			
Yes	102/132	77.27	1.08 (0.65–1.79)	0.758		
Cough:						
No	212/280	75.71	1.00 (reference)		1.00 (reference)	
Yes	69/88	78.41	1.16 (0.65–2.07)	0.604	1.75 (0.88–3.46)	0.110
Education **:			1.18 (0.86–1.62)	0.306	1.24 (0.88–1.74)	0.220
Italian Language **°:			1.26 (0.85–1.89)	0.254	1.30 (0.82–2.04)	0.265
Employment Status:						
No	201/258	77.91	1.00 (reference)		1.00 (reference)	
Yes	75/102	73.53	0.79 (0.46–1.34)	0.377	0.59 (0.30–1.17)	0.133
Years in Italy:						
<5	188/252	74.60	1.00 (reference)		1.00 (reference)	
≥5	88/109	80.73	1.43 (0.82–2.48)	0.209	2.94 (1.44–6.01)	0.003
Continuous OR			1.06 (0.99–1.13)	0.119	1.30 (0.90–1.86)	0.160
Homelessness:						
No	267/345	77.39	1.00 (reference)		1.00 (reference)	
Yes	14/23	60.87	0.45 (0.19–1.09)	0.077	0.35 (0.13–0.96)	0.042
Refugees:						
No	258/340	75.88	1.00 (reference)			
Yes	23/28	82.14	1.46 (0.54–3.97)	0.456		
Prostitution:						
No	273/357	76.47	1.00 (reference)			
Yes	8/11	72.73	0.82 (0.21–3.16)	0.774		
Pregnancy:						
No	272/359	75.84	-			
Yes	9/9	91.67	-	0.091 ***		
TST Result:						
Negative	210/210	100.00	-			
Positive	71/139	51.08	-	0.000 ***		

° TB incidence in the country of origin is sub-divided into 4 categories: 0–49, 50–99, 100–199, ≥200 per 100,000 population/year. * Meaning at least one among cough, hemoptysis, fever, chest pain, weight loss, fatigue, night sweats, chills or loss of appetite. ** Education is sub-divided into three categories: illiterate/primary school, secondary school, high school/degree. **° Knowledge of Italian is sub-divided into three categories: none, sufficient, good. *** Two-sample Mann–Whitney test. TST: tuberculin skin test; SE Asia: South-East Asia; TB: tuberculosis.

4. Discussion

In our study, adherence with the screening of active TB and LTBI and with pulmonary TB therapy was very high compared to other recent investigations conducted in the USA, Switzerland, and Italy [13–17]. However, a direct comparison between these studies is difficult because study design, countries of origin and sample size were different.

In our study, we identified few fragile groups on which to focus to improve the performance of the screening program [18]. In particular, immigrants from America, Europe, Nigeria, Tunisia, and Georgia showed very low rates of compliance with screening, highlighting the need to target these subgroups with tailored procedures. Immigrants who were employed had lower overall compliance, probably due to the absence of health protection plans and the difficulties in acquiring health permits for these irregular workers. Conversely, knowledge of the Italian language was not significantly associated with compliance. The systematic presence of cultural mediators may have played a key role in determining this outcome, leading the immigrants to fully understand the significance of the screening program and the treatments they underwent, and to freely express questions and concerns [9,19,20].

Our study has a few strengths. First of all, the decision to include in the study the whole population of illegal immigrants should be mentioned. Had we not included eligible immigrants from countries with low TB incidence, we would have excluded about 48% of the subjects and missed the 80% of TB cases. Therefore the model adopted in our centre is innovative because it is based on the presence of a stable team of specialists and a nurse dedicated to the periodic monitoring of compliance and to the recovery of the subjects that don't show up for appointments. Finally, the fact that all the investigations were for free and that there was a direct delivery of the therapy are important elements in ensuring such a high adherence to screening and treatment.

Our study also has limitation. To begin with, the immigrants' self-reported data on education, employment status and symptoms are likely to imply an information bias that may have limited the accuracy of such data. Furthermore, we have not studied the correlation between adherence to screening and characteristics such drinking habits and drug abuse, smoking, or possible congenital or acquired immunodeficiency disorders. Moreover, the small size of the study population led to a lack of statistical precision in our results, which need to be confirmed in larger populations. In addition, in patients with negative TST, QFT was not systematically performed and as reported in recent studies [21], QFT might increase the identification of LTBI cases in recent immigrants. Finally, the findings of this study and the investigated setting may not be easily extended to other contexts, since the model adopted in the study centre was based on the fact that all procedures were for free, and were characterized by a stable health care team. This team also included a nurse dedicated to the periodic monitoring of compliance and to making contact with subjects who failed to show up for appointments.

5. Conclusions

In conclusion, irregular immigrants appear to be a population with a high prevalence of LTBI and active TB and risk of progression from LTBI to active TB, being also characterized by delays in diagnosis and impaired referral to the health services. Consequently, they should be actively targeted with appropriate screening and follow-up procedures, particularly for selected subgroups showing lower compliance with screening and follow-up. Because of the high probability of diagnostic delay and poor accessibility to health services, systematic active TB screening should be implemented in each clinic dealing with seeking care irregular immigrants. In addition, the simultaneous execution of both the TB and LTBI screenings may optimize allocation and use of the resources.

Supplementary Materials: The following are available online at http://www.mdpi.com/1660-4601/16/1/28/s1, Table S1: Demographic characteristics of the study population. Table S2: Association between countries of origin, sex and compliance with screening for active tuberculosis (TB), screening for latent tuberculosis infection (LTBI) and the full protocol (multivariate analysis). Table S3: Association between selected characteristics and compliance with Tuberculin Skin Test (TST) reading (bivariate and multivariate analysis). Table S4: Association between selected characteristics and compliance with chest-X-ray (CXR) execution (bivariate and multivariate analysis).

Table S5: Association between selected characteristics and compliance with Quantiferon-TB gold (QFT) execution (bivariate and multivariate analysis). Table S6: Association between selected characteristics and outcome of Latent Tuberculosis Infection (LTBI) treatment (bivariate analysis).

Author Contributions: Conceptualization, B.F., C.S. and V.M.; Methodology, B.F., C.S., F.R. and V.M.; Software and Formal Analysis, B.F., C.S. and V.M.; Validation, V.M.; Investigation and Resources, B.F., F.R., C.C., G.M., M.L., M.C. and N.F.; Data Curation, B.F., C.S. and V.M.; Writing—original draft preparation, B.F., C.S., and V.M.; Writing—Review and Editing, B.F., C.S., F.R., C.C., M.C., G.M., M.L., N.F., and V.M.; Visualization, B.F., C.S., V.M.; Supervision and Project Administration, B.F. and V.M.; Funding Acquisition, V.M.

Funding: Funding provided by the Local Health Unit of Reggio Emilia.

Acknowledgments: In this section you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Floyd, K.; Glaziou, P.; Zumla, A.; Raviglione, M. The global tuberculosis epidemic and progress in care, prevention, and research: An overview in year 3 of the End TB era. *Lancet Respir. Med.* **2018**, *6*, 299–314. [CrossRef]
2. WHO. *Global Tuberculosis Report 2016*; WHO: Geneva, Switzerland, 2016.
3. Matteelli, A.; Rendon, A.; Tiberi, S.; Al-Abri, S.; Voniatis, C.; Carvalho, A.C.C.; Centis, R.; D'Ambrosio, L.; Visca, D.; Spanevello, A.; et al. Tuberculosis elimination: Where are we now? *Eur. Respir. Rev.* **2018**, *27*. [CrossRef] [PubMed]
4. Lonnroth, K.; Migliori, G.B.; Abubakar, I.; D'Ambrosio, L.; de Vries, G.; Diel, R.; Douglas, P.; Falzon, D.; Gaudreau, M.A.; Goletti, D.; et al. Towards tuberculosis elimination: An action framework for low-incidence countries. *Eur. Respir. J.* **2015**, *45*, 928–952. [CrossRef] [PubMed]
5. Sanchez-Montalva, A.; Salvador, F.; Molina-Morant, D.; Molina, I. Tuberculosis and immigration. *Enfermedades Infecciosas y Microbiologia Clinica* **2018**, *36*, 446–455. [CrossRef] [PubMed]
6. WHO. *The End TB Strategies*; WHO: Geneva, Switzerland, 2014.
7. Kaushik, N.; Lowbridge, C.; Scandurra, G.; Dobler, C.C. Post-migration follow-up programme for migrants at increased risk of developing tuberculosis: A cohort study. *ERJ Open Res.* **2018**, *4*. [CrossRef] [PubMed]
8. IsdseAsesdRE-R. *La Tubercolosi in Italia, Rapporto 2008*; Ministero della Salute: Roma, Italy, 2010.
9. Salute, M.D. *Aggiornamento Delle Raccomandazioni per le Attività di Controllo Della Tubercolosi. Gestione dei Contatti e Della Tubercolosi in Ambito Assistenziale*; Istituto Nazionale Malattie Infettive Lazzaro Spallanzani IRCCS: Roma, Italy, 2010.
10. Salute, M.D. *Aggiornamento Delle Raccomandazioni per le Attività di Controllo Della Tubercolosi: Politiche Efficaci a Contrastare la Tubercolosi Nella Popolazione Immigrata*; Ministero della Salute: Roma, Italy, 2010.
11. CDC. *Targeted Tubercolin Testing and Treatment of Latent Tubercolosis Inferction*; CDC: Atlanta, GA, USA, 2000.
12. WHO. *Definitions and Reporting Framework for Tuberculosis—2013 Revision*; WHO: Geneva, Switzerland, 2014.
13. Morano, J.P.; Walton, M.R.; Zelenev, A.; Bruce, R.D.; Altice, F.L. Latent tuberculosis infection: Screening and treatment in an urban setting. *J. Community Health* **2013**, *38*, 941–950. [CrossRef] [PubMed]
14. Bennett, R.J.; Brodine, S.; Waalen, J.; Moser, K.; Rodwell, T.C. Prevalence and treatment of latent tuberculosis infection among newly arrived refugees in San Diego County, January 2010-October 2012. *Am. J. Public Health* **2014**, *104*, e95–e102. [CrossRef] [PubMed]
15. Bodenmann, P.; Vaucher, P.; Wolff, H.; Favrat, B.; de Tribolet, F.; Masserey, E.; Zellweger, J.P. Screening for latent tuberculosis infection among undocumented immigrants in Swiss healthcare centres; a descriptive exploratory study. *BMC Infect. Dis.* **2009**, *9*, 34. [CrossRef] [PubMed]
16. El-Hamad, I.; Casalini, C.; Matteelli, A.; Casari, S.; Bugiani, M.; Caputo, M.; Bombana, E.; Scolari, C.; Moioli, R.; Scarcella, C.; et al. Screening for tuberculosis and latent tuberculosis infection among undocumented immigrants at an unspecialised health service unit. *Int. J. Tuberc. Lung Dis.* **2001**, *5*, 712–716. [PubMed]
17. Schepisi, M.S.; Gualano, G.; Piselli, P.; Mazza, M.; D'Angelo, D.; Fasciani, F.; Barbieri, A.; Rocca, G.; Gnolfo, F.; Olivani, P.; et al. Active Tuberculosis Case Finding Interventions Among Immigrants, Refugees and Asylum Seekers in Italy. *Infect. Dis. Rep.* **2016**, *8*, 6594. [CrossRef] [PubMed]

18. WHO. *Guidelines on the Management of Latent Tuberculosis Infection*; WHO: Geneva, Switzerland, 2015.
19. Suphanchaimat, R.; Kantamaturapoj, K.; Putthasri, W.; Prakongsai, P. Challenges in the provision of healthcare services for migrants: A systematic review through providers' lens. *BMC Health Serv. Res.* **2015**, *15*, 390. [CrossRef] [PubMed]
20. Cheng, I.H.; Drillich, A.; Schattner, P. Refugee experiences of general practice in countries of resettlement: A literature review. *Br. J. Gen. Pract.* **2015**, *65*, e171–e176. [CrossRef] [PubMed]
21. Saracino, A.; Scotto, G.; Fornabaio, C.; Martinelli, D.; Faleo, G.; Cibelli, D.; Tartaglia, A.; Di Tullio, R.; Fazio, V.; Prato, R.; et al. QuantiFERON-TB Gold In-Tube test (QFT-GIT) for the screening of latent tuberculosis in recent immigrants to Italy. *New Microbiol.* **2009**, *32*, 369–376. [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

Premature Adult Death and Equity Impact of a Reduction of NO_2, PM_{10}, and $PM_{2.5}$ Levels in Paris—A Health Impact Assessment Study Conducted at the Census Block Level

Wahida Kihal-Talantikite [1,*], Pierre Legendre [2,†], Pauline Le Nouveau [2] and Séverine Deguen [2,3,†]

1 LIVE UMR 7362 CNRS (Laboratoire Image Ville Environnement), University of Strasbourg, 67000 Strasbourg, France
2 School of Public Health (EHESP), 35043 Rennes CEDEX, France; legendrepierre@LIVE.fr (P.L.); pauline.le-nouveau@eleve.ensai.fr (P.L.N.); severine.deguen@ehesp.fr (S.D.)
3 Sorbonne Universités, UPMC Univ Paris 06, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique (IPLESP UMRS 1136), 75646 Paris, France
* Correspondence: wahida.kihal@live-cnrs.unistra.fr; Tel.: +33-368-850-977
† These authors contributed equally to this work.

Received: 20 November 2018; Accepted: 21 December 2018; Published: 24 December 2018

Abstract: *Background*: To support environmental policies aiming to tackle air pollution, quantitative health impact assessments (HIAs) stand out as one of the best decision-making tools. However, no risk assessment studies have quantified or mapped the health and equity impact of air pollution reduction at a small spatial scale. *Objectives*: We developed a small-area analysis of the impact of air pollution on "premature" death among an adult population over 30 years of age to quantify and map the health and equity impact related to a reduction of air pollution. *Methods*: All-cause mortality data of an adult population (>30 years) from January 2004 to December 2009 were geocoded at the residential census block level in Paris. Each census block was assigned socioeconomic deprivation levels and annual average ambient concentrations of NO_2, PM_{10}, and $PM_{2.5}$. HIAs were used to estimate, at a small-area level, the number of "premature" deaths associated with a hypothetical reduction of NO_2, PM_{10}, and $PM_{2.5}$ exposure. In total, considering global dose response function for the three pollutants and socioeconomic deprivation specific dose response function, nine HIAs were performed for NO_2 and six and four HIAs for PM10 and PM2.5, respectively. Finally, a clustering approach was used to quantify how the number of "premature" deaths could vary according to deprivation level. *Results*: The number of deaths attributable to NO_2, PM_{10}, and $PM_{2.5}$ exposure were equal to 4301, 3209, and 2662 deaths, respectively. The most deprived census blocks always appeared as one of the groups most impacted by air pollution. Our findings showed that "premature" deaths attributable to NO_2 were not randomly distributed over the study area, with a cluster of excess "premature" deaths located in the northeastern area of Paris. *Discussion*: This study showed the importance of stratifying an environmental burden of disease study on the socioeconomic level, in order to take into consideration the modifier effect of socioeconomic status on the air pollution-mortality relationship. In addition, we demonstrated the value of spatial analysis to guide decision-making. This shows the need for tools to support priority-setting and to guide policymakers in their choice of environmental initiatives that would maximize health gains and reduce social inequalities in health.

Keywords: health impact assessments; premature death; equity impact; health impact; reduction of air pollution; environmental inequalities; spatial analysis; small area; AirQ

1. Introduction

Despite considerable improvement in prevention, management, and regulation, air pollution remains a leading environmental health issue worldwide. From a recent air quality model, the World Health Organization (WHO) estimates that 92% of the global population lives in places where air quality levels exceed WHO limits [1]. Air pollution has been identified as a health priority in the sustainable development agenda. Clean air is one of the fundamental requirements for human health and well-being [2].

While the increased risk of air pollution to health is relatively low compared to other risk factors, the total number of people affected is significant. According to the Organization for Economic Cooperation and Development [3], air pollution is known to be the main environmental cause of "premature" death. In 2012, WHO estimated from Global Health Observatory data that ambient air pollution contributed to 5.4% of all deaths worldwide [4]. However, while most studies have focused on estimating a relationship between pollution and health, less attention has been given to the differential health effects of air pollution according to the socioeconomic status, measured at individual and/or neighborhood levels [5,6]. Identifying population subgroups that are the most vulnerable to the effects of air pollution remains a public health research concern. Recent studies have suggested that several contextual or individual characteristics (such as gender and socioeconomic position, for example) could modify the association between exposure and mortality. Chen et al. in 2005 [7] found a significant increase of coronary death risk with $PM_{2.5}$ exposure in women only, while Deguen et al. in 2015 [5] revealed a stronger association between short term variations of NO_2 concentrations and all-cause mortality for subjects living in areas with low socioeconomic status.

Today, to support environmental policies aiming to tackle air pollution, quantitative health impact assessments (HIAs) stand out as one of the best decision-making tools, because they provide valuable information regarding the future health effects of a potential plan or policy. HIAs are already routinely used by the U.S. Environmental Protection Agency [8] in order to revise national ambient air quality standards. For instance, an increase in life expectancy of 0.61 years associated with a reduction of 10 $\mu g/m^3$ in PM was estimated in the U.S. by Pope et al. in 2009 [9].

A study conducted in the Lausanne-Morges [10] urban area in Switzerland quantified the reduction in "premature" deaths due to air pollution reduction over a period of 10 years, and estimated a decrease of 1% to 2% of total all-cause annual deaths. In two French areas (the Grenoble and Lyon areas) [11], a recent study estimated at census block level that about 3–8% of deaths and 3–10% of lung cancer cases were attributable to $PM_{2.5}$ exposure [11]. An HIA was also recently used to evaluate the health and economic impacts of a potential public transportation modification in terms of proposed fare increases and service cuts conducted in the U.S. state of Massachusetts [12]. To our knowledge, only a few epidemiological studies have investigated the health impact of reducing air pollution according to socioeconomic deprivation measured at a small spatial scale [13,14], ignoring within-city variations of air pollutants. In addition, in order to build efficient policies, it is crucial to establish a full and detailed socioeconomic and health-related assessment at the local scale and identify the categories of citizens who have multiple risk factors. However, no risk assessment studies have quantified or mapped the health impact of air pollution reduction at a small spatial scale to develop targeted policies, and more specifically, environmental policies. This study attempts to remedy this by developing a novel small-area approach combining an HIA and the clustering approach to map the health impact by socioeconomic deprivation level, and to investigate the equity impact of a reduction of ambient NO_2, PM_{10}, and $PM_{2.5}$ concentrations.

In this context, this study has two objectives. First, we will estimate the number of "premature" deaths among an adult population older than 30 years associated with a reduction of NO_2, PM_{10}, and $PM_{2.5}$ concentrations at the census block level in Paris, based on the counterfactual method [11]. Second, we will investigate the spatial distribution of the estimates number of "premature" deaths using a clustering approach to quantify how the number of "premature" deaths could vary according to neighborhood socioeconomic deprivation status measured at census block level.

2. Materiel and Methods

2.1. Study Area

The study area is the city of Paris (the capital of France). The population is about 2,250,000 inhabitants and about 1,360,000 inhabitants are over 30 years old. Paris is subdivided into 992 census blocks with a mean population of about 2199 inhabitants and a mean area of 0.11 km^2.

2.2. Health Data

All-cause mortality data from January 2004 to December 2009 were considered and geocoded at the residential census block level in our study. The data were provided by the death registry of Paris. For confidentiality reasons, it was not possible to distinguish causes of mortality. According to the French demographic institute [15], the mortality rate is very low during childhood, then increases exponentially from age 30. In addition, causes of death for the population less than 30 years old are recognized to be mostly road injuries, domestic injuries, and suicide. For these reasons, and also because it was not possible to obtain the causes of death for reasons of confidentiality, we decided to exclude all the deaths of people aged under than 30 years [16]. The census block of residence was available for each case. In order to estimate death rate, we obtained the population size from the French National Census Bureau (INSEE: http://www.insee.fr). Ethical approval was obtained from the French commission on data privacy and public liberties (CNIL—Commission Nationale de l'Informatique et des Libertés, N 914118).

2.3. Air Pollution

Annual average ambient concentrations of NO_2, PM_{10}, and $PM_{2.5}$ were modeled at census block level by the local air quality monitoring networks, corresponding to the Ile de France region for two different periods: from January 2004 to December 2009 for NO_2 and from January 2007 to December 2009 for PM_{10} and $PM_{2.5}$. The ESMERALDA Atmospheric Modeling system was used. This model integrates several data sources: meteorological data, linear emission sources, surface and major point sources, and background pollution measurements.

2.4. Socioeconomic Deprivation Index

To characterize the neighborhood socioeconomic deprivation at the census block level, an index was created in a previous study [17] (more details elsewhere by Lalloué et al. [17]). Briefly, Principal Component Analysis (PCA) was used to select 15 variables out of 41 initial socioeconomic and demographic variables provided by the 2006 national census at the census block level. Previous ecological studies have demonstrated this index's ability to capture environment-related socio-spatial inequalities in France [6,18,19]. In order to capture the spatial variability of the pollutants, the socioeconomic index was categorized into 10 groups according to the decile of its distribution.

2.5. Health Impact Assessments (HIAs)

HIAs follow a methodology that requires diverse data sources. We combined information related to: (i) size of the population and level of their exposure (population exposure), (ii) the death rate in our study (baseline health rate), and (iii) dose-response function (the relative risk: RR).

The dose-response function was derived from epidemiological studies assessing the relative risk associated with the observed and/or modelled exposure [20]. In this study, the relative risk comes from WHO recommendations; the dose-response function relating all-cause mortality and long term NO_2, $PM_{2.5}$, and PM_{10} concentrations is quantified by RR = 1.041, 95% CI [1.019; 1.064], RR = 1.064, 95% CI [1.043; 1.085] and RR = 1.077, 95% CI [1.068; 1.086], respectively for 10 $\mu g/m^3$ increase in exposure to the pollutant.

In our study, the health effects were evaluated for hypothetical air pollution reductions, according to WHO recommendations. The guideline values identified for each pollutant were 40 μg/m^3 for NO_2, 10 μg/m^3 for $PM_{2.5}$ and 20 μg/m^3 for PM_{10}.

The benefits of the air pollutant reduction scenarios are expressed in terms of attributable number of deaths per year (ΔY) estimated from the following equation:

$$\Delta Y = Y0 \times (1 - e^{-\beta \times \Delta x}) \tag{1}$$

Where:

Y0 is the total number of observed deaths,

Δx is the difference between the yearly observed average of the air pollutant and the reference value (counterfactual), and

β is the natural logarithm of the dose-response function (the relative risk) expressed for a 10 μg/m^3 increase in exposure to the air pollutant (β = ln(RR)/10).

The attributable number of deaths was estimated by AirQ+ software which was developed by the WHO European Centre for Environment and Health (http://www.euro.who.int/en/health-topics/environment-and-health/air-quality/activities/airq-software-tool-for-health-risk-assessment-of-air-pollution).

Tables 1 and 2 present the input data required by the AirQ+ software for our two different periods: 2004 to 2009 for NO_2 (Table 1) and 2007 to 2009 for PM_{10} and $PM_{2.5}$ (Table 2).

To conduct an HIA per socioeconomic deprivation class, we used two studies which investigated the associations between all-cause mortality and long term air pollutant exposure by socioeconomic group: a Dutch study investigated NO_2 and PM_{10} across 5 socioeconomic groups [14] and a Italian study investigated NO_2 and $PM_{2.5}$ across only 3 socioeconomic groups [13] (Tables 3 and 4). Therefore, we estimated the attributable death rates separately for each socioeconomic class based on the 5 dose-response functions of the Dutch study and on the 3 dose-responses functions of the Italian study.

Table 1. Descriptive statistics by socioeconomic deprivation class over the period 2004–2009. It corresponds to the input data of the health impact assessment (HIA) for NO_2 exposure, only.

2004–2009	Population > 30 Years	Death Rate (per 100,000 Inhabitants)	NO_2 (μg/m^3)	Surface (km^2)
Total	8,152,966	1007.20	53.39	105.4
1 (Less deprived)	718,290	1129.91	52.66	9.20
2	724,338	977.86	54.01	9.87
3	781,848	1006.33	54.10	7.54
4	802,248	985.23	53.53	7.14
5	782,280	1005.14	54.02	10.04
6	874,356	1002.57	53.73	6.66
7	872,640	917.90	52.63	7.17
8	862,818	1010.06	50.96	12.37
9	863,826	960.96	51.00	13.62
10 (More deprived)	819,006	1071.05	54.11	14.75

Death rate is the ratio between the total number of observed deaths older than 30 years and the total population older than 30 years. The death rate is expressed per 100,000 inhabitants. NO_2 value corresponds to the mean of the annual average concentrations of the census blocks included in a given socioeconomic deprivation class. NO_2: nitrogen dioxide. More precisely, values of NO_2 are equal to $\frac{\sum_{i=1}^{N}\sum_{j=1}^{T} C_{ij}}{N*T}$, where N is the number of census block, T the number of years over the study period, and C the annual average concentrations of NO_2 of a given census block (i) in a given year (j).

Table 2. Descriptive statistics by socioeconomic deprivation class over the period 2007–2009. It corresponds to the input data of the HIA for PM_{10} and $PM_{2.5}$ exposure.

2007–2009	Population > 30 Years	Death Rate (per 100,000 Population)	PM_{10} (g/m^3)	$PM_{2.5}$ (g/m^3)	Surface (km^2)
Total	4,103,250	992.14	31.07	20.9	105.4
1 (Less deprived)	370,011	1112.67	30.8	20.74	9.20
2	370,800	924.22	31.24	21.01	9.87
3	395,459	979.87	31.34	21.07	7.54
4	407,266	982.4	31.16	21	7.14
5	394,149	987.19	31.31	21.09	10.04
6	442,374	964.57	31.15	20.96	6.66
7	443,739	901.21	30.9	20.83	7.17
8	435,338	986.82	30.29	20.41	12.37
9	442,427	933.94	30.34	20.42	13.62
10 (More deprived)	414,934	1048.12	31.17	20.84	14.75

Death rate is the ratio between the total number of observed deaths older than 30 years and the total population older than 30 years. The death rate is expressed per 100,000 inhabitants. PM_{10} (idem $PM_{2.5}$ value corresponds to the mean of the annual average concentrations of the census blocks included in a given socioeconomic deprivation class. PM_{10}: particulate matter 10 µm or less in diameter. $PM_{2.5}$: particulate matter 2.5 µm or less in diameter. More precisely, values of PM_{10} (and $PM_{2.5}$) are equal to $\frac{\sum_{i=1}^{N}\sum_{j=1}^{T} C_{ij}}{N*T}$, where N is the number of census block, T the number of years over the study period, and C the annual average concentrations of PM_{10} ($PM_{2.5}$) of a given census block (*i*) in a given year (*j*).

Table 3. Associations between long-term NO_2 and PM_{10} exposure and mortality all-causes by socioeconomic class extracted from the Dutch study [14].

Air Pollutants		Socioeconomic Deprivation Class				
		High (1–2)	Moderate/High (3–4)	Medium (5–6)	Moderate/ Low (7–8)	Low (9–10)
NO_2	HR	1.013	1.036	1.022	1.027	1.031
	LL	1.002	1.027	1.014	1.02	1.025
	UL	1.023	1.046	1.03	1.034	1.038
PM_{10}	HR	1.048	1.099	1.052	1.091	1.094
	LL	1.013	1.064	1.024	1.067	1.074
	UL	1.084	1.137	1.08	1.114	1.113

Socioeconomic deprivation class: decile; class 'high' = less deprived versus 'low' = more deprived; HR: Hazard Ratio; LL: lower limit of 95% confidence interval of the hazard ratio; UL: upper limit of 95% confidence interval of the hazard ratio; NO_2: nitrogen dioxide; PM_{10}: particulate matter 10 micrometers or less in diameter.

Table 4. Associations between long-term NO_2 and $PM_{2.5}$ exposure and mortality all-causes by socioeconomic class extracted from the Italian study [13].

Air Pollutants		Socioeconomic Deprivation Class		
		High (1–3)	Medium (4–7)	Low (8–10)
NO_2	HR	1.024	1.016	1.034
	LL	1.012	1.002	1.024
	UL	1.036	1.03	1.045
$PM_{2.5}$	HR	1.04	1.018	1.05
	LL	1.02	0.99	1.03
	UL	1.06	1.04	1.07

Socioeconomic deprivation class: decile; class 'high' = less deprived (decile 1, 2 and 3) versus 'low' = more deprived (decile 8, 9 and 10); HR: Hazard Ratio; LL: lower limit of 95% confidence interval of the hazard ratio; UL: upper limit of 95% confidence interval of the hazard ratio; NO_2: nitrogen dioxide; $PM_{2.5}$: particulate matter 2.5 µm or less in diameter.

2.6. Spatial Analysis

The number of attributable deaths (estimated following the methodology described in Section 2.5) was distributed in each census block, proportionally to the adult population size living in the census block. To investigate the spatial distribution of "premature" deaths at census block level in Paris, we used a spatial scan statistic approach. The Poisson probability model used in the SaTScan software [21] was chosen as a cluster analysis method to detect the presence of high avoidable death spatial clusters (called 'most likely clusters').

The null hypothesis (H0) tested was that the risk is equi-probable throughout the study area. In other words, the expected "premature" death rate would be randomly distributed over the area. The alternative hypothesis (H1) was that there is an elevated risk within the cluster in comparison with census blocks outside the cluster. The procedure works as follows: a circle or window of variable radius (from 0 up to 50% of the population size as recommended by Kulldorf [22]) is placed at every centroid of the census block and moves across the whole study area. For each window, the "premature" death risk estimated in the window is compared with expected "premature" death rate under the hypothesis of a random distribution. The statistically significant most likely clusters are identified using the likelihood ratio test [23]. The *p*-value associated to each detected cluster was obtained from a Monte Carlo replication [24]. ArcGis software was used to map and visualize the spatial location of the statistically significant most likely clusters.

3. Results

3.1. Description of the Study Area and the Population

A total of 82,117 deaths of people over 30 were registered in Paris between January 2004 and December 2009 (and 40,710 deaths between January 2007 and December 2009). The yearly average of NO_2 concentrations was 53.4 μg/m^3 (min = 38.6 μg/m^3; max = 83.13 μg/m^3). The yearly average of PM_{10} concentrations was 31.07μg/m^3 (min = 24.4μg/m^3; max = 43.3 μg/m^3) and 20.9 μg/m^3 (min = 14.9 μg/m^3; max = 28.7 μg/m^3) for $PM_{2.5}$. The spatial distribution of the NO_2 concentrations (Figure 1) shows a gradient from the south of the Seine River, with lower levels (the majority of the census blocks exhibit values less than 51 μg/m^3), to the north, with higher levels (several of the census blocks, in dark colors, show values greater than 55 μg/m^3). The spatial distributions of the annual average level of NO_2, PM_{10} and $PM_{2.5}$ exceeded the WHO threshold presented a similar pattern to the NO_2 spatial distribution (see Appendix A Figures A1–A3). The spatial distribution of the socioeconomic index (Figure 2) revealed a clear gradient from southwest (the least deprived areas), to northeast (the most deprived areas).

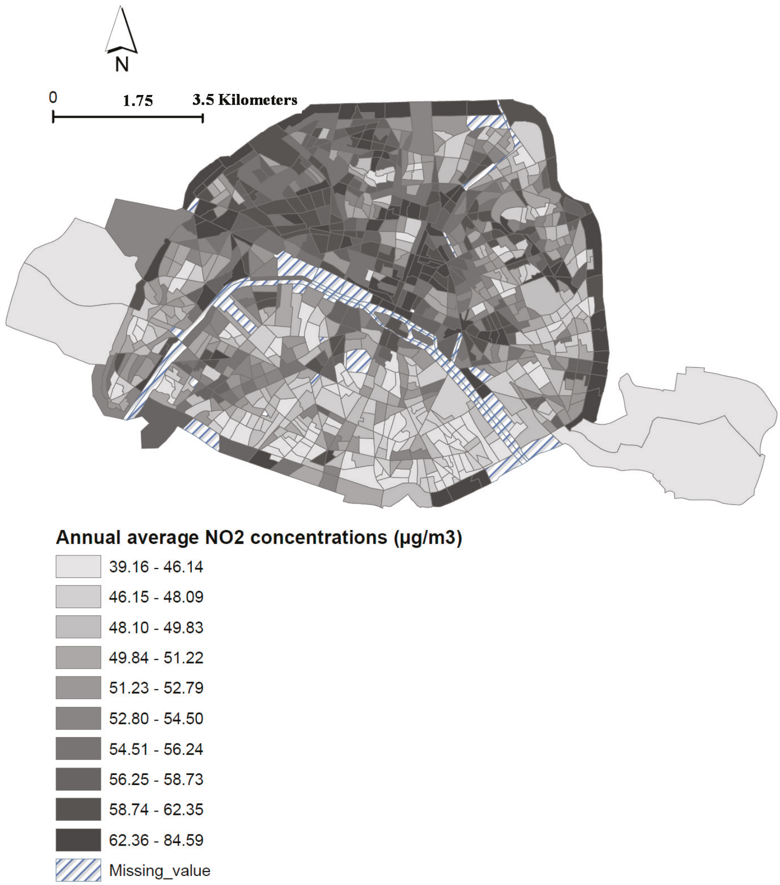


Figure 1. Spatial distribution of the annual average NO_2 concentrations according to the deciles of its distribution, at the census block level, Paris city (Period: 2004–2009).

Figure 2. Spatial distribution of the socioeconomic deprivation index according to the decile of its distribution, at the census block level, Paris city.

3.2. Health Impact

3.2.1. Overall Estimates

Over the period 2004–2009, the number of deaths attributable to NO_2 exposure was 4301 (95% CI [2044; 6545]), which corresponds to about 717 death per year (95% CI [340; 1091]) and to a rate of 52.8 (95% CI [25.1; 80.3]) per 100,000 inhabitants (aged > 30 years). It represents about 5% of total deaths among the adult population over 30 years of age.

For particulate matter, over the period 2007–2009, the number of attributable deaths was equal to 3209 (95% CI [1938; 3355]) and 2,662 (95% CI [2859; 3553]) for PM_{10} and $PM_{2.5}$, respectively. This corresponds to a rate of 78.2 (95% CI [69.7; 86.6]) and 64.9 (95% CI [44.5; 84.4]) per 100,000 inhabitants (about 7.8% and 6.5% of total deaths for PM_{10} and $PM_{2.5}$, respectively).

3.2.2. Estimates by Socioeconomic Deprivation Class

Tables 5 and 6 show the rate of attributable deaths estimated for the three air pollutants by decile of the socioeconomic deprivation index distribution.

Table 5. Rate of attributable deaths per socioeconomic class for two dose-response function.

Socioeconomic Deprivation Classes	Rate of Attributable Deaths (per 100,000 Population) [95% CI]	
	NO_2 Dose-Response Function 1	NO_2 Dose-Response Function 2
Decile 1 (less deprived)	18.3 [2.8; 21.1]	33.4 [16.9; 49.5]
Decile 2	17.5 [2.7; 30.7]	31.9 [16.2; 47.3]
Decile 3	48.9 [37.1; 61.8]	33.1 [16.8; 48.9]
Decile 4	46.0 [34.9; 56.2]	20.9 [2.7; 38.6]
Decile 5	30.2 [19.4; 40.8]	22.1 [2.8; 40.8]
Decile 6	29.5 [18.9; 39.9]	21.6 [2.7; 39.9]
Decile 7	30.4 [22.7; 37.9]	18.1 [2.3; 33.6]
Decile 8	33.2 [26.9; 40.4]	36.3 [25.9; 47.6]
Decile 9	31.7 [25.7; 38.6]	34.7 [25.7; 45.4]
Decile 10 (more deprived)	45.2 [36.7; 54.9]	49.4 [35.2; 64.5]

Dose response function 1 based on Dutch study [14] and dose response function 2 based on Italian study [13]. 95% CI: 95% Confidence Interval.

Table 6. Rate of attributable deaths per socioeconomic class for two dose-response function.

Socioeconomic Deprivation Classes	Rate of Attributable Deaths (per 100 000 Population) [95% CI]	
	PM_{10} Dose-Response Function 1	$PM_{2.5}$ Dose-Response Function 2
Decile 1 (less deprived)	54.9 [15.4; 92.8]	45.9 [23.4; 67.5]
Decile 2	47.4 [13.3; 80.1]	39.1 [19.9; 57.4]
Decile 3	99.5 [66.5; 132.8]	41.6 [21.2; 61.2]
Decile 4	98.2 [65.7; 131.1]	19.1 [0; 41.5]
Decile 5	55 [26.1; 82.3]	19.3 [0; 42]
Decile 6	53 [25.1; 79.3]	18.7 [0; 40.6]
Decile 7	81.6 [61.5; 100]	17.2 [0; 37.5]
Decile 8	84.6 [63.7; 103.7]	48.9 [29.9; 67.1]
Decile 9	82.8 [66.5; 97.9]	46.3 [28.3; 63.6]
Decile 10 (more deprived)	100.1 [80.3; 118.1]	54.0 [33.3; 74.1]

Dose response function 1 based on Dutch study [14] and dose response function 2 based on Italian study [13]. 95% CI: 95% Confidence.

Whatever the pollutant, the most deprived census blocks (decile 10) always appeared as one of the groups most impacted by air pollution. With an annual average of NO_2 equal to 54.11 $\mu g/m^3$ (one of the highest values), the attributable death rates are estimated to be 45.2 and 49.4 per 100,000 inhabitants using the dose-response function of the Dutch and Italian studies, respectively. With an annual average of PM_{10} and $PM_{2.5}$ equal to 31.17 $\mu g/m^3$ and 20.84 $\mu g/m^3$ (in the high range of the annual average of air pollutants), the attributable death rates are estimated to be 100.1 and 54.0 per 100,000 inhabitants, using the dose-response function of the Dutch and Italian studies, respectively.

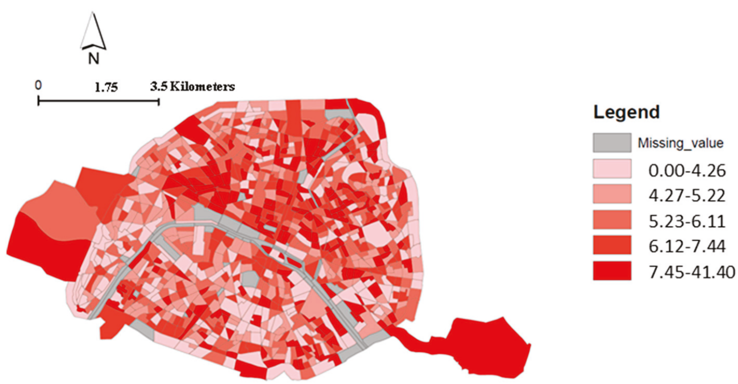
Populations living in less deprived census blocks (decile 3 and 4, in particular) also appear highly impacted by air pollution. These findings are consistent with the increase of the dose-response function and the level of air pollutant exposure in the high range, whatever the pollutant of interest.

3.3. Spatial Distribution

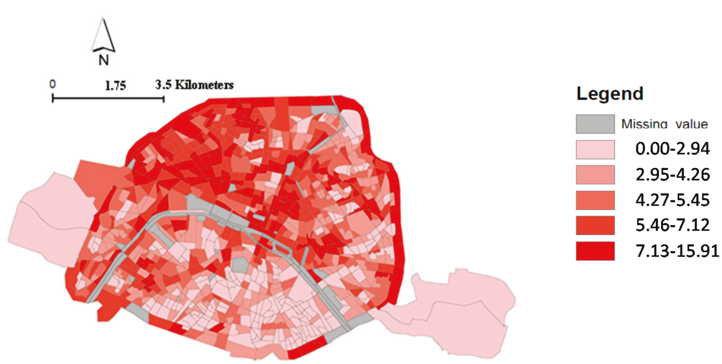
The rate of "premature" deaths per census block (Figure 3) was estimated according to 3 different scenarios: (a) without spatial variability of NO_2 in Paris, (b) with spatial variability of NO_2 between census blocks, and (c) with spatial variability of NO_2 and socio-economic level between census blocks. The rate corresponds to the number of premature death divided by the number of total adult deaths.

Unlike the Figure 3a, the Figure 3b,c reveal a spatial pattern with a higher rate of "premature" deaths attributable to NO_2 located in the north part of Paris in comparison with the south part. A difference also appears between Figure 3b,c: considering the spatial variability of NO_2 combined with the level of socio-economic deprivation (based on Dutch study [14]), the higher rate of "premature" deaths among total adult death shifted in northeastern Paris (Figure 3c).

For particulate matter, the spatial distribution of the rate of the "premature" adult deaths attributable to PM_{10} and $PM_{2.5}$, respectively, among total death, show the same pattern (see Appendix A Figures A4 and A5).

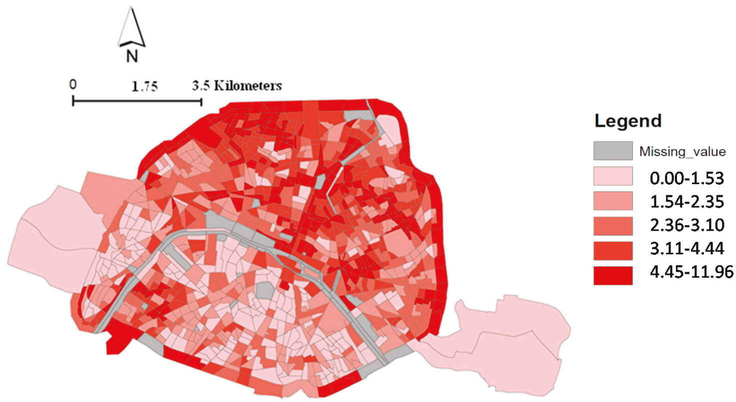


(**a**)



(**b**)

Figure 3. *Cont.*



(**c**)

Figure 3. Spatial distribution of the rate of adults deaths attributable to NO_2 among total death, at the census block level, Paris city; (**a**) without spatial variability of NO_2 exposure in Paris; (**b**) with spatial variability of NO_2 between census blocks; (**c**) with spatial variability of NO_2 and socio-economic level between census blocks (according Dutch study [14]).

The statistical spatial approach confirms that the spatial aggregation of "premature" deaths in the northeast is significant (see Figure 4). This means that "premature" deaths are not randomly distributed across the study area. This most likely cluster comprises an area of 459 census blocks with a risk 1.12 times higher than in the rest of the study area (p-value = 0.029). This cluster hosts a total of 4,038,108 inhabitants and has 3455 "premature" deaths (about 80% of the total number of "premature" deaths estimated in Paris). The spatial approach did not reveal any statistically significant aggregation of "premature" deaths attributable to PM_{10} and to $PM_{2.5}$ (data not shown).

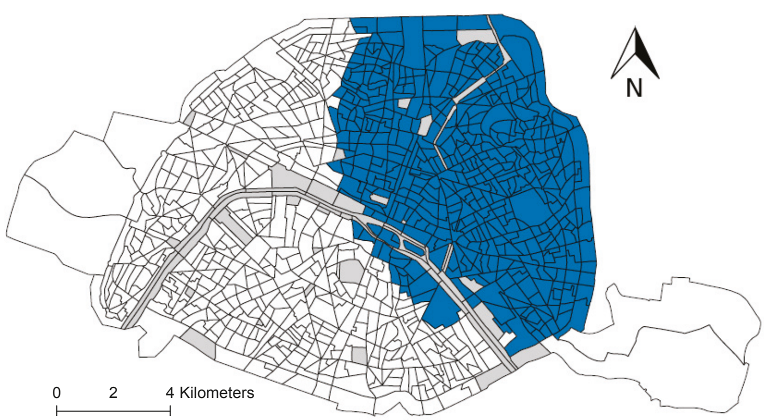


Figure 4. Mapping of the most likely cluster of the number of avoidable deaths. Blue areas identify census block included in the most likely cluster.

4. Discussion

In this study, we developed a small-area analysis of the impact of air pollution on "premature" death to quantify and map the health and equity impact related to a reduction of air pollution. We evaluated the health impact of hypothetical air pollution reductions according to WHO recommendations. This allowed us to estimate at a small-area level the rate of "premature" deaths attributable to NO_2, PM_{10}, and $PM_{2.5}$ taking into account the level of socioeconomic deprivation, and to visualize the spatial distribution of the risk of "premature" deaths.

First, we predicted an overall mortality attributable to long-term NO_2 exposure equal to 4301 deaths (5% of the total deaths registered in Paris over the period 2004 to 2009). Over the shorter period 2007–2009, the number of deaths attributable to PM_{10} and $PM_{2.5}$ were comparatively higher: 3209 and 2662 deaths, which corresponds to about 7.8% and 6.5% of total deaths. This percentage was consistent with the Global Burden of Disease published in 2015 [25], which estimated that about 7.6% of total deaths were attributable to long-term exposure to $PM_{2.5}$.

A recent study conducted in greater Cairo, Egypt estimated that about 11% and 8% of non-accidental mortality (in the population over 30 years old) could be attributed to $PM_{2.5}$ and NO_2, respectively [26]. The higher level of $PM_{2.5}$ concentrations varying between 50 $\mu g/m^3$ and 100 $\mu g/m^3$ in this megacity may partially explain the difference observed with our estimate, the maximum concentrations of $PM_{2.5}$ being equal to 28.7 $\mu g/m^3$ in Paris. In contrast, because the NO_2 concentration was found to be below the 40 $\mu g/m^3$ air quality guideline of WHO, the author used another limit equal to 10 $\mu g/m^3$, according to the recommendation of the Health Risks of Air Pollution in Europe project [27]. While in Paris the annual average NO_2 concentration is higher, the stricter limit used in the Egyptian study may partially explain the difference with our estimate of deaths attributable to NO_2. A study conducted in the Lausanne-Morges urban area of Switzerland estimated the health benefits of a reduction of PM_{10} and NO_2 exposure after implementing a clean air plan [10]. Over a period of 10 years, the reduction of PM_{10} and NO_2 exposure was equal to 3.3 $\mu g/m^3$ and 5.6 $\mu g/m^3$. These air quality improvements reduced total mortality by about 1% to 2%. Applying a similar reduction of PM_{10} and NO_2 exposure in Paris produced comparable estimates of the percentage of "premature" deaths.

Second, our study demonstrated that the burden of mortality varied according to the level of socioeconomic deprivation. Populations living in the most deprived census blocks (those of the decile 10) appear particularly at risk of death related to NO_2 exposure. Indeed, while the level of NO_2 exposure decreases between the decile 5 and 9, population living in the census blocks of the decile 10 (the most deprived) accumulate a high level of exposure and a particular vulnerability to the adverse effect of air pollution. Consequently, for this population group, the two issues (exposure differential and vulnerability differential) may explain the high rate of death due to air pollution. However, it is not easy to draw a general statement about the most probable explanation between exposure differential, vulnerability differential, both because what we observed between socioeconomic level and NO_2 exposure is not as clear with PM_{10} and $PM_{2.5}$ exposure. Maybe, it could be partially explained by the lower spatial variability of PM.

Finally, our study showed that "premature" deaths attributable to NO_2 were not randomly distributed over the study area, with a cluster of excess "premature" deaths located in the northeastern area of Paris.

To our knowledge, our study is the first to stratify an environmental burden of disease by the socioeconomic deprivation level measured at the residential census block level, making it difficult to compare our findings with those of others.

Several limitations of this study should be addressed here.

First, the methodology used to estimate attributable deaths is based on the AirQ+ software, which is based on a reference model developed by WHO. However, one weakness is that it does not take into consideration the effects caused by exposure to several pollutants in combination or their synergistic effects. In our study, as in the majority of scientific literature, the effects of pollutants are investigated individually, which could bias our estimates.

Secondly, the exposure level attributed to the population was approximated by the annual average ambient concentrations of the pollutants estimated at the place of residence provided at the date of death. This is a common limitation of numerous epidemiological studies which investigate the health impact of long-term exposure to air pollution, ignoring temporal and spatial variability due to mobility of the population and it could lead to a misclassification of the exposure. A conceptual model has been recently proposed aiming to assess cumulative exposure to air pollution at a fine scale and applied in Paris at the census block level [28]. The findings revealed that the level of population exposure to NO_2 decreased when including the population mobility within the census block. However, the decrease was lower for the arrondissements located in northeastern Paris where the level of socioeconomic deprivation is the highest. This finding further supports the hypothesis of differential exposure.

Third, the socioeconomic deprivation status was estimated at the census block level rather than the individual level. However, census blocks are defined to maximize their uniformity in terms of population size, socioeconomic and demographic characteristics, land use, and zoning, thus reducing the risk of ecological bias.

Finally, the major limitation of our paper is the lack of studies that stratify their analysis based on socioeconomic deprivation status. Indeed, to produce a robust dose-response function per socioeconomic deprivation class, a meta-analysis is recommended. However, only two studies conducted in areas comparable to Paris were identified in the literature. Using the dose (air pollution)-response (mortality) function (relative risk) of these studies, our findings revealed that the number of "premature" deaths varied according to the socioeconomic deprivation level measured at the place of residence. This reflects not only the different dose-response functions used, but also the level of air pollution exposure and the population density. However, our findings tend to show a higher impact of air pollution exposure among the more deprived areas.

Benefits of this Research for Public Health

This study provides answers to socioeconomic and environmental inequalities highlighted as an important public health issue by WHO. The research that formed the basis of public health policy provides little evidence for effective initiatives aiming to improve population health and tackle environmental and social inequalities in health. This paper is an attempt to fill the gap regarding the need for the development of powerful tools to support priority-setting and guide policymakers in their choice of environmental policies.

In this context, this study produced crucial information for policymakers to prioritize actions to investigate social health inequalities:

- Quantification of the number of "premature" deaths attributable to a reduction of NO_2, PM_{10}, and $PM_{2.5}$ stratified by residential socioeconomic deprivation status.
- Spatial distribution of health and equity impacts of reducing these three pollutants.

In addition, this study illustrates the value of socio-spatial analysis implemented at a small spatial scale to pinpoint the areas where action is needed. In our study, for instance, we identified that an action conducted in northeastern Paris would be highly effective, since this area accounts for about 80% of the total number of "premature" deaths estimated.

At middle- and long-term, it could be really useful to perform the same study again with recent health and air pollution data, in order to investigate if the spatial distribution of the premature death changes over time, or if despite of the decrease of air pollution, cluster counting of a higher number of premature deaths related to air pollution is located in the same place.

5. Conclusions

This study showed the importance of stratifying an environmental burden of disease study on the socioeconomic level in order to take into consideration the modifier effect of socioeconomic status on the air pollution-mortality relationship. In addition, we demonstrated the value of spatial analysis to guide decision-making. Indeed, given today's budgetary constraints, it can be quite challenging for policymakers to select an initiative. This shows the need for tools to support priority-setting and to guide policymakers in their choice of environmental initiatives that would maximize health gains and reduce social inequalities in health.

Author Contributions: Conceptualization, W.K.-T. and S.D.; Formal analysis, W.K.-T., P.L. and S.D.; Investigation, W.K.-T., P.L., P.L.N. and S.D.; Methodology, W.K.-T., P.L., P.L.N. and S.D.; Supervision, W.K.-T. and S.D.; Validation, W.K.-T. and S.D.; Writing—original draft, W.K.-T., P.L., P.L.N. and S.D.; Writing—review & editing, W.K.-T. and S.D.

Funding: This research received no external funding.

Acknowledgments: The authors thank all scientific researchers of the Equit' Area project and the personnel of the local association in air quality of Paris Metropolitan Area. This study was supported by Foundation de France and Interdisciplinary research program on health crises and health protection of University of Sorbonne Paris City (U.S.PC).

Conflicts of Interest: The authors declare they have no actual or potential competing financial interests.

Appendix A.

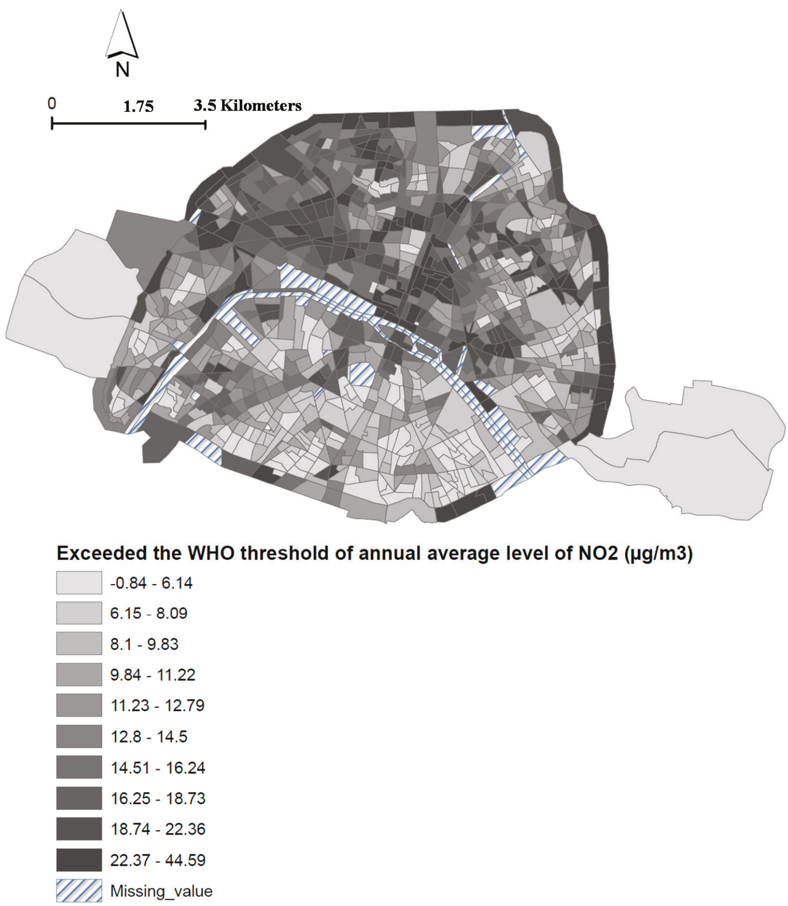


Figure A1. Spatial distribution of the exceeded the WHO threshold of annual average level of NO_2 from 2000–2009 (the difference between level of NO_2 exposure in each census block and WHO guidelines (40 µg/m^3)).

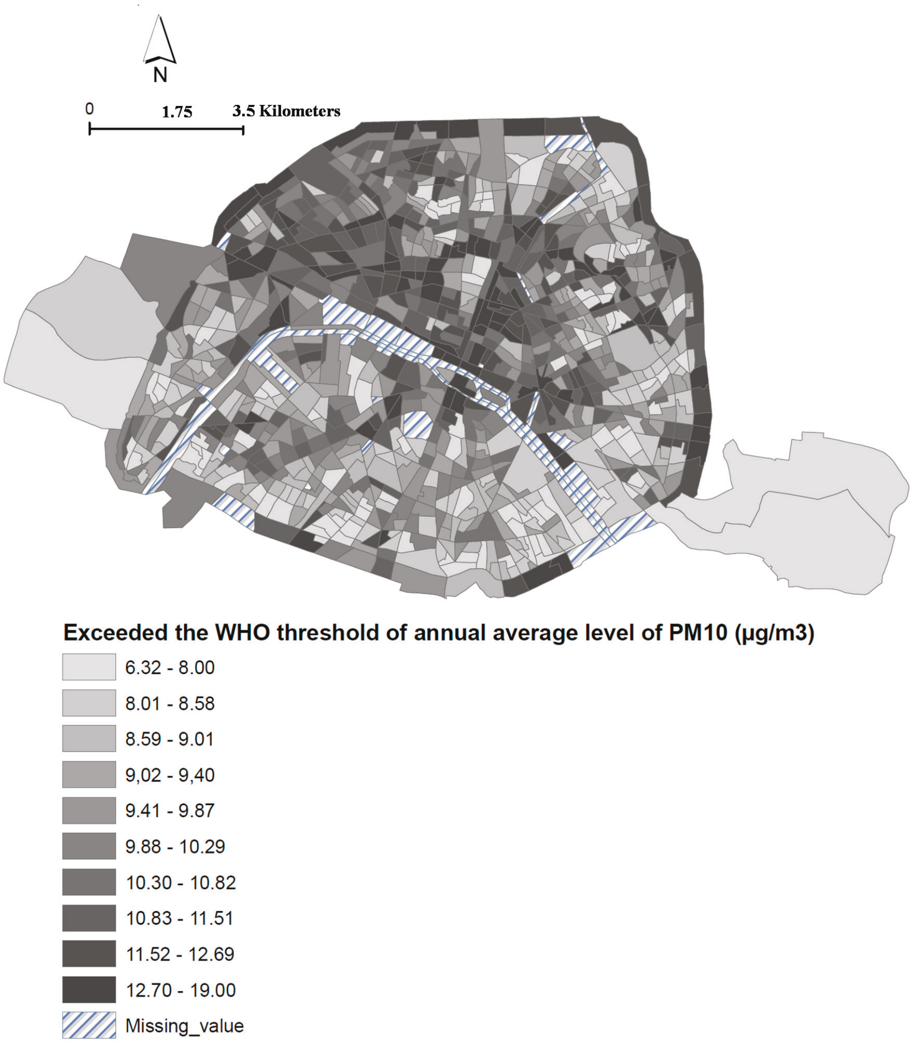


Figure A2. Spatial distribution of the exceeded the WHO threshold of annual average level of PM_{10} from 2000–2009 (the difference between level of PM_{10} exposure in each census block and WHO guidelines (20 µg/m^3)).

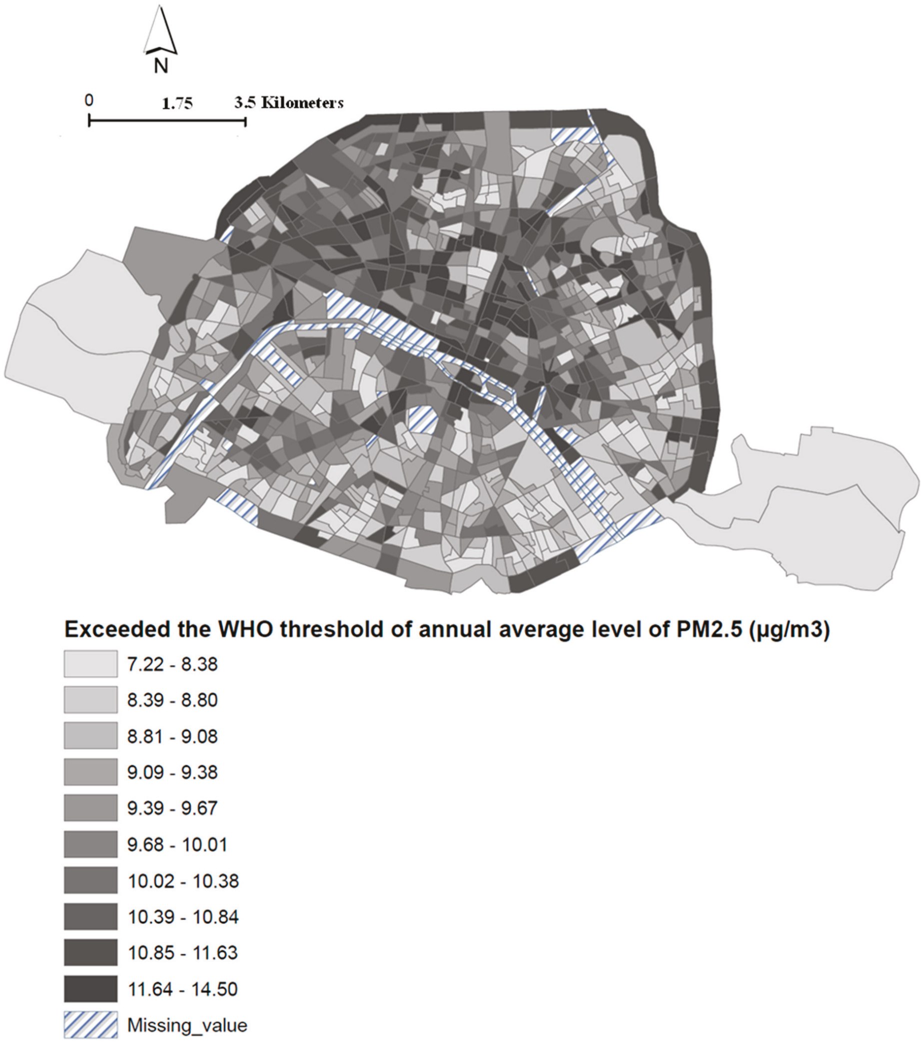
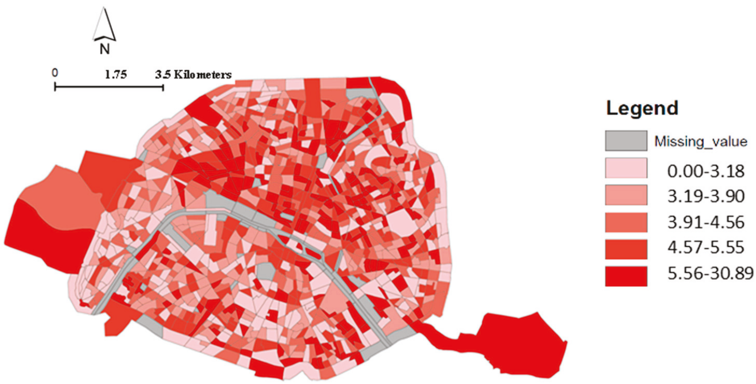
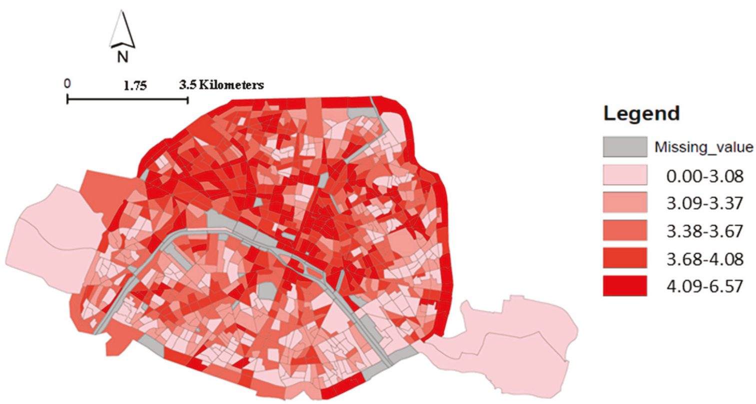


Figure A3. Spatial distribution of the exceeded the WHO threshold of annual average level of $PM_{2.5}$ from 2000–2009 (the difference between level of $PM_{2.5}$ exposure in each census block and WHO guidelines (10 $\mu g/m^3$)).



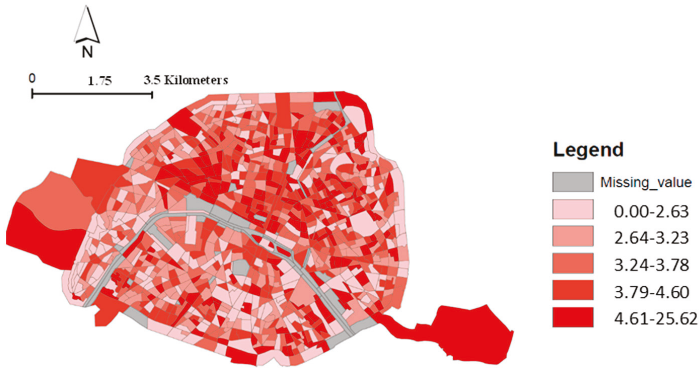
(**a**)



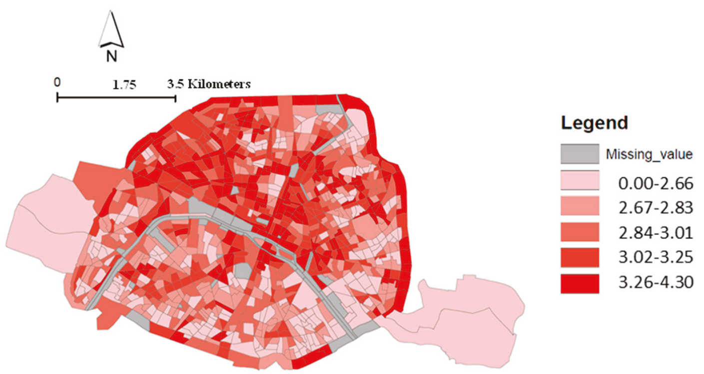
(**b**)

(**c**)

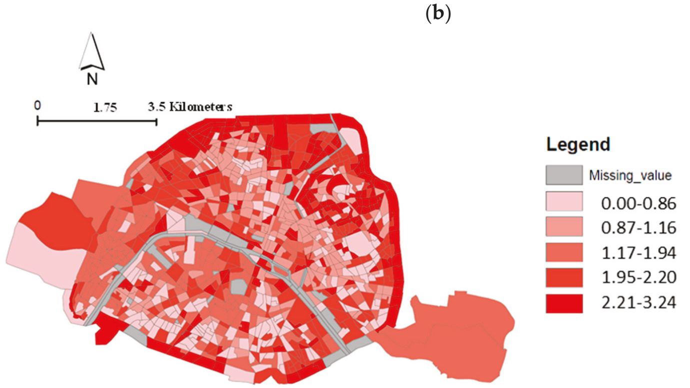
Figure A4. Spatial distribution of the rate of deaths attributable to PM_{10} among total death, at the census block level, Paris City; (**a**) without spatial variability of PM_{10} exposure in Paris; (**b**) with spatial variability of PM_{10} between census block; (**c**) with spatial variability of PM_{10} and socio-economic level between census block.



(**a**)



(**b**)



(**c**)

Figure A5. Spatial distribution of the rate of deaths attributable to $PM_{2.5}$ among total death, at the census block level, Paris City; (**a**) without spatial variability of $PM_{2.5}$ exposure in Paris; (**b**) with spatial variability of $PM_{2.5}$ between census block; (**c**) tacking account the spatial variability of $PM_{2.5}$ and socio-economic level between census block.

References

1. World Health Organization. *WHO Releases Country Estimates on Air Pollution Exposure and Health Impact*; WHO: Geneva, Switzerland, 2016. Available online: http://www.who.int/news-room/detail/27-09-2016-who-releases-country-estimates-on-air-pollution-exposure-and-health-impact (accessed on 23 May 2018).
2. Guillerm, N.; Cesari, G. Fighting ambient air pollution and its impact on health: From human rights to the right to a clean environment. *Int. J. Tubercul. Lung Dis.* **2015**, *19*, 887–897. [CrossRef] [PubMed]
3. OECD. Organisation for Economic Co-Operation and Development—The Cost of Air Pollution. Health Impacts of Road Transport. 2014. Available online: http://www.keepeek.com/Digital-Asset-Management/oecd/environment/the-cost-of-air-pollution_9789264210448-en#page1 (accessed on 9 February 2015).
4. WHO. *WHO Mortality and Burden of Disease from Ambient Air Pollution*; WHO: Geneva, Switzerland, 2012. Available online: http://www.who.int/gho/phe/outdoor_air_pollution/burden/en/ (accessed on 23 May 2018).
5. Deguen, S.; Petit, C.; Delbarre, A.; Kihal, W.; Padilla, C.; Benmarhnia, T.; Lapostolle, A.; Chauvin, P.; Zmirou-Navier, D. Neighbourhood Characteristics and Long-Term Air Pollution Levels Modify the Association between the Short-Term Nitrogen Dioxide Concentrations and All-Cause Mortality in Paris. *PLoS ONE* **2015**, *10*, e0131463. [CrossRef] [PubMed]
6. Padilla, C.M.; Deguen, S.; Lalloue, B.; Blanchard, O.; Beaugard, C.; Troude, F.; Navier, D.Z.; Vieira, V.M. Cluster analysis of social and environment inequalities of infant mortality. A spatial study in small areas revealed by local disease mapping in France. *Sci. Total Environ.* **2013**, *454–455*, 433–441. [CrossRef] [PubMed]
7. Chen, L.H.; Knutsen, S.F.; Shavlik, D.; Beeson, W.L.; Petersen, F.; Ghamsary, M.; Abbey, D. The association between fatal coronary heart disease and ambient particulate air pollution: Are females at greater risk? *Environ. Health Perspect.* **2005**, *113*, 1723–1729. [CrossRef] [PubMed]
8. U.S. Environmental Protection Agency. *Final Regulatory Impact Analysis (RIA) for the NO_2 National Ambient Air Quality Standards (NAAQS)*; Office of Air Quality Planning and Standards: Washington, DC, USA, 2010.
9. Pope, C.A.; Ezzati, M.; Dockery, D.W. Fine-particulate air pollution and life expectancy in the United States. *N. Engl. J. Med.* **2009**, *360*, 376–386. [CrossRef] [PubMed]
10. Castro, A.; Künzli, N.; Götschi, T. Health benefits of a reduction of PM10 and NO_2 exposure after implementing a clean air plan in the Agglomeration Lausanne-Morges. *Int. J. Hyg. Environ. Health* **2017**, *220*, 829–839. [CrossRef] [PubMed]
11. Morelli, X.; Rieux, C.; Cyrys, J.; Forsberg, B.; Slama, R. Air pollution, health and social deprivation: A fine-scale risk assessment. *Environ. Res.* **2016**, *147*, 59–70. [CrossRef] [PubMed]
12. James, P.; Ito, K.; Buonocore, J.J.; Levy, J.I.; Arcaya, M.C. A health impact assessment of proposed public transportation service cuts and fare increases in Boston, Massachusetts (USA). *Int. J. Environ. Res. Public Health* **2014**, *11*, 8010–8024. [CrossRef] [PubMed]
13. Cesaroni, G.; Badaloni, C.; Gariazzo, C.; Stafoggia, M.; Sozzi, R.; Davoli, M.; Forastiere, F. Long-term exposure to urban air pollution and mortality in a cohort of more than a million adults in Rome. *Environ. Health Perspect.* **2013**, *121*, 324–331. [CrossRef] [PubMed]
14. Fischer, P.H.; Marra, M.; Ameling, C.B.; Hoek, G.; Beelen, R.; de Hoogh, K.; Breugelmans, O.; Kruize, H.; Janssen, N.A.; Houthuijs, D. Air Pollution and Mortality in Seven Million Adults: The Dutch Environmental Longitudinal Study (DUELS). *Environ. Health Perspect.* **2015**, *123*, 697–704. [CrossRef] [PubMed]
15. Ined—Inst. Natl. D'études Démographiques. 2004. Available online: https://www.ined.fr/en/everything_about_population/graphs-maps/interpreted-graphs/age-risk-mortality/ (accessed on 13 December 2018).
16. INSEE. 2016. Available online: https://www.insee.fr/fr/statistiques/3134763 (accessed on 13 December 2018).
17. Lalloué, B.; Monnez, J.-M.; Padilla, C.; Kihal, W.; Le Meur, N.; Zmirou-Navier, D.; Deguen, S. A statistical procedure to create a neighborhood socioeconomic index for health inequalities analysis. *Int. J. Equity Health* **2013**, *12*, 21. [CrossRef] [PubMed]
18. Kihal-Talantikite, W.; Padilla, C.M.; Lalloué, B.; Gelormini, M.; Zmirou-Navier, D.; Deguen, S. Green space, social inequalities and neonatal mortality in France. *BMC Pregnancy Childbirth* **2013**, *13*, 191. [CrossRef] [PubMed]

19. Kihal-Talantikite, W.; Padilla, C.M.; Lalloue, B.; Rougier, C.; Defrance, J.; Zmirou-Navier, D.; Deguen, S. An exploratory spatial analysis to assess the relationship between deprivation, noise and infant mortality: An ecological study. *Environ. Health Glob. Access. Sci. Source* **2013**, *12*, 109. [CrossRef] [PubMed]
20. Ostro, B. *Outdoor Air Pollution—Assessing the Environmental Burden of Disease at National and Local Levels*; World Health Organization: Geneva, Switzerland, 2004; Available online: http://www.who.int/quantifying_ehimpacts/national/en/ (accessed on 23 December 2018).
21. Kulldorff, M. Information Management Services, Inc. SaTScan: Software for the Spatial, Temporal, and Space-Time Scan Statistics, Version 6.0. 2005. Available online: http://www.satscan.org/ (accessed on 8 October 2018).
22. Kulldorff, M. A spatial scan statistic. *Commun. Stat. Theory Methods* **1997**, *26*, 1481–1496. [CrossRef]
23. Kulldorff, M.; Nagarwalla, N. Spatial disease clusters: Detection and inference. *Stat. Med.* **1995**, *14*, 799–810. [CrossRef] [PubMed]
24. Dwass, M. Modified Randomization Tests for Nonparametric Hypotheses. *Ann. Math. Stat.* **1957**, *28*, 181–187. [CrossRef]
25. GBD 2015 LRI Collaborators. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory tract infections in 195 countries: A systematic analysis for the Global Burden of Disease Study 2015. *Lancet. Infect. Dis.* **2017**, *17*, 1133–1161. [CrossRef]
26. Wheida, A.; Nasser, A.; El Nazer, M.; Borbon, A.; Abo El Ata, G.A.; Abdel Wahab, M.; Alfaro, S.C. Tackling the mortality from long-term exposure to outdoor air pollution in megacities: Lessons from the Greater Cairo case study. *Environ. Res.* **2018**, *160*, 223–231. [CrossRef] [PubMed]
27. Hensche, S.; Chan, G. Health Risks of Air Pollution in Europe—HRAPIE Project New Emerging Risks to Health from Air Pollution—Results from the Survey of Experts. 2013. Available online: http://www.euro.who.int/__data/assets/pdf_file/0017/234026/e96933.pdf (accessed on 8 October 2018).
28. Wahida, K.-T.; Padilla, C.M.; Denis, Z.-N.; Olivier, B.; Géraldine, L.N.; Philippe, Q.; Séverine, D. A Conceptual Framework for the Assessment of Cumulative Exposure to Air Pollution at a Fine Spatial Scale. *Int. J. Environ. Res. Public Health* **2016**, *13*, 319. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

Beyond Commuting: Ignoring Individuals' Activity-Travel Patterns May Lead to Inaccurate Assessments of Their Exposure to Traffic Congestion

Junghwan Kim and Mei-Po Kwan *

Department of Geography and Geographic Information Science, Natural History Building, 1301 W Green Street, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA; jk11@illinois.edu
* Correspondence: mkwan@illinois.edu; Tel.: +1-217-244-4853

Received: 3 December 2018; Accepted: 28 December 2018; Published: 30 December 2018

Abstract: This research examines whether individual exposures to traffic congestion are significantly different between assessments obtained with and without considering individuals' activity-travel patterns in addition to commuting trips. We used crowdsourced real-time traffic congestion data and the activity-travel data of 250 individuals in Los Angeles to compare these two assessments of individual exposures to traffic congestion. The results revealed that individual exposures to traffic congestion are significantly underestimated when their activity-travel patterns are ignored, which has been postulated as a manifestation of the uncertain geographic context problem (UGCoP). The results also highlighted that the probability distribution function of exposures is heavily skewed but tends to converge to its average when individuals' activity-travel patterns are considered when compared to one obtained when those patterns are not considered, which indicates the existence of the neighborhood effect averaging problem (NEAP). Lastly, space-time visualizations of individual exposures illustrated that people's exposures to traffic congestion vary significantly even if they live at the same residential location due to their idiosyncratic activity-travel patterns. The results corroborate the claims in previous studies that using data aggregated over areas (e.g., census tracts) or focusing only on commuting trips (and thus ignoring individuals' activity-travel patterns) may lead to erroneous assessments of individual exposures to traffic congestion or other environmental influences.

Keywords: traffic congestion; activity-travel patterns; real-time traffic data; the uncertain geographic context problem (UGCoP); the neighborhood effect averaging problem (NEAP)

1. Introduction

Traffic congestion has long been a serious transportation-related issue that people confront in their daily life in the U.S. [1–3]. Drivers in the U.S. wasted 7 billion hours on roads in 2015 due to delays caused by traffic congestion [4]. Moreover, it is expected that traffic congestion will intensify in the future as more people will move to urban areas [5]. Transportation and health researchers have thus considered traffic congestion exposure as a critical factor that influences individuals' physical and mental health. For example, a number of studies revealed that higher exposure to traffic congestion may be associated with escalated heart rate and blood pressure [6,7], heightened urinary catecholamine (a stress-related hormone) [8], and negative health outcomes [9–11]. In addition to these physical tolls, studies have also shown that exposures to traffic congestion may be linked to psychological stress [12–17] and reduced well-being [18,19]. Furthermore, some studies have argued that longer commuting time, which is worsened by traffic congestion, may harm people's work-family balance [20] or even increase the likelihood of being involved in domestic violence [21].

To accurately assess the effects of traffic congestion exposures on human health, it is important to accurately measure individual exposures to traffic congestion. Specifically, in terms of measuring traffic congestion exposures, most previous studies used area-based aggregate data (e.g., [9,21,22]) or focused only on commuting trips (and thus disregarded non-commuting trips) (e.g., [11,18,19]). We argue that these approaches in previous research may lead to erroneous assessments of individual exposures to traffic congestion, which may in turn lead to erroneous evaluations of the health impacts of traffic congestion because they did not consider individuals' unique activity-travel patterns.

To address the limitations of previous studies, this research examines whether individual exposures to traffic congestion are significantly different between assessments obtained with and without considering individuals' activity-travel patterns in addition to commuting trips. We used crowdsourced real-time traffic congestion data and the activity-travel data of 250 individuals in Los Angeles to compare these two assessments of individual exposures to traffic congestion. The results indicate that ignoring individuals' activity-travel patterns may lead to inaccurate assessments of their exposures to traffic congestion.

2. Limitations of Previous Studies

Previous studies on traffic congestion have several limitations, some of which this study seeks to address. First, existing research used data aggregated over areas such as metropolitan areas or zip (postal) code areas. For example, Levy et al. [9] assessed the health impact of $PM_{2.5}$ exposure associated with traffic congestion on mortality and monetized the value of mortality risk in 83 metropolitan areas in the U.S. By using zip code areas in the Los Angeles Metropolitan Area, Beland and Brent [21] noted that traffic congestion may lead to a higher risk of being involved in domestic violence. Brauer et al. [22] examined people's traffic-related air pollution "exposure zones" (defined as a 500-m buffer zone from highways or a 100-m buffer zone from major arterials) and observed that 32% of people in Canada live in exposure zones.

Although these studies provide useful insights into the health impacts of traffic congestion, their estimations of traffic congestion exposure may be erroneous. Part of the reason for the error is because the units of analysis are areal units for which individual data are aggregated (e.g., buffer zones, zip code areas, or metropolitan areas); and the modifiable areal unit problem (MAUP) may contribute to some of such error. This means that previous studies presumed that individuals in the same areal unit are exposed to the same level of traffic congestion when estimating traffic congestion exposure. However, this assumption is problematic because each individual may have distinctive activity-travel patterns and thus may have different levels of exposure to traffic congestion and experience different health impacts [23,24]. In other words, since individuals have idiosyncratic activity-travel patterns, different individuals may be exposed to different levels of traffic congestion in complex and unique ways even when they live in the same area.

Second, previous research focused only on commuting trips while ignoring other types of trips, where individuals may also be exposed to significant traffic congestion. For instance, Olsson et al. [25] found that commuting satisfaction may affect overall happiness based on a survey of 713 commuters. Using a 23-year longitudinal dataset of 2736 commuters, Sandow et al. [11] showed that there may be gender differences in mortality risks due to longer commuting. Despite the meaningful results of these previous studies, focusing exclusively on commuting trips while not considering other components of individuals' activity-travel patterns may lead to erroneous estimations of their exposure to traffic congestion for the following two reasons.

First, commuting trips account for only a small proportion of individuals' total travel. Although commuting trips seem to constitute the most significant portion of our trips at first glance, almost 70% of trips in the U.S. consists of non-work trips according to the 2009 National Household Travel Survey [26]. Thus, considering only commuting trips may result in erroneous estimations of traffic congestion exposure because people may also experience traffic congestion when undertaking other types of trips, which also comprise their travel and are overlooked by previous studies.

Second, and more specifically, individuals also experience traffic congestion when undertaking trips during non-peak hours. At first sight, it sounds reasonable that commuting trips in peak hours (e.g., 7–9 A.M., 3–6 P.M.) are the only time when individuals are exposed to traffic congestion. However, this may not be true for large cities, where serious traffic congestion occurs almost all day long in certain road segments (e.g., [27–29]). This suggests that individuals are exposed to traffic congestion in complex ways in their daily life. They may be exposed to different levels of traffic congestion when undertaking not only commuting trips during peak hours but also other types of travels during non-peak hours.

We thus argue that, for these two reasons, the approaches used in previous studies on individual exposure to traffic congestion and its health impact may exacerbate the uncertain geographic context problem (UGCoP). The UGCoP is a critical methodological issue, and one of the ways in which it may be encountered is when people's exposure to environmental contexts or risk factors (e.g., traffic congestion and air pollution) is inaccurately estimated as a result of ignoring their activity-travel patterns [23,24]. Recent studies have shown that using inaccurately estimated exposures to environmental pollutants may lead to serious inferential errors or misleading results when examining the health impacts of such exposures [30–35].

Further, these approaches in previous studies may aggravate the neighborhood effect averaging problem (NEAP) that arises when individuals' unique activity-travel patterns are ignored [36]. The NEAP suggests that when the probability distribution of residence-based exposures approximates a bell-shaped distribution, individual exposures tend to converge toward the average if individuals' activity-travel patterns are considered when compared to exposures obtained when such patterns are not considered. As one of the manifestations of the UGCoP, the NEAP thus suggests that ignoring individuals' activity-travel patterns may lead to erroneous exposure estimations because of neighborhood effect averaging [36]. Eventually, this may also result in inferential errors or misleading results when researchers investigate the health effects of these exposures [23,24].

In traffic congestion exposure and health studies, these two methodological issues—the UGCoP and the NEAP—arise because of the following two reasons. First, the issues are caused by the spatiotemporal variations in traffic congestion intensities. Although at first glance, the levels of traffic congestion seem to be homogeneous over space and time in large metropolitan areas, this may not be true. For example, intensities of traffic congestion in non-peak hours may not be as severe as those in peak hours (i.e., temporal variations). Also, even for the same time of day, people may be exposed to different traffic congestion intensities based on where trips occur (i.e., spatial variations). Second, the issues arise because individuals are undertaking various types of trips rather than undertaking commuting trips only. Therefore, individuals may be exposed to traffic congestion in complex spatiotemporal ways when they are undertaking commuting trips and other trips [23,37,38].

To sum up, the approaches in previous research examining the effects of traffic congestion exposure on human health largely ignore the UGCoP and the NEAP. This is because both the spatiotemporal variations in traffic congestion and individuals' unique activity-travel patterns increase contextual complexities when measuring individual exposures to traffic congestion. These complexities may lead to erroneous assessments of individual exposures to traffic congestion if their activity-travel patterns are overlooked. Eventually, using inaccurately estimated traffic congestion exposure may exacerbate inferential errors or lead to misleading results when investigating the health effects of traffic congestion exposure.

In light of the relative ignorance of critical methodological issues like the UGCoP and the NEAP in past research on traffic congestion exposure and health, this research seeks to fill this gap. Specifically, this research examines whether individual exposure to traffic congestion is significantly underestimated when individuals' travels beside commuting trips are ignored. To achieve this research goal, we seek to answer the following three questions in this paper by utilizing crowdsourced real-time traffic congestion data and activity-travel data of 250 individuals in Los Angeles, California.

The first question is: Do spatiotemporal variations exist in traffic congestion intensities? In other words, we will investigate whether traffic congestion intensities are different over space and time. Spatiotemporal variations in traffic congestion intensities indicate that two approaches adopted by previous studies (i.e., using data aggregated over areas like census tracts and focusing on commuting trips only) may exacerbate the UGCoP and the NEAP, which is problematic.

The second question is: Will exposures to traffic congestion based on commute-only assessments be significantly lower than those obtained from assessments that also consider individuals' activity-travel patterns in addition to commuting trips? In other words, we will compare individuals' exposure to traffic congestion obtained from two assessments: one that only considers commuting trips and one that also considers individuals' activity-travel patterns in addition to considering commuting trips. If we find that exposures to traffic congestion based on commute-only assessments are significantly lower than those obtained from assessments that also consider individuals' activity-travel patterns in addition to commuting trips, more attention is needed to address the UGCoP and the NEAP in future research on traffic congestion exposure and health.

Lastly, the third question is: How are individuals uniquely exposed to traffic congestion as a result of their activity-travel patterns? In other words, can we observe individuals' idiosyncratic activity-travel patterns and their associated exposures to traffic congestion? Answering this question will help researchers develop an in-depth understanding of the unique characteristics of individuals' activity-travel patterns and how such unique characteristics may affect the accurate assessment of traffic congestion exposure.

3. Study Area and Data

3.1. Study Area

The study area for this research is the Los Angeles-Long Beach-Anaheim Metropolitan area in California, consisting of Los Angeles County and Orange County (Figure 1). We chose the Los Angeles Metropolitan Area for its renowned severe traffic congestion. According to a recent traffic congestion assessment study, the Los Angeles region is ranked as the first out of 297 cities in the U.S. as well as the first out of 1360 cities in the world regarding its severity of traffic congestion [29]. Additionally, the entire metropolitan area is selected as a study unit as it comprehensively captures individuals' socioeconomic characteristics and activity-travel patterns [39].

Figure 1. The Los Angeles Metropolitan Area and its major highways.

3.2. Data

3.2.1. Individual Activity-Travel Survey Data

Since the study seeks to obtain results based on people's actual activity-travel behaviors (not to compare the congested and non-congested travel speeds/times of all road segments in the study area), we employed individual activity-travel survey data accessed via the Transportation Secure Data Center [40]. This individual activity-travel survey is a part of the "National Household Travel Survey California Add-On" survey conducted in 2017. The survey data were collected from around 55,800 individuals (about 26,000 households) in California. Participants were asked to report their activity-travel patterns (e.g., the location of activities, duration of activities, trip purposes, modes of travel, and the number of accompanying passengers) for one survey day and their socioeconomic attributes. Note that the survey did not collect or provide any global positioning system (GPS) data for constructing the space-time trajectories of participants' trips. Therefore, to estimate the travel time of participants' trips and their exposures to traffic congestion, we assumed that they used the shortest path (in terms of travel time) to travel between all locations and used the Google Maps Application Programming Interface (API) to derive the travel time for each trip based on the time of day and the shortest travel route for the trip.

We selected individuals according to the research goal as follows. First, we selected individuals whose trips were all in the study area (i.e., the Los Angeles Metropolitan Area) on weekdays. We only focused on weekdays because weekend activity-travel patterns typically consist of non-routine patterns such as recreational trips and often do not involve any commuting trips [26]. Second, we selected individuals who were actively employed because unemployed people do not undertake any commuting trips. Third, we selected individuals who made trips by driving alone without any accompanying passengers. In other words, all trips in this study were traveled by driving, and trips made by public transit and non-motorized modes (including buses, taxis, bicycles, and walking) were not considered. We focused on these individuals to control other possible travel-mode related factors that may also influence how traffic congestion exposure affects health. For instance, previous studies found that the effects of traffic congestion exposure on health may be different when individuals are drivers rather than passengers (e.g., [41,42]). Moreover, they found that the existence of accompanied passengers may affect drivers' stress (e.g., [7,42,43]).

Lastly, individuals who did not undertake any commuting trips or who had only commuting trips were excluded because we seek to generate two exposure assessments, one that considers only commuting trips and the other considers both commuting and non-commuting trips. Note that we define commuting trips as trips that are anchored at a workplace so that we can consider trip-chaining travel behaviors. Before this exclusion criterion was applied, there were 729 individuals in our subsample. As a result of applying this exclusion criterion, 77 individuals (11%) were excluded because they did not make any commuting trips (e.g., having a day-off from the work), and an additional 402 individuals were removed since they only made commuting trips (i.e., no other type of trips). Note that a considerable portion of the survey participants (34%, 250 individuals) reported that they made commuting as well as non-commuting trips. This provides a compelling rationale that individuals' activity-travel patterns still should be considered to accurately assess their exposure to traffic congestion.

After this selection process, 250 individuals were finally included in the subsample used in this study. To avoid sample selections that do not have similar sociodemographic characteristics as the larger population in the Los Angeles Metropolitan Area, we compared their socio-economic attributes with those of the larger population in the study area. Note that since our research focuses on employed individuals, the statistics reported in Table 1 represent only employed workers. Overall, descriptive statistics of the selected participants showed similarity to those of the larger population in the study area. The only discrepancy we found is that the median age of the selected participants (45.2 years old) is higher than that of the larger population in the Los Angeles Metropolitan Area (39.9 years old).

This can be explained by underrepresentation of the younger generations in our subsample as we focus on workers who drive their own cars. Recent travel behavior studies revealed that the younger generation (e.g., millennials) may drive less or not own cars (e.g., [44,45]). Therefore, it is likely that the younger generation may be underrepresented in our subsample.

Table 1. Comparison of the sociodemographic attributes of the 250 selected participants with those of the larger population of the Los Angeles Metropolitan Area.

Sociodemographic Attributes	The 250 Selected Individuals			Los Angeles Metropolitan Area [(a)]		
Gender	Male 53%		Female 47%	Male 55%		Female 45%
Race	White 59%	African-American 6%	Asian 14%	White 58%	African-American 6%	Asian 17%
Age (mean)	45.2 years old			39.9 years old		
% of people with higher education [(b)]	62%			65%		

[(a)] American Community Survey (ACS) 2016 5-year estimates, [(b)] Higher education indicates education attainment that is equal to or higher than bachelor's degree.

3.2.2. Real-Time Traffic Congestion Data

In this study, we estimate individual exposure to traffic congestion for each trip by subtracting its free-flow travel time from its estimated travel time that considers traffic congestion, following the framework used in previous studies (e.g., [29,46,47]). For example, imagine an individual who undertakes 5 trips in his or her daily life (Figure 2). This person travels from the home location to a workplace (Trip 1), goes back from the workplace to home (Trip 2), goes grocery shopping from home (Trip 3), goes to the gym after the grocery shopping (Trip 4), and finally goes back home from the gym (Trip 5). For each of these trips, by subtracting the free-flow travel time from the estimated travel time (obtained using the Google Maps API (Google, Mountain View, CA, USA) based on time of day and the origin and destination of the trip), we can estimate this person's exposure to traffic congestion for each trip.

Recall that the primary goal of this research is to compare individual exposures to traffic congestion obtained from two assessments: one that only considers commuting trips and one that also considers individuals' activity-travel patterns in addition to considering commuting trips. Assume that the person in this example is exposed to traffic congestion for 10 min for each of the five trips. An assessment that only considers the two commuting trips (Trips 1 and 2) estimates that the duration of exposure to traffic congestion is 20 min (10 + 10), while an assessment that also considers the non-commuting trips (Trips 3, 4, and 5) as well as the commuting trips (Trips 1 and 2) estimates that the duration of exposure to traffic congestion is 50 min (10 + 10 + 10 + 10 + 10).

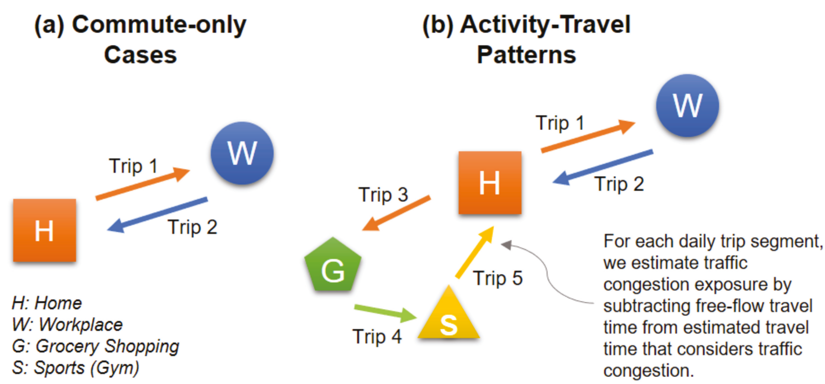


Figure 2. An example of estimating an individual's exposure to traffic congestion for two types of assessments: (**a**) commute-only versus (**b**) activity-travel patterns in addition to commuting trips.

To obtain free-flow travel time and estimated travel time, we utilized the Google Maps API (Figure 3). The Google Maps API estimates driving time between two points when API users provide departure time, departure/arrival/waypoints locations, and route options (e.g., avoiding toll roads or highways) [48,49]. The API computes driving time based on two data sources: (1) crowdsourced real-time traffic data that were submitted by anonymous drivers who consent to send their location information to Google Maps via their smartphones and (2) historical traffic flow databases that Google has established [48]. Free-flow travel time is derived as if trips occurred at 2 A.M. when traffic volumes practically approach 0. To the best of our knowledge, while no study has compared the accuracy of Google Maps data with those from other sources, it seems that travel times provided by Google Maps are highly accurate based on several sources on the web. For instance, in one assessment that used 56 trips with an average journey time of 32 min, the average travel time difference between actual and estimated travel times is 1.8 min (see https://blog.ancoris.com/how-accurate-is-google-maps-journey-time).

In this research, the departure time and geographic coordinates (e.g., longitude and latitude) of the origin and destination of each trip of the participants recorded in the travel survey were used to obtain free-flow travel time and estimated travel time through the Google Maps API. Note that we did not use the reported travel times from the survey as the actual travel time in this study because the survey did not provide the travel routes of participants' trips and it is not possible to estimate the corresponding free-flow travel time for each of the participants' trip, which in turn renders the comparison between the free-flow travel time and the (estimated) travel time that considers traffic congestion for each trip of the participants impossible. Further, estimating actual travel time using the Google Maps API serves to avoid the recall and rounding errors common in the reported travel times of travel surveys.

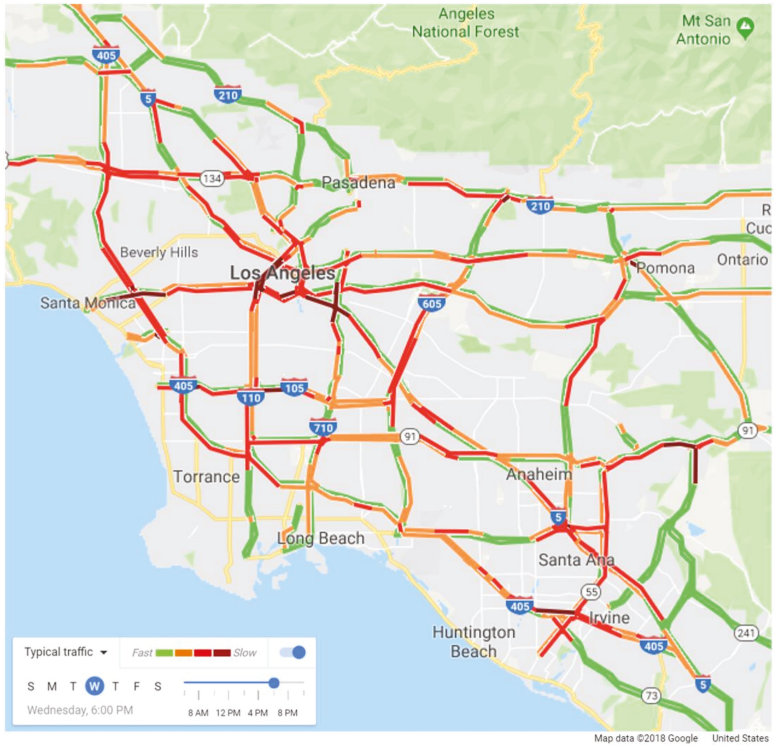


Figure 3. A screenshot of a map of typical traffic congestion levels at 6 P.M. on Wednesday in the Los Angeles Metropolitan Area (Source: Google Maps).

Using the Google Maps API service has several advantages over traditional desktop-based GIS programs. One compelling advantage is that using the API service does not require researchers to prepare a considerable amount of data and use considerable computing resources. For example, the Google Maps API promptly provides users with travel time that considers traffic congestion between any two given locations (i.e., longitude, latitude) at 20-min intervals.

To obtain this detailed travel time estimate, conventional desktop-based GIS programs require researchers to prepare a considerable amount of network data (e.g., [50,51]). For example, researchers need to prepare road network files, estimate the traffic volume and speed on each road segment at each time interval, and generate penalty information for each street intersection (e.g., one-way roads, no-left-turn penalty, and so on). Preparing these datasets may not be feasible for large metropolitan areas such as Los Angeles. Additionally, even if researchers can prepare the required data, it may take substantial time to run the shortest-path algorithm since the road networks are large and complex. However, by using the Google Maps API service, researchers only need to develop a simple program based on easily accessible programming languages (e.g., Python, Java, and so on). Moreover, since the calculations of travel time are performed inside the API service (where the API uses its own high-performance computing facilities), researchers can get results immediately. For these reasons, there has recently been a growing number of studies that extensively employed the Google Maps API and other map-based API services (e.g., [52,53]).

However, it should be noted that the Google Maps API service has several limitations. One limitation is that users may not know the detailed mechanism of how it estimates travel time. However, documentation from API service providers may mitigate this issue (e.g., [49]).

Another limitation is that API services may charge a fee based on the number of API requests. For example, Google Maps API users can use 40,000 API requests per month for free. Beyond the 40,000 free requests, users need to pay a fee per single API request (e.g., a single query of travel time estimation for a single pair of origin and destination) [49]. Thus, the Google Maps API service may not be a viable option for researchers who want to obtain travel times for a larger number of origin-destination pairs [54]. However, this limitation did not significantly affect our research because we did not need a large number of requests; we requested travel time estimates for approximately 1000 trips, which the 250 selected participants undertook.

4. Results

In the first part of Section 4.1 below, we explore whether traffic congestion intensities are different across space and time in the study area based on data from the INRIX 2017 Global Traffic Scorecard [29]. In the second part of Section 4.1, we examine how traffic congestion levels are different over space and time based on the 1022 trips made by the 250 selected participants of the survey. In Section 4.2, we compare two assessments of individual exposures to traffic congestion for the 250 participants: one that considers only commuting trips and the other one that considers both commuting and non-commuting trips. In Section 4.3, we explore how three individuals from the same household are exposed to traffic congestion in unique ways over space and time through visualizations of their space-time trajectories.

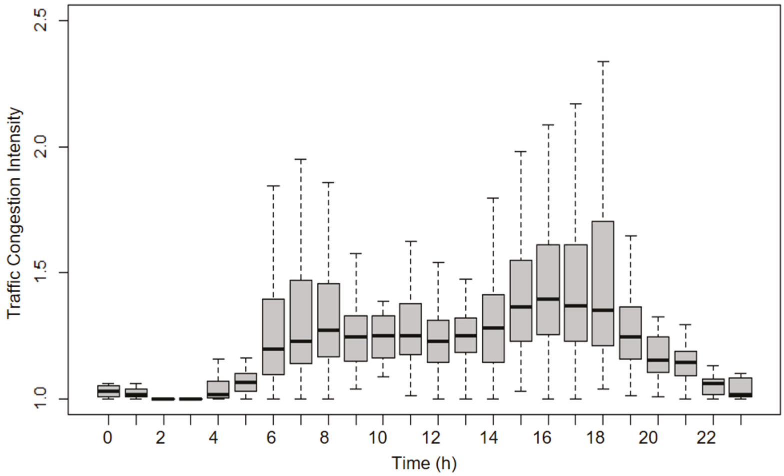
4.1. Spatiotemporal Variations in Traffic Congestion Intensities

In this section, we answer the first research question: Do spatiotemporal variations exist in traffic congestion intensities? We empirically examine whether traffic congestion intensities are different across space and time in the study area based on data from the INRIX 2017 Global Traffic Scorecard [29] and the trips made by the 250 selected participants. Using the first data source, traffic congestion intensity is assessed in terms of the percentage of congestion travel time that drivers experience out of gross travel time; using the second data source, traffic congestion intensity is derived as the ratio of travel time that considers traffic congestion to free-flow travel time. This is an important question because spatiotemporal variabilities in traffic congestion intensities may exacerbate the UGCoP. Although it is widely known that traffic congestion intensities are different over space and time [55], here we empirically investigate its precise spatiotemporal configurations in the study area.

First, we investigate general spatiotemporal variations of traffic congestion intensities in the study area using data from the INRIX 2017 Global Traffic Scorecard, which provides data on traffic congestion for over 1360 cities around the world (Table 2) [29]. Although the Los Angeles Metropolitan Area is globally notorious for its severe traffic congestion, and as these data indicate, we can observe the spatiotemporal heterogeneity of the traffic congestion in the area: (1) There are temporal variations in traffic congestion. For instance, inter-city drivers experience traffic congestion for 22% of their gross travel time during peak hours while 10% of their gross time occurs during non-peak hours (e.g., around noon) [29]. (2) There are spatial variations in traffic congestion. For example, during non-peak hours, intra-city drivers experience traffic congestion for 13% of their gross travel time, while inter-city drivers experience traffic congestion only for 10% of the gross travel time [29].

Second, we examine the spatiotemporal variations in traffic congestion intensities based on the 1022 trips made by the 250 selected participants (note that, here, traffic congestion intensity for a trip is derived as the ratio of travel time that considers traffic congestion to free-flow travel time; see Figure 2 and earlier description on how these two travel times for each trip are derived). Figure 4 illustrates traffic congestion intensity (the vertical axis) variations of these 1022 trips by trip departure time (the horizontal axis). The ratio of congestion and free-flow travel times is widely used in practice to represent the severity of traffic congestion at the road-segment level (e.g., [29,46,47]). For example, if traffic congestion intensity is 1.5, it means that the travel time that considers traffic congestion is

1.5 times longer than the free-flow travel time due to traffic congestion. Therefore, the minimum value of the traffic congestion intensity is 1.0 because free-flow travels give the minimum travel time.



Mean	1.032	1.000	1.047	1.284	1.377	1.323	1.255	1.329	1.512	1.475	1.207	1.085
Standard Deviation	0.027	N/A	0.062	0.251	0.338	0.255	0.165	0.244	0.423	0.363	0.156	0.096

Figure 4. Observed traffic congestion intensity over space and time.

We can observe the following two things in Figure 4. First, there are temporal variations in traffic congestion intensity. Not surprisingly, traffic congestion is generally severe during peak hours (e.g., 7–9 A.M., 3–6 P.M.). Moreover, although traffic congestion during non-peak hours is less severe than that of peak-hours, traffic congestion is still observed at any time during a day. This corroborates our earlier observation that there are temporal variations in traffic congestion in the Los Angeles Metropolitan Area. Second, there are spatial variations in traffic congestion intensity. Vertical distributions of the observations (i.e., indicated by the range of the boxes) represent various traffic congestion intensities at different locations at each hour. For example, the range of the box at 6 P.M. is wider than that at 12 P.M., which means more spatial variations in traffic congestion intensities exist at 6 P.M.

These two findings answer our first research question: Do spatiotemporal variations exist in traffic congestion intensities? Based on the observations made from Table 2 and Figure 4, it is clear that spatiotemporal variations of traffic congestion intensities exist in the study area. In what follows, we continue our analysis to compare individuals' exposures to traffic congestion obtained from two assessments based on the trips made by the 250 selected participants: one that only considers commuting trips and one that considers both commuting and non-commuting trips (i.e., taking into account individuals' activity-travel patterns).

Table 2. Spatiotemporal variations in traffic congestion in the study area (Source: [29]).

Types of Trips	Peak Hours	Non-Peak Hours (Daytime)	Non-Peak Hours (Late Nighttime)
Intra-city trips [(b)]	20% [(a)]	13%	8%
Inter-city trips [(c)]	22%	10%	3%

[(a)] Percentage of congestion travel time that drivers experience out of gross travel time. [(b)] Intra-trips refer trips that occur within the city. [(c)] Inter-trips refer trips that occur into/out of the city.

4.2. Differences in Individual Exposures to Traffic Congestion between the Two Assessments

We conduct a paired sample *t*-test to see whether individual exposures to traffic congestion for each participant are significantly different between the commute-only assessment and the assessment that also considers participants' activity-travel patterns (i.e., considering both commuting and non-commuting trips) (please see Section 3.2.2 and Figure 2 for a detailed explanation of the method). Here, individual exposure to traffic congestion for each trip is estimated by subtracting its free-flow travel time from its estimated travel time (which considers traffic congestion and is estimated using the Google Maps API). This is the additional travel time for a trip due to traffic delay or congestion.

For each participant, we obtain a commute-only exposure measure by adding the additional travel times incurred by the commuting trips and another exposure measure by adding the additional travel times of both commuting and non-commuting trips. We then compare the difference between these two exposure measures for each participant (and thus a paired sample *t*-test is used).

Table 3 indicates that the mean difference in participants' exposure to traffic congestion is 6.66 min, which means that the duration for which a participant experiences traffic congestion increases on average by 6.66 min (47.78%) when participants' activity-travel patterns are considered, compared to the commute-only assessment. The result of the paired sample *t*-test confirms that the differences in exposures to traffic congestion between the two assessments are statistically significant ($p < 0.001$). Figure 5 visualizes the results presented in Table 3. The box plots also show that the average exposure to traffic congestion when individuals' activity-travel patterns are considered is higher than the average exposure obtained in the commute-only assessment.

Table 3. Mean exposure to traffic congestion in the commute-only assessment and the assessment that also considers participants' activity-travel patterns.

Statistics	Commute-Only Assessment	Activity-Travel Patterns Considered	Mean of Differences [a]
Mean	13.94 (min)	20.60 (min)	6.66 *** [b]
Standard Deviation	14.87	18.13	-

[a] Paired sample *t*-test result, [b] *** *p*-value < 0.001.

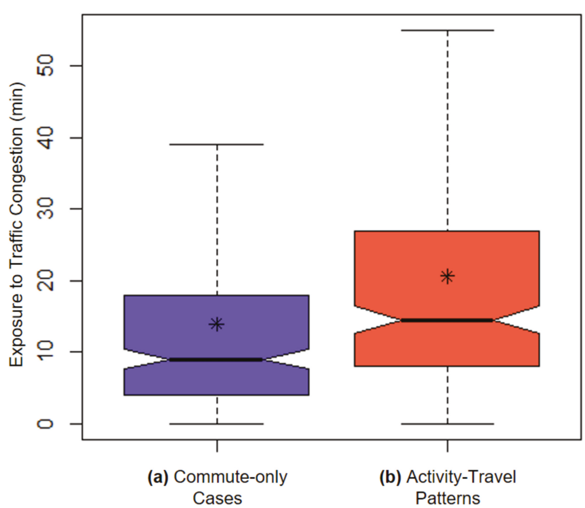


Figure 5. Box plots of individuals' traffic congestion exposure for (**a**) the commute-only assessment and (**b**) the assessment that also considers participants' activity-travel patterns. (Outliers are not presented.)

Moreover, we examine the probability distribution function of individual exposure to traffic congestion. Figure 6 presents the histograms of individuals' traffic congestion exposure levels for the two assessments. As the histograms show, when individuals' activity-travel patterns are considered (right histogram), the shape of the probability distribution function becomes less skewed and converges to its mean. Table 4 shows that skewness (from 2.313 to 1.724) and kurtosis (from 9.766 to 5.982) of the histogram decrease after activity-travel patterns are considered. The results indicate that the probability distribution function of individual exposures to traffic congestion shows a tendency to converge to its average when individuals' activity-travel patterns are considered.

This phenomenon can also be understood as a manifestation of the neighborhood effect averaging problem (NEAP) observed by Kwan [36]. However, there are two important differences between our observations here and the original interpretation of the NEAP put forward by Kwan [36]. First, both exposure assessments (i.e., one that considers only commuting trips and one that considers people's entire activity-travel patterns) in this research are mobility-based. In other words, the commute-only exposure assessment is not residence-based because it already included some portion of individuals' daily mobility (i.e., commuting trips). However, the original articulation of the NEAP compares residence-based exposures with mobility-based exposures. This indicates that the NEAP can also be encountered in environmental exposure assessments when only parts (instead of all) of people's daily mobility are ignored.

Second, the probability distribution functions of the two exposure assessments in this study are not bell-shaped but heavily skewed. The original notion concerning the NEAP only focuses on distributions of individual exposures that approximate a bell-shaped distribution (one such distributions is the normal distribution), but distribution functions in our research are heavily skewed. This indicates that the neighborhood effect averaging problem can also be encountered when the probability distributions of individual exposures are not bell-shaped. These two differences between our observations here and the original interpretation of the NEAP extends the original interpretation of the NEAP in important ways.

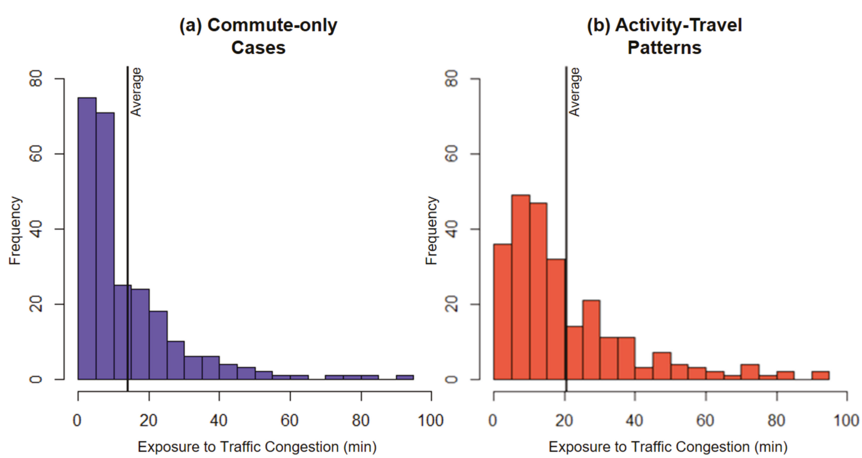


Figure 6. The histograms of traffic congestion exposure in (**a**) the commute-only assessment and (**b**) the assessment that considers participants' activity-travel patterns.

Table 4. Skewness and kurtosis of the histograms for both assessments.

Statistics	Commute-Only Assessment	Activity-Travel Patterns Considered
Skewness	2.313	1.724
Kurtosis	9.766	5.982

Additionally, we examine in detail how considering individuals' activity-travel patterns impact the relative levels of individual exposure to traffic congestion. Figure 7 illustrates the standardized (*z*-score) individual exposures to traffic congestion. The horizontal axis displays individual exposures of the commute-only cases, while the vertical axis represents individual exposures obtained by the assessment that considers participants' activity-travel patterns. For example, points in the first quadrant (top-right) represent cases when individual exposures to traffic congestion are higher than the average in both assessments. On the contrary, points in the third quadrant (bottom-left) indicates that individual exposures to traffic congestion are lower than its average in both assessments. A closer examination of the graph yields a couple of important findings.

First, the standardized (*z*-score) individual exposures to traffic congestion of most participants (203 participants, 81% of the selected subsample) range between −1 and 0 (see the focused area in the inset). Second, a majority of participants (48 out of 69) in the first quadrant are located in the blue triangular area. This indicates that individual exposures shift much closer to its mean value when activity-travel patterns are considered. These findings confirm an earlier observation that many individuals have exposure levels around the average value while fewer individuals have very high or low exposure levels, and considering individuals' activity-travel patterns leads the exposure level to converge to its mean [36].

Based on these results, we also answer the second question: Will exposures to traffic congestion based on commute-only assessments be significantly lower than those obtained from assessments that also consider individuals' activity-travel patterns in addition to commuting trips? We found that this is indeed the case. There are statistically significant differences between exposures evaluated with and without considering individuals' activity-travel patterns, indicating that the UGCoP is a serious issue. We also found that ignoring individuals' activity-travel patterns may exacerbate the NEAP. Therefore, we can conclude that overlooking people's activity-travel patterns may lead to serious methodological issues in the form of the UGCoP and the NEAP when assessing their exposures to traffic congestion.

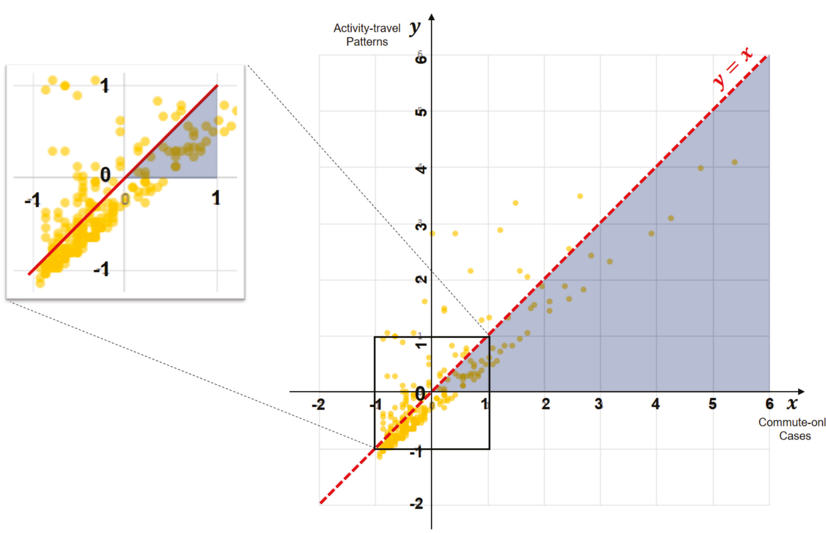


Figure 7. Standardized (*z*-score) individual exposures to traffic congestion.

4.3. Space-Time Visualizations of Individual Exposures to Traffic Congestion

Figure 8 illustrates the cumulative traffic congestion exposures of 3 individuals from the same residence (i.e., the same household) over the 24 hours of the survey day. Line A (blue) indicates the mother's exposure to traffic congestion, while Lines B (yellow) and C (red) represent the older son and the younger daughter respectively. Also, space-time visualizations of these individuals' activity-travel patterns are presented in Figure 9. The vertical axis (t) represents time, and the horizontal plane displays space (x, y). Each dot represents a 1-min interval in the trip trajectories obtained from the Google Maps API. The size of the dots indicates traffic congestion intensity. For instance, larger dots represent more intense traffic congestion. The vertical solid lines indicate durations when individuals are performing activities at fixed locations, as their location (x, y) does not change over time.

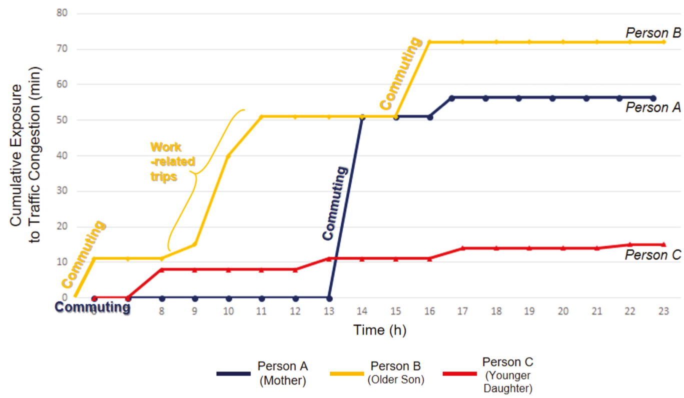


Figure 8. Cumulative exposures to traffic congestion of three selected individuals from the same household (A: Mother, B: Older Son, C: Younger Daughter).

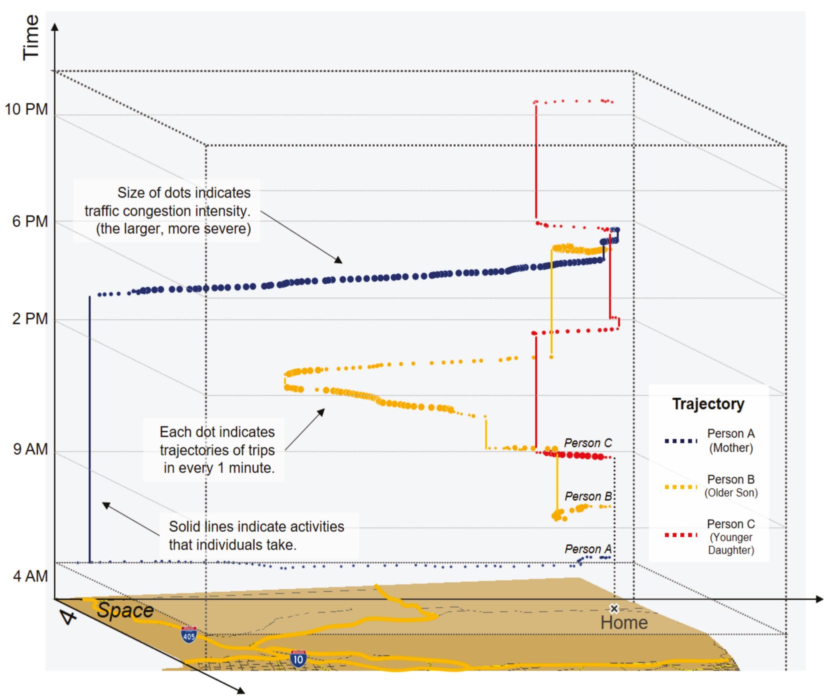


Figure 9. Space-time visualizations of traffic congestion exposures of three individuals from the same household (A: Mother, B: Older Son, C: Younger Daughter).

The space-time illustration of individuals' trajectories clearly shows that individuals are exposed to traffic congestion in unique ways over space and time. First, although all 3 individuals have a similar travel-demand environment (e.g., actively employed and driving their own cars), each family member's traffic congestion exposure varies because of their idiosyncratic activity-travel patterns.

As Figure 9 shows, for example, the traffic congestion exposure of Person A (mother) ranges between 10 and 70, while that of Person C (younger daughter) varies between 10 and 20. This difference can be explained by the different activity-travel patterns between these two persons. Person A takes longer commuting trips, and she is heavily exposed to traffic congestion especially during her way back home. Person A is exposed to severe traffic congestion when she takes non-commuting trips, but the non-commuting trips do not significantly contribute to the total exposure because the length of the trips is relatively short. By contrast, Person C is less exposed to traffic congestion than Person A. Most trips that Person C takes are near her residence and relatively short, which enables her to avoid heavy exposure to traffic congestion.

Specifically, we can observe that the younger son (Person B) is exposed to heavier traffic congestion than the others in the household. His traffic congestion exposure becomes more severe when his activity-travel patterns are considered. This drastic increase is mainly because of his work-related trips happening near areas in South Los Angeles during non-peak hours (e.g., 9 A.M.–2 P.M.), when traffic congestion there is still severe.

Further, Figure 10 depicts the exposures to traffic congestion of 32 individuals from 15 households in the subsample for both the commute-only assessments and assessments that also consider their activity-travel patterns. The bar graph clearly shows that individuals from the same household are differently exposed to traffic congestion because individuals' activity-travel patterns are idiosyncratic.

Based on these results, we answer the third question: How are individuals uniquely exposed to traffic congestion as a result of their activity-travel patterns? The results confirm that individuals are idiosyncratically exposed to traffic congestion due to their distinctive activity-travel patterns. The results also corroborate previous studies [37,56], which argue that individuals from the same household are differently exposed to environmental influences or contexts. Therefore, these results may cast doubt on the validity of previous studies' exclusive focus on commuting trips during peak hours and using data aggregated over areas (e.g., census tracts) because they did not fully reflect individuals' unique activity-travel patterns.

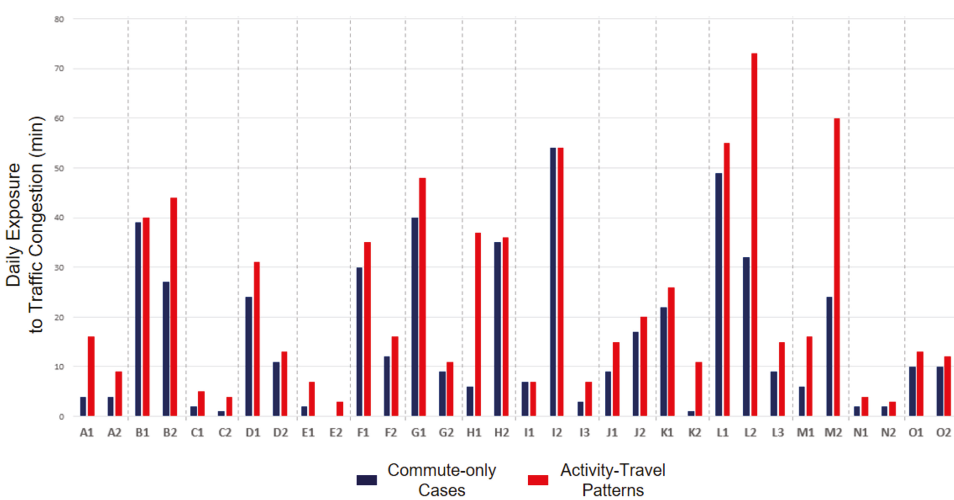


Figure 10. Individual exposures to traffic congestion of 32 individuals from 15 households in the subsample (Same alphabet means the same household.)

5. Conclusions

This research empirically examined whether the uncertain geographic context problem (UGCoP) and the neighborhood effect averaging problem (NEAP) are encountered in research on individual exposures to traffic congestion. We used crowdsourced real-time traffic congestion data and activity-travel data of 250 individuals in Los Angeles to compare two assessments of individual exposure to traffic congestions: one that only considers commuting trips and one that also considers individuals' non-commuting trips in addition to considering commuting trips (thus taking people's activity-travel patterns into account).

First, the results indicated that spatiotemporal variations in traffic congestion intensity exist in the study area, which calls for the consideration of individuals' activity-travel patterns when assessing their exposures to traffic congestion in future research.

Second, the paired sample *t*-test results revealed that individual exposures to traffic congestion are significantly underestimated when individuals' activity-travel patterns are ignored. Further, the results highlighted that the probability distribution function of individual exposures is heavily skewed but tends to converge to its average value when individuals' activity-travel patterns are considered, which is a manifestation of the neighborhood effect averaging problem (NEAP). These results indicated that both the UGCoP and the NEAP are critical methodological issues in traffic congestion and health studies.

Lastly, we presented space-time visualizations of the traffic congestion exposures of 3 individuals from the same household. The results illustrated that since individuals have idiosyncratic activity-travel patterns, their exposures to traffic congestion vary significantly even if they live at the same residential location. These results corroborate the claim in previous studies that using residence-based methods or data aggregated over areas (e.g., census tracts) may lead to erroneous assessments of individual exposures to traffic congestion or other environmental influences [31].

The results of our research imply that epidemiological studies should pay more attention to individuals' activity-travel patterns when assessing people's environmental exposures. As the results of this study illustrated, ignoring individuals' daily mobility (i.e., activity-travel patterns) may result in erroneous assessments of their exposures to traffic congestion. Eventually, using inaccurately estimated traffic congestion exposures may exacerbate inferential errors or lead to significantly misleading results when investigating the effects of traffic congestion exposures on human health. In addition, the results imply that researchers who study environmental exposures should focus more on individual-level analysis, since this study shows that people who live in the same residential location may have different traffic congestion exposures due to their distinct activity-travel patterns. This implies that using residence-based methods or data aggregated over areas (e.g., census tracts), which are popular approaches in previous studies, may lead to critical methodological issues.

Although this study significantly advances our knowledge about two critical methodological issues (i.e., the UGCoP and the NEAP) in traffic congestion and health studies, it has several limitations that should be addressed in further studies. First, we presumed that individuals used the shortest path (in terms of travel time) to travel from one location to another, which may not fully capture their true activity-travel patterns. Although this assumption is reasonable, people may not necessarily use the shortest path. One possible solution to this issue may be to employ a space-time prism [57], as illustrated in Figure 11. A space-time prism consists of points that individuals may possibly visit given their spatiotemporal constraints [50,58]. Considering possible alternative routes in a space-time prism may help researchers comprehensively assess traffic congestion exposures.

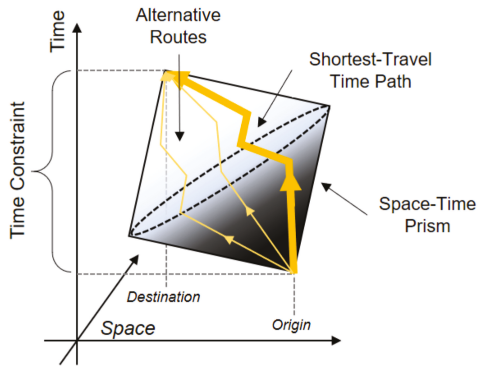


Figure 11. An example of a space-time prism.

Second, further studies need to consider the subjective aspects of individual exposures to traffic congestion. Previous studies revealed that exposures to traffic congestion may go through subjective perception filters [12,16,59]. This indicates that although people are exposed to the same level of objective traffic congestion (e.g., 20 min in traffic congestion), the effects of the objective traffic congestion on health may follow different mechanisms for each individual. However, due to the limitations of the survey data used in this study (which did not collect or provide any data on participants' perceptions), we were not able to address this issue. One possible solution may be to integrate in-depth interviews about subjective factors with activity-travel surveys [25,60]. By combining the subjective experiences of traffic congestion with objective measures, future research may further advance our knowledge of the health impacts of exposures to traffic congestion.

Third, more scrutiny is required to unveil the temporal dimension of the effects of traffic congestion exposures on health [24]. We computed total traffic congestion exposure in minutes because we presumed that cumulative exposure may influence health. However, the effects of traffic congestion exposures on health may show "time-lagged response" [24], which means that it may have health effects afterward. Moreover, not only the duration of traffic congestion but also the variability in driving time may negatively affect health. Since several epidemiological studies reported such evidence [15,61], more attention is required to clearly understand the temporal aspect of the effects of traffic congestion on health. One possible solution is to utilize real-time global positioning system (GPS) technology to gain clearer pictures and more detailed understanding of these temporal effects of traffic congestion on health [62,63].

Author Contributions: The authors confirm contribution to the paper as follows: study conception and design: J.K., M.-P.K.; data collection and computing: J.K.; analysis and interpretation of results: J.K., M.-P.K.; draft manuscript preparation: J.K., M.-P.K.

Funding: Junghwan Kim was supported by a Foster Fellowship. Mei-Po Kwan was supported by grants from the U.S. National Science Foundation (BCS-1832465) and the National Natural Science Foundation of China (41529101).

Acknowledgments: Junghwan Kim would like to thank Brennan Borlaug and the Transportation Secure Data Center (TSDC) for their kind support on data access. The authors thank the three anonymous reviewers for their helpful comments.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. O'sullivan, A. *Urban Economics*; McGraw-Hill/Irwin: New York City, NY, USA, 2007.
2. Hartgen, D.T.; Fields, M.G.; Layzell, A.L.; Jose, E.S. How employers view traffic congestion: Results of National Survey. *Transp. Res. Rec. J. Transp. Res. Board* **2012**, *2319*, 56–66. [CrossRef]

3. Anderson, M. How to beat the traffic. *Science* **2017**, *357*, 36–37. [CrossRef] [PubMed]
4. Traffic Gridlock Sets New Records for Traveler Misery. Available online: https://tti.tamu.edu/news/traffic-gridlock-sets-new-records-for-traveler-misery/ (accessed on 27 November 2018).
5. United Nations Department of Economic and Social Affairs. Available online: http://www.un.org/en/development/desa/news/population/2015-report.html (accessed on 27 November 2018).
6. Stokols, D.; Novaco, R.W. Transportation and Well-Being. In *Transportation and Behavior*; Altman, I., Wohlwill, J.F., Everett, P.B., Eds.; Springer: Boston, MA, USA, 1981; pp. 85–130. ISBN 978-1-4684-4003-4.
7. Schaeffer, M.H.; Street, S.W.; Singer, J.E.; Baum, A. Effects of control on the stress reactions of commuters. *J. Appl. Soc. Psychol.* **1988**, *18*, 944–957. [CrossRef]
8. Lundberg, U. Urban commuting: Crowdedness and catecholamine excretion. *J. Hum. Stress* **1976**, *2*, 26–32. [CrossRef] [PubMed]
9. Levy, J.I.; Buonocore, J.J.; von Stackelberg, K. Evaluation of the public health impacts of traffic congestion: A health risk assessment. *Environ. Health* **2010**, *9*, 1–12. [CrossRef] [PubMed]
10. Currie, J.; Walker, R. Traffic congestion and infant health: Evidence from E-ZPass. *Am. Econ. J. Appl. Econ.* **2011**, *3*, 65–90. [CrossRef]
11. Sandow, E.; Westerlund, O.; Lindgren, U. Is your commute killing you? On the mortality risks of long-distance commuting. *Environ. Plan. A* **2014**, *46*, 1496–1516. [CrossRef]
12. Stokols, D.; Novaco, R.W.; Stokols, J.; Campbell, J. Traffic congestion, Type A behavior, and stress. *J. Appl. Psychol.* **1978**, *63*, 467–480. [CrossRef]
13. Hennessy, D.A.; Wiesenthal, D.L. The relationship between traffic congestion, driver stress and direct versus indirect coping behaviours. *Ergonomics* **1997**, *40*, 348–361. [CrossRef]
14. Hennessy, D.A.; Wiesenthal, D.L. Traffic congestion, driver stress, and driver aggression. *Aggress. Behav.* **1999**, *25*, 409–423. [CrossRef]
15. Kluger, A.N. Commute variability and strain. *J. Organ. Behav.* **1998**, *19*, 147–165. [CrossRef]
16. Gottholmseder, G.; Nowotny, K.; Pruckner, G.J.; Theurl, E. Stress perception and commuting. *Health Econ.* **2009**, *18*, 559–576. [CrossRef] [PubMed]
17. Gulian, E.; Debney, L.; Glendon, A.I.; Davies, D.R. Coping with driver stress. In *Stress and Tension Control 3*; Springer: Boston, MA, USA, 1989; pp. 173–186.
18. Stutzer, A.; Frey, B.S. Stress that Doesn't Pay: The commuting paradox. *Scand. J. Econ.* **2008**, *110*, 339–366. [CrossRef]
19. Hansson, E.; Mattisson, K.; Björk, J.; Östergren, P.-O.; Jakobsson, K. Relationship between commuting and health outcomes in a cross-sectional population survey in southern Sweden. *BMC Public Health* **2011**, *11*, 1–14. [CrossRef] [PubMed]
20. Denstadli, J.M.; Julsrud, T.E.; Christiansen, P. Urban commuting—A threat to the work-family balance? *J. Transp. Geogr.* **2017**, *61*, 87–94. [CrossRef]
21. Beland, L.-P.; Brent, D.A. Traffic and crime. *J. Public Econ.* **2018**, *160*, 96–116. [CrossRef]
22. Brauer, M.; Reynolds, C.; Hystad, P. Traffic-related air pollution and health in Canada. *Can. Med. Assoc. J.* **2013**, *185*, 1557–1558. [CrossRef] [PubMed]
23. Kwan, M.-P. The uncertain geographic context problem. *Ann. Assoc. Am. Geogr.* **2012**, *102*, 958–968. [CrossRef]
24. Kwan, M.-P. The limits of the neighborhood effect: Contextual uncertainties in geographic, environmental health, and social science research. *Ann. Am. Assoc. Geogr.* **2018**, *108*, 1482–1490. [CrossRef]
25. Olsson, L.E.; Gärling, T.; Ettema, D.; Friman, M.; Fujii, S. Happiness and satisfaction with work commute. *Soc. Indic. Res.* **2013**, *111*, 255–263. [CrossRef]
26. Santos, A.; McGuckin, N.; Nakamoto, D.; Gray, D.; Liss, S. *Summary of Travel Trends: 2009 National Household Travel Survey*; (No. FHWA-PL-11-022); U.S. Department of Transportation: Washington, DC, USA, 2011.
27. Grynbaum, M.M. In Manhattan, driving is toughest on Wednesdays. *The New York Times*, 23 March 2010.
28. Zhan, X.; Hasan, S.; Ukkusuri, S.V.; Kamga, C. Urban link travel time estimation using large-scale taxi data with partial information. *Transp. Res. Part C Emerg. Technol.* **2013**, *33*, 37–49. [CrossRef]
29. INRIX Global Traffic Scorecard. Available online: http://inrix.com/scorecard/ (accessed on 27 November 2018).
30. Dewulf, B.; Neutens, T.; Lefebvre, W.; Seynaeve, G.; Vanpoucke, C.; Beckx, C.; Van de Weghe, N. Dynamic assessment of exposure to air pollution using mobile phone data. *Int. J. Health Geogr.* **2016**, *15*, 14. [CrossRef] [PubMed]

31. Park, Y.M.; Kwan, M.-P. Individual exposure estimates may be erroneous when spatiotemporal variability of air pollution and human mobility are ignored. *Health Place* **2017**, *43*, 85–94. [CrossRef] [PubMed]
32. Shafran-Nathan, R.; Levy, I.; Broday, D.M. Exposure estimation errors to nitrogen oxides on a population scale due to daytime activity away from home. *Sci. Total Environ.* **2017**, *580*, 1401–1409. [CrossRef] [PubMed]
33. Kwan, M.-P.; Wang, J.; Tyburski, M.; Epstein, D.H.; Kowalczyk, W.J.; Preston, K.L. Uncertainties in the geographic context of health behaviors: A study of substance users' exposure to psychosocial stress using GPS data. *Int. J. Geogr. Inf. Sci.* **2018**, 1–20. [CrossRef]
34. Yu, H.; Russell, A.; Mulholland, J.; Huang, Z. Using cell phone location to assess misclassification errors in air pollution exposure estimation. *Environ. Pollut.* **2018**, *233*, 261–266. [CrossRef]
35. Zhao, P.; Kwan, M.-P.; Zhou, S. The uncertain geographic context problem in the analysis of the relationships between obesity and the built environment in Guangzhou. *Int. J. Environ. Res. Public. Health* **2018**, *15*, 308. [CrossRef]
36. Kwan, M.-P. The neighborhood effect averaging problem (NEAP): An elusive confounder of the neighborhood effect. *Int. J. Environ. Res. Public. Health* **2018**, *15*, 1841. [CrossRef]
37. Kwan, M.-P. Beyond space (as we knew it): Toward temporally integrated geographies of segregation, health, and accessibility. *Ann. Assoc. Am. Geogr.* **2013**, *103*, 1078–1086. [CrossRef]
38. Helbich, M. Toward dynamic urban environmental exposure assessments in mental health research. *Environ. Res.* **2018**, *161*, 129–135. [CrossRef] [PubMed]
39. Wilson, S.G.; Plane, D.A.; Mackun, P.J.; Fischetti, T.R.; Goworowska, J. *Patterns of Metropolitan and Micropolitan Population Change: 2000 to 2010*; 2010 Census Special Reports; US Census Bureau: Suitland, MD, USA, 2012; pp. 1–102.
40. Transportation Secure Data Center. Available online: www.nrel.gov/tsdc (accessed on 27 November 2018).
41. Evans, G.W.; Wener, R.E.; Phillips, D. The morning rush hour: Predictability and commuter stress. *Environ. Behav.* **2002**, *34*, 521–530. [CrossRef]
42. Bou Samra, P.; El Tomb, P.; Hosni, M.; Kassem, A.; Rizk, R.; Shayya, S.; Assaad, S. Traffic congestion and blood pressure elevation: A comparative cross-sectional study in Lebanon. *J. Clin. Hypertens.* **2017**, *19*, 1366–1371. [CrossRef]
43. Novaco, R.W.; Collier, C. Commuting stress, ridesharing, and gender: Analyses from the 1993 state of the commute study in Southern California. *Transp. Res. Rec. J. Transp. Res. Board* **1994**, *1433*, 170–176.
44. Garikapati, V.M.; Pendyala, R.M.; Morris, E.A.; Mokhtarian, P.L.; McDonald, N. Activity patterns, time use, and travel of millennials: A generation in transition? *Transp. Rev.* **2016**, *36*, 558–584. [CrossRef]
45. Zhong, L.; Lee, B. Carless or car later?: Declining car ownership of millennial households in the Puget Sound Region, Washington State. *Transp. Res. Rec. J. Transp. Res. Board* **2017**, *2664*, 69–78. [CrossRef]
46. Schrank, D.; Eisele, B.; Lomax, T.; Bak, J. *2015 Urban Mobility Scorecard*; Texas A&M Transportation Institute: College Station, TX, USA, 2015.
47. Urban Congestion Reports. Available online: https://ops.fhwa.dot.gov/perf_measurement/ucr/ (accessed on 27 November 2018).
48. Barth, D. *The Bright Side of Sitting in Traffic: Crowdsourcing Road Congestion Data*; Google Official Blog: Mountain View, CA, USA, 2009.
49. Google Maps Directions API. Available online: https://developers.google.com/maps/documentation/directions/start (accessed on 27 November 2018).
50. Kwan, M.-P. Space-time and integral measures of individual accessibility: A comparative analysis using a point-based framework. *Geogr. Anal.* **1998**, *30*, 191–216. [CrossRef]
51. Weber, J.; Kwan, M.-P. Bringing time back in: A study on the influence of travel time variations and facility opening hours on individual accessibility. *Prof. Geogr.* **2002**, *54*, 226–240. [CrossRef]
52. Kim, S.H.; Chung, J.-H. Exploration on origin–destination-based travel time variability: Insights from Seoul metropolitan area. *J. Transp. Geogr.* **2018**, *70*, 104–113. [CrossRef]
53. Weiss, D.J.; Nelson, A.; Gibson, H.S.; Temperley, W.; Peedell, S.; Lieber, A.; Hancher, M.; Poyart, E.; Belchior, S.; Fullman, N.; et al. A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **2018**, *553*, 333–336. [CrossRef]
54. Hu, Y.; Downs, J. Measuring and visualizing place-based space-time job accessibility. *J. Transp. Geogr.* **2019**, *74*, 278–288. [CrossRef]

55. Levinson, H.S.; Lomax, T.J. Developing a travel time congestion index. *Transp. Res. Rec. J. Transp. Res. Board* **1996**, *1564*, 1–10. [CrossRef]
56. Dons, E.; Int Panis, L.; Van Poppel, M.; Theunis, J.; Willems, H.; Torfs, R.; Wets, G. Impact of time–activity patterns on personal exposure to black carbon. *Atmos. Environ.* **2011**, *45*, 3594–3602. [CrossRef]
57. Hägerstrand, T. What about people in Regional Science? *Pap. Reg. Sci. Assoc.* **1970**, *24*, 6–21. [CrossRef]
58. Kim, H.-M.; Kwan, M.-P. Space-time accessibility measures: A geocomputational algorithm with a focus on the feasible opportunity set and possible activity duration. *J. Geogr. Syst.* **2003**, *5*, 71–91. [CrossRef]
59. Novaco, R.W.; Stokols, D.; Milanesi, L. Objective and subjective dimensions of travel impedance as determinants of commuting stress. *Am. J. Commun. Psychol.* **1990**, *18*, 231–257. [CrossRef]
60. Bergstad, C.J.; Gamble, A.; Gärling, T.; Hagman, O.; Polk, M.; Ettema, D.; Friman, M.; Olsson, L.E. Subjective well-being related to satisfaction with daily travel. *Transportation* **2011**, *38*, 1–15. [CrossRef]
61. Wener, R.; Evans, G.W.; Boately, P. Psychophysiological effects of a trip and spillover into the workplace. *Transp. Res. Rec. J. Transp. Res. Board* **2005**, *1924*, 112–117.
62. Vhaduri, S.; Ali, A.; Sharmin, M.; Hovsepian, K.; Kumar, S. Estimating drivers' stress from GPS traces. In Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications—AutomotiveUI '14, Seattle, WA, USA, 17–19 September 2014; pp. 1–8.
63. Susilo, Y.O.; Liotopoulos, F.K. Measuring door-to-door journey travel satisfaction with a mobile phone app. In *Quality of Life and Daily Travel*; Friman, M., Ettema, D., Olsson, L.E., Eds.; Applying Quality of Life Research; Springer International Publishing: Cham, Switzerland, 2018; pp. 119–138. ISBN 978-3-319-76623-2.

International Journal of
Environmental Research and Public Health

MDPI

Article

Elimination of *Schistosoma japonicum* Transmission in China: A Case of Schistosomiasis Control in the Severe Epidemic Area of Anhui Province

Linhan Li [1,2,3,†], Yibiao Zhou [1,2,3,†], Tianping Wang [4], Shiqing Zhang [4], Gengxin Chen [5], Genming Zhao [1], Na He [1], Zhijie Zhang [1], Dongjian Yang [1,2,3], Ya Yang [1,2,3], Yu Yang [1,2,3], Hongchang Yuan [1], Yue Chen [6] and Qingwu Jiang [1,2,3,*]

1 Department of Epidemiology, School of Public Health, Fudan University, 130 Dong'an Road, Shanghai 200032, China; 16111020005@fudan.edu.cn (L.L.); ybzhou@fudan.edu.cn (Y.Z.); gmzhao@shmu.edu.cn (G.Z.); nhe@shmu.edu.cn (N.H.); zhj_zhang@fudan.edu.cn (Z.Z.); 17211020111@fudan.edu.cn (D.Y.); yayang16@fudan.edu.cn (Y.Y.); 16211020071@fudan.edu.cn (Y.Y.); hcyfduepi@163.com (H.Y.)
2 Key Laboratory of Public Health Safety, Ministry of Education, 130 Dong'an Road, Shanghai 200032, China
3 Center of Tropical Disease Research, 130 Dong'an Road, Shanghai 200032, China
4 Institute for the Prevention and Control of Schistosomiasis in Anhui Province, 337 Wuhu Road, Hefei 230061, China; wangtianping@hotmail.com (T.W.); zhangdj1994@aliyun.com (S.Z.)
5 Guichi Antischistosomiasis Station, West Qiupu Road, Chizhou 247100, China; cgx5611@163.com
6 School of Epidemiology and Public Health, University of Ottawa, 600 Peter Morand, Ottawa, ON K1G 5Z3, Canada; Yue.Chen@uottawa.ca
* Correspondence: jiangqw@fudan.edu.cn
† These authors contributed equally to this work.

Received: 30 November 2018; Accepted: 2 January 2019; Published: 7 January 2019

Abstract: Over the several decades, China has been incessantly optimizing control strategies in response to the varying epidemic situations of schistosomiasis. We evaluated continuously the changing prevalence under different control strategies of two villages, Sanlian and Guifan, in China through five phases lasting 37 years. We tested residents, calculated prevalence and discussed change causes. We found the prevalence in Sanlian did not differ significant from that of Guifan ($p = 0.18$) in 1981, but decreased to 2.66%, much lower than Guifan's 11.25%, in 1984 ($p = 0$). Besides, prevalence in Guifan increased to 21.25% in 1987, while in Sanlian it rose to 20.78% until 1989. Those data confirmed that praziquantel combined with snail control could better reduce the prevalence. From 1992 to 1994, the prevalence in the two villages displayed downtrends, which showed the World Bank Loan Project worked. From 1995 to 2004, repeated oscillations with no obvious change trend was seen. Since 2005, the prevalence in both villages has shown a significant downtrend ($p < 0.05$), which suggests the integrated strategy is effective. We considered the control strategies were implemented suitably in the study area under changing social circumstances. Adjusting the strategy in consideration of social transformations is necessary and vital. The experience may be useful for policy making of other epidemic areas with an analogous situation.

Keywords: evaluation; method; policy; strategy; social; trend

1. Introduction

The archeological discovery of Schistosoma eggs in a corpse in the Changsha Mawangdui Han tomb in Hunan Province revealed that schistosomiasis has been prevalent in China for over 2000 years. Not until the American physician Logan first reported the clinical diagnosis of schistosomiasis in China [1], did Chinese people begin to consider schistosomiasis as a disease. *Schistosoma* *japonicum*, one of the five species of schistosome, is widely distributed in China, Indonesia, and The Philippines [2]. As it needs to go through several growth stages in water, the transmission and infection of humans is closely related to specific geographic conditions, which causes difficulties in prevalence control [3]. Schistosomiasis brings a heavy burden to China's population health, social stability, and economic development.

After the establishment of the People's Republic of China in 1949, the central government organized and delegated expert teams to conduct a national survey that found 12 provinces and 351 counties where schistosomiasis prevailed [4] with some 10.5–11.8 million people infected. Zedong Mao, the Chinese president back then, with immense political influence, wrote a famous poem "Farewell to the god of plague" and launched a massive country-wide movement to eliminate schistosomiasis [5]. With medical treatment, environmental modification, and molluscicide application, by 1980, four out of the 12 provinces and two thirds of the affected counties successfully repressed the prevalence of schistosomiasis [6]. As for the remaining eight provinces, epidemic areas were mostly distributed in marshlands and mountainous areas with complicated environmental factors that hindered control.

In order to address this predicament and search for a new and effective schistosomiasis control method, Warren and Su conducted a study in Guichi district, Anhui Province, one of the ten historically most severely affected areas in China [7] and confirmed praziquantel therapy could effectively reduce the schistosomiasis prevalence in humans. When combined with snail control in epidemic areas, the prevalence decreased even more significantly. In this study, we performed a 37-years longitudinal study from 1981 to 2017 on the basis of the study conducted by Warren and Su to understand the long-term effects of schistosomiasis transmission control strategies.

2. Materials and Methods

2.1. Study Area

The study was conducted in a marshland area including Sanlian village (30.642447, 117.346828 and Guifan village (30.574982, 117.410960), in Guichi district, Anhui Province, which is located in the middle and lower reaches of the Yangtze River Basin, the largest river in China (Figure 1).

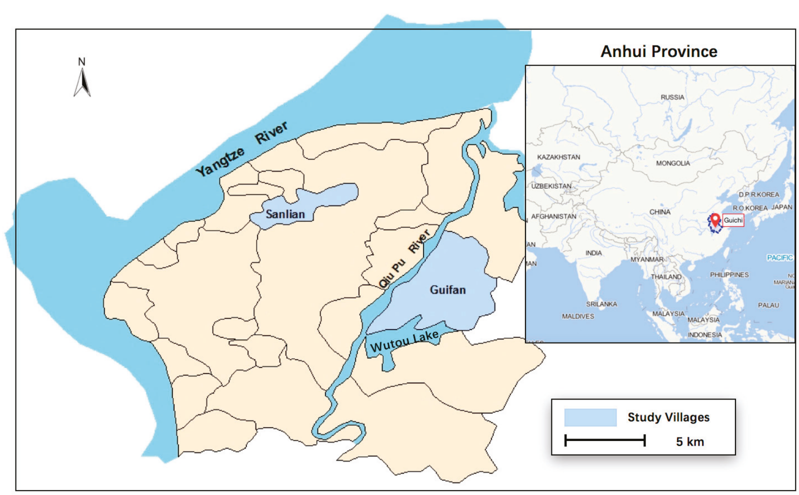


Figure 1. Location of the study area in Guichi district, Anhui Province, China.

Its humid subtropical monsoon climate has annual average temperatures of 16 degrees Celsius and annual average rainfall of 1600 mm. This provides a suitable environment for the survival of Schistosoma japonicum. There are many small hills full of swamps, ponds and ditches in Guichi. This feature makes Guichi a typical schistosomiasis epidemic area. In addition, people in Guichi live along rivers or lakes, and their lifestyles area mainly involve farming and grazing, fishing and lake grassing. These labor activities expose local residents to be infected by schistosomiasis [7]. At the end of 2012, the total habitat area of snails in Guichi district was approximately 26.2715 km^2 [8]. Of 650,000 people in Guichi, about 400,000 people were at risk of schistosomiasis, and the endemic area covered 16 towns and 165 villages [8]. All subjects gave their informed consent for inclusion before they participated in the study. The research was approved by the Medical Ethics Committee of Fudan University School of Public Health (IRB#2011-03-0270).

2.2. Intervention Measures

Sanlian village applied molluscicide (solid pentachlorophenol) on snails' habitats every spring during 1981–1984, and simultaneously administered praziquantel to people with positive fecal examination. Guifan village only administered praziquantel to subjects with positive fecal tests. After 1984, the different control measures were all stopped, and two villages began to follow the same control strategies: (1) by local government in 1985–1991. (2) recommended by the World Bank Loan Project (WBLP) in 1992–2001. (3) by the continued WBLP measures in 2002–2004. (4) since 2005, the integrated schistosomiasis control strategy has been carried out. Table 1 provides detailed information on the intervention measures from 1981 to 2017.

2.3. Disease Diagnosis

During the study period, after the transmission season (October–November), residents aged 5 to 65 years (743 in Sanlian, 1479 in Guifan in 1981) were tested by Kato-Katz (KK) fecal test and/or indirect hemagglutination assay (IHA) [9]. After 1981, the number of examined people was about 600 to 1200 in Sanlian and Guifan, respectively. From 1981 to 1994, all subjects in the two villages were tested by the two Kato-Katz thick smears test. During the period of 1995–2004, subjects were screened by indirect hemagglutination assay method. From 2005 to 2014, all subjects were tested by three Kato-Katz thick smears test. All the subjects were screened by indirect hemagglutination assay and the seropositive were confirmed by Kato-Katz three smears test from 2015 to 2017.

2.4. Statistical Analysis

The prevalence of human infection was estimated by the formula: (1) Observed prevalence (%) = No. of egg-positive/No. of examined people × 100% (by using Kato-Katz method). (2) Observed prevalence (%) = (No. of blood-positive/No. of examined people × 100% (by using IHA test). (3) Observed prevalence (%) = (No. blood-positive/No. of examined people by using IHA test) × (No. of egg-positive/No. of examined people by using Kato-Katz method) × 100% (by using Kato-Katz and IHA).

The adjusted prevalence is the observed prevalence of human infection after taking the sensitivity and specificity of the Kato-Katz method into consideration. The adjusted formula: $p' = \frac{P_0+\beta-1}{\alpha+\beta-1}$ p': adjusted prevalence, p_0: observed prevalence, α: sensitivity, β: specificity [10].

Two stool samples with three thick smears each are used as the diagnostic gold standard [11]. The sensitivities of two smears and three smears Kato-Katz methods are 64% and 75% separately. The specificities are 100% for both of them; the sensitivity and specificity of IHA and the observed prevalence measured by it are not satisfy the condition of adjusted formula [10], therefore, the results of IHA cannot be adjusted.

All data were entered and built a database with Microsoft Excel, and statistical analyses were performed with the use of SPSS (version25, IBM, New York, NY, USA) and SAS software (University Edition, The Statistical Analysis System, Raleigh, NC, USA). We used the chi-square test, including testing for trend when appropriate, to examine any difference and change over time. Two-sided p values were calculated for all comparisons [12–14].

Table 1. The varieties of control strategies of schistosomiasis from 1981–2017 in Guifan and Sanlian villages in China.

Time	Phase	Specific Methods
1981–1984	Warren's Study Control	Guifan Village: People with a positive stool examination result should receive drug chemotherapy. Sanlian Village: 1. People with positive result of stool examination should receive drug chemotherapy. 2. The areas where snail lived were sprayed with molluscicide every spring.
1985–1991	Local Government Control	1. Chemotherapy for people and cattle was the basic measure. 2. Health education for people. 3. Molluscicide was directed at the susceptible areas.
1992–2001	World Bank Loan Project	Goal: Enhanced morbidity control through praziquantel to human and bovines. Details: 1. Chemotherapy was complemented by health education. 2. Snails were controlled by environmental management. 3. Molluscicide was a key way to sustain transmission control.
2002–2004	Extended Period of WBLP Control	1. Achievements from WBLP were not reinforced. 2. The extent of chemotherapy shrinked and compliance of chemotherapy became worse. 3. Returning farmland to lake was advocated due to the molluscicide and previously reclaimed land from the lake.
2005–2017	Integrated Strategy Control	2005: 1. Chemotherapy with praziquantel for people and cattle. 2. Snail control with molluscicides. 3. Health education programmes—residents should stay away from snail-infested areas and water Since 2006: 4. All cattle were replaced with mechanized equipment. 5. Other domestic animals were fenced in. 6. Piped water and lavatories were supplied to improve sanitation.

3. Results

Infection in Humans

The trend of schistosomiasis infection rates in the residents of Sanlian and Guifan villages over the last 37 years was shown in Figure 2. In 1981, the prevalence of schistosomiasis was 32.81% in Sanlian village and 30.00% in Guifan village (χ^2 = 1.839, p = 0.175). With the intervention measures of human chemotherapy and snail control in Sanlian village, the prevalence dropped to 2.66% in 1984, while in Guifan village with chemotherapy only, the prevalence was only reduced to 11.25% (χ^2 = 41.576, p = 0). By 1987, the prevalence in Guifan village increased to 21.25%, while the rate in Sanlian village rose to 20.78% until 1989. Since 1987, the local government in Guifan village has strengthened its control work so the rate declined again. Since the implementation of WBLP in 1992, the prevalence in Sanlian and Guifan villages had continued to decrease from 12.51% and 14.10% in 1992 to 3.29% and 2.89% in 1994. There was a volatile stage, with no obvious change trend in both villages from 1995 to 2004. In 1998 and 1999, infection rates in Sanlian and Guifan reached their respective peaks. A comprehensive control strategy has been implemented since 2005, and the prevalence decreased from 8.12% in 2005 to 0 in 2014 in Sanlian and from 14.39% to 0.70% for the same period in Guifan. There was no infected person detected in either village from 2015 to 2017. Both prevalences had significant downtrends since 2005 (Sanlian χ^2 = 7.995, p = 0.005; Guifan χ^2 = 9.433, p = 0.002).

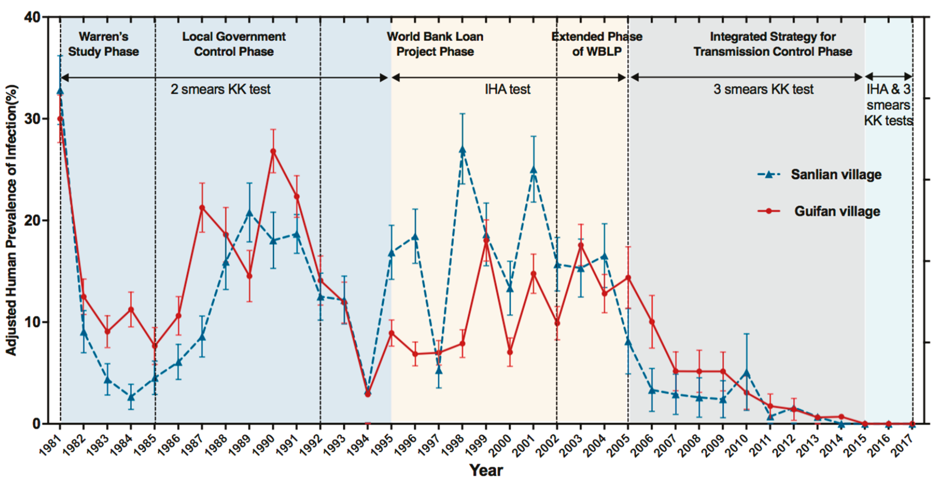


Figure 2. Prevalence of *Schistosoma japonicum* infection in humans in two villages during 1981–2017. (1) The adjusted prevalence formula: $p' = \frac{P_0+\beta-1}{\alpha+\beta-1}$ p': Adjusted prevalence; p_0: Observed prevalence; α: Sensitivity; β: Specificity [10]; (2) The Kato-Katz (KK) 2 slides method was used in 1981–1994: α = 0.64, β = 1 [11]; The indirect hemagglutination assay (IHA) method was used in 1995–2004, but IHA not meet the condition of adjusted formula [10]; The Kato-Katz 3 slides method was used in 2005–2017, α = 0.75, β = 1 [11]; (3) Prevalence were assessed after the transmission season for each year in Sanlian and Guifan villages. (4) The I bars represent 95% confidence intervals.

4. Discussion

This study provides a long and continuous evaluation of the prevalence of schistosomiasis in China. Using disease data from 1981 to 2017, we found the changes of infection rates of humans within each historical phase were on account of different and complex causes, particularly social, economic, political, and health aspects. The schistosomiasis control strategies were suitable for the respective situations back then. Compared to most other studies concentrating on a single epidemic period or data from specific time points, infection rates and control strategies of schistosomiasis in this study are evaluated in a more comprehensive, systematic, and direct way. We confirmed that snail control and chemotherapy in combination has better effect than chemotherapy alone. In addition, we esteemed that the morbidity control measures taken in each historical period were appropriate and suited to the changing disease and social-economic status situations in China. We believed that WBLP is a successful example of disease control aided by foreign funds. Besides, we concluded the integrated strategy has greatly reduced the disease burden of schistosomiasis in China.

After different control measures were stopped in 1984, the prevalence in Sanlian village started to rise the next year, but still remained lower than that of the Guifan village until 1988. This suggests chemotherapy combined with snail control achieved robust and longer-term effects in reducing the prevalence. This conclusion corroborates control projects undertaken in other countries [15,16]. However, from 1987 to 1992, the prevalence oscillated in both villages, which was similar to other regions in China at that time. There are several possible reasons. First of all, since 1978, China's rural areas have begun to shift from a centralized system to a contracted responsibility system for joint production. The mass extermination of snails is no longer practical. Moreover, as a result of responsibility privatization and the subsequent rise of labor cost, the previous schistosomiasis control measures become too expensive. Second of all, with the reform of China's social and economic system in the 1980s, the source of schistosomiasis control budgets shifted from the previous three-level system (county, provincial, and the central governments) to the local government only, significantly intensifying the burden on local finances and therefore limiting corresponding expenditure on schistosomiasis control. Third of all, previous drug eradication of snails and environmental modification brought many problems such as environmental pollution and flooding, further increased economic burden and social impact [17]. Last but not least, praziquantel, a very effective chemotherapy drug started to be introduced to and widely used in China in the late 1980s [3], the Chinese government and the public dedicated less attention to schistosomiasis, which also partly led to the oscillation in prevalence.

In 1992, large-scale drug treatment, health education and snails control were implemented in both villages, which led to obviously prevalence declined in the next two years. In 1990s, following the WHO's recommendation that developing countries should utilize chemotherapy to achieve better morbidity control [18,19], a wide application of praziquantel, indirect hemagglutination, and the Kato-Katz method were initiated. This marked the starting point when China's schistosomiasis control strategy switched from environmental modification to praziquantel-based control. Particularly, in 1992, with the help of the WHO, China began to implement the World Bank Loan Project (No. P003624) for "Infectious and Endemic Disease Control Project" [20,21], which provided 71 million dollars to help and support the government's financial expenditure on schistosomiasis control, the application of parasite diagnostic techniques and the promotion of chemotherapeutic drugs [22]. Following a sharp decline in 1992–1994, the prevalence in the two villages rebounded in 1995 due to flooding in Guichi district. On the contrary, in 1997, the drought in Guichi decreased the risk of infection and therefore the prevalence. Similarly, prevalence peaked again in 1998 and 1999 as a result of severe flooding in China's Yangtze River Basin that year [23]. Many new snail areas also appeared after the flood and this aggravated the disease burden [24]. As shown in Figure 2, prevalence in two villages oscillated at a high level from 1995 to 2004 even with continuous praziquantel-based control, indicating human chemotherapy alone might not be sufficient to completely eliminate schistosomiasis. This was also supported by another study [25]. Fluctuation of prevalence in the late WBLP period was also confirmed by another study [26]. There are four possible reasons for the changing prevalence at that time. Firstly, the reform of government institutions and the loss of well-trained technicians caused difficulties in normal operation of schistosomiasis control and budget reduction. Secondly, with the termination of the WBLP, the public's compliance [27] with drug treatment had also decreased. Thirdly, although the prevalence and incidence of acute schistosomiasis dropped to a relatively low level in the early stage of this period [28], praziquantel could not stop transmission in high-epidemic areas and reinfection [3]. Lastly, natural factors such as floods also contributed heavily to the prevalence fluctuation.

From 2005 to 2017, the infection rates in both villages decreased significantly. This might result from four aspects. First of all, in 2004, Chinese government set up four priority control programs of communicable diseases including schistosomiasis, AIDS, tuberculosis and hepatitis B [14], and schistosomiasis received an increased attention. Second of all, Chinese government launched the "National long-term plan for control of schistosomiasis (2004–2015)" program [29] in response to the rebounded trend of prevalence, which marked the beginning of the integrated control strategy in the country. Third of all, as the government increased its public health budget for schistosomiasis, the government transfer payment program was started and the national health system was strengthened. Lastly, the comprehensive control strategy, including praziquantel treatment of human and animals, molluscide application, health education, sanitation improvement, removal of bovines, etc., were implemented in our study area in 2006 [30] which greatly helped to reduce the disease burden of schistosomiasis. This result is consistent with the conclusions of many other studies [6,30–38]. However, several issues remain as well. Firstly, the comprehensive control strategy is not able to further reduce the infection rate to less than 3% in some areas [30,39]. Secondly, although removal of agriculture main livestock, such as bovines, sheep, and horses, can reduce human prevalence to a large extent [24], the existing control measures need to be adjusted in China since 45 species of mammal can be infected by *S. japonicum* in total, including but not limited to dogs, cats, and rodents [40–42]. Thirdly, the broadly used Kato-Katz method has relatively low sensitivity, particularly for subjects with low infectiosity [43,44]. This could lead to false negative results. Recently, several studies confirmed that Helmintex method is a more sensitive egg detection procedure than Kato-Katz method and other diagnostic tests like the point-of-care immunodiagnostic for detecting schistosome cathodic circulating antigen (POC-CCA) and Saline gradient [45,46]. Perhaps widely used of Helmintex in China is feasible and necessary. Fourthly, impact of water conservancy project on schistosomiasis control still remains controversial, due to many influencing factors such as water level, temperature, and plants change [47]. Lastly, environmental change, migration of population, and effectiveness of drug treatment could also influence the transmission of *S. japonicum*. Further investigations and effective long-term policies in controlling these factors are needed.

The main limitation of our study is that the two villages are both typical bottomland areas and cannot represent all types of epidemic areas. In addition, a comparative study of the two villages was only run from 1981 to 1984. Therefore, there is no comparison group afterwards. Moreover, the study used three different diagnostic methods: two slides Kato-Katz, indirect hemagglutination assay (IHA), and three slides Kato-Katz, however, infection rate tested by IHA cannot be adjusted with its sensitivity and specificity [10].

5. Conclusions

Our results demonstrated that Guichi district has implemented suitable schistosomiasis control strategies for the respective social circumstances in different periods. This study suggested that adjusting the control strategy in according to social, economic, and political changes is necessary and vital. Schistosomiasis control in Guichi district is one of the successful examples in schistosomiasis control over the past 40 years in China. It could provide valuable experience on future policy making not only for China, but also epidemic areas with similar social conditions around the world.

Author Contributions: L.L., Y.Z., and Q.J. participated in structuring this manuscript. T.W., S.Z. and G.C. collected the data. L.L. and Y.Z. analyzed the data and wrote the first draft. G.Z., N.H., Z.Z. and H.Y. provided advices for the subject of this manuscript. L.L., Y.Z., D.Y., Y.Y. (Ya Yang) and Y.Y. (Yu Yang) searched the scientific literature. L.L., Y.Z., Y.C. and Q.J. revised the manuscript. All authors contributed to, reviewed, and approved the final draft of the paper.

Funding: This manuscript was supported by National office for philosophy and social sciences for the Chinese Social Science Fund Major Project (project number: 16ZDA237).

Conflicts of Interest: The authors declare no conflict of interest. The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

References

1. Logan, O.T. A case of dysentery in hunan province caused by the trematode *Schistosoma japonicum*. *China Mission. Med. J.* **1905**, *19*, 243–245.
2. WHO. Schistosomiasis. Different Type of Schistosome. Available online: https://www.who.int/schistosomiasis/epidemiology/table3/en/ (accessed on 11 October 2018).
3. King, C.H. Toward the elimination of schistosomiasis. *N. Engl. J. Med.* **2009**, *360*, 106–109. [CrossRef] [PubMed]
4. Mao, S.P.; Shao, B.R. Schistosomiasis control in the People's Republic of China. *Am. J. Trop. Med. Hyg.* **1982**, *31*, 92–99. [CrossRef] [PubMed]
5. Wang, L.; Utzinger, J.; Zhou, X.N. Schistosomiasis control: Experiences and lessons from China. *Lancet* **2008**, *372*, 1793–1795. [CrossRef]
6. Zhou, X.N.; Wang, L.Y.; Chen, M.G.; Wu, X.H.; Jiang, Q.W.; Chen, X.Y.; Zheng, J.; Utzinger, J. The public health significance and control of schistosomiasis in China—Then and now. *Acta Trop.* **2005**, *96*, 97–105. [CrossRef]
7. Warren, K.S.; Su, D.L.; Xu, Z.Y.; Yuan, H.C.; Peters, P.A.; Cook, J.A.; Mott, K.E.; Houser, H.B. Morbidity in schistosomiasis japonica in relation to intensity of infection. A study of two rural brigades in Anhui province, China. *N. Engl. J. Med.* **1983**, *309*, 1533–1539. [CrossRef]
8. Qian, R.F.; Han, S.M.; Ke, Y.F. The epidemic situation of schistosomiasis in Guichi District, Chizhou City, in 2008–2012. *J. Trop. Dis. Parasitol.* **2014**, *12*, 108–110. (In Chinese)
9. Katz, N.; Chaves, A.; Pellegrino, J. A simple device for quantitative stool thick-smear technique in schistosomiasis mansoni. *Rev. Inst. Med. Trop. Sao Paulo* **1972**, *14*, 397–400.
10. Rogan, W.J.; Gladen, B. Estimating prevalence from the results of a screening test. *Am. J. Epidemiol.* **1978**, *107*, 71–76. [CrossRef]
11. Lin, D.D.; Liu, J.X.; Liu, Y.M.; Hu, F.; Zhang, Y.Y.; Xu, J.M.; Li, J.Y.; Ji, M.J.; Bergquist, R.; Wu, G.L.; et al. Routine Kato-Katz technique underestimates the prevalence of *Schistosoma japonicum*: A case study in an endemic area of the People's Republic of China. *Parasitol. Int.* **2008**, *57*, 281–286. [CrossRef]
12. Bosso, L.; Lacatena, F.; Varlese, R.; Nocerino, S.; Cristinzio, G.; Russo, D. Plant pathogens but not antagonists change in soil fungal communities across a land abandonment gradient in a mediterranean landscape. *Acta Oecol.* **2017**, *78*, 1–6. [CrossRef]
13. Steinbaum, L.; Kwong, L.H.; Ercumen, A.; Negash, M.S.; Lovely, A.J.; Njenga, S.M.; Boehm, A.B.; Pickering, A.J.; Nelson, K.L. Detecting and enumerating soil-transmitted helminth eggs in soil: New method development and results from field testing in Kenya and Bangladesh. *PLoS Negl. Trop. Dis.* **2017**, *11*, e0005522. [CrossRef]
14. Wang, L.D.; Chen, H.G.; Guo, J.G.; Zeng, X.J.; Hong, X.L.; Xiong, J.J.; Wu, X.H.; Wang, X.H.; Wang, L.Y.; Xia, G.; et al. A strategy to control transmission of *Schistosoma japonicum* in China. *N. Engl. J. Med.* **2009**, *360*, 121–128. [CrossRef] [PubMed]
15. Katz, N.; Rocha, R.S.; Pereira, J.P. Schistosomiasis control in Peri-Peri (Minas Gerais, Brazil) by repeated clinical treatment and molluscicide application. *Rev. Inst. Med. Trop. Sao Paulo* **1980**, *22*, 85–93, 203–211.
16. McCullough, F.S.; Gayral, P.; Duncan, J.; Christie, J.D. Molluscicides in schistosomiasis control. *Bull. World Health Organ.* **1980**, *58*, 681–689. [PubMed]
17. Spear, R.C.; Seto, E.; Remais, J.; Carlton, E.J.; Davis, G.; Qiu, D.; Zhou, X.; Liang, S. Fighting waterborne infectious diseases. *Science* **2006**, *314*, 1081–1083. [CrossRef] [PubMed]
18. WHO. *The Control of Schistosomiasis: Report of a WHO Expert Committee*; World Health Organization: Geneva, Switzerland, 1985.
19. WHO. *The Control of Schistosomiasis: Second Report of the WHO Committee*; World Health Organization: Geneva, Switzerland, 1993.
20. Bank, T.W. Project: Infectious and Endemic Disease Control Project. Available online: http://projects.worldbank.org/P003624/infectious-endemic-disease-control-project?lang=en (accessed on 10 May 2018).
21. Utzinger, J.; Bergquist, R.; Shu-Hua, X.; Singer, B.H.; Tanner, M. Sustainable schistosomiasis control—The way forward. *Lancet* **2003**, *362*, 1932–1934. [CrossRef]

22. Chen, X.Y.; Wang, L.Y.; Cai, J.M.; Zhou, X.N.; Zheng, J.; Guo, J.G.; Wu, X.H.; Engels, D.; Chen, M.G. Schistosomiasis control in China: The impact of a 10-year world bank loan project (1992–2000). *Bull. World Health Organ.* **2005**, *83*, 43–48.
23. Wu, X.H.; Zhang, S.Q.; Xu, X.J.; Huang, Y.X.; Steinmann, P.; Utzinger, J.; Wang, T.P.; Xu, J.; Zheng, J.; Zhou, X.N. Effect of floods on the transmission of schistosomiasis in the Yangtze River valley, People's Republic of China. *Parasitol. Int.* **2008**, *57*, 271–276. [CrossRef]
24. Yang, G.J.; Liu, L.; Zhu, H.R.; Griffiths, S.M.; Tanner, M.; Bergquist, R.; Utzinger, J.; Zhou, X.N. China's sustained drive to eliminate neglected tropical diseases. *Lancet Infect. Dis.* **2014**, *14*, 881–892. [CrossRef]
25. Zhou, Y.B.; Zhao, G.M.; Jiang, Q.W. Effects of the praziquantel-based control of schistosomiasis japonica in China. *Ann. Trop. Med. Parasitol.* **2007**, *101*, 695–703. [CrossRef] [PubMed]
26. Zhou, X.N.; Wang, L.Y.; Chen, M.G.; Wang, T.P.; Guo, J.G.; Wu, X.H.; Jiang, Q.W.; Zheng, J.; Chen, X.Y. An economic evaluation of the national schistosomiasis control programme in China from 1992 to 2000. *Acta Trop.* **2005**, *96*, 255–265. [CrossRef]
27. Jiang, Q.W.; Wang, L.Y.; Guo, J.G.; Chen, M.G.; Zhou, X.N.; Engels, D. Morbidity control of schistosomiasis in China. *Acta Trop.* **2002**, *82*, 115–125.
28. Guo, J.G.; Cao, C.L.; Hu, G.H.; Lin, H.; Li, D.; Zhu, R.; Xu, J. The role of 'passive chemotherapy' plus health education for schistosomiasis control in China during maintenance and consolidation phase. *Acta Trop.* **2005**, *96*, 177–183. [CrossRef] [PubMed]
29. Zhou, X.N.; Jiang, Q.W.; Wang, T.P.; Lin, D.D.; Wu, X.H. Status and strategy for research development of schistosomiasis control in China. *Chin. J. Schistosomiasis Control* **2005**, *17*, 1–3.
30. Zhou, Y.B.; Liang, S.; Chen, G.X.; Rea, C.; He, Z.G.; Zhang, Z.J.; Wei, J.G.; Zhao, G.M.; Jiang, Q.W. An integrated strategy for transmission control of *Schistosoma japonicum* in a marshland area of China: Findings from a five-year longitudinal survey and mathematical modeling. *Am. J. Trop. Med. Hyg.* **2011**, *85*, 83–88. [CrossRef] [PubMed]
31. Liu, R.; Dong, H.F.; Jiang, M.S. The new national integrated strategy emphasizing infection sources control for schistosomiasis control in China has made remarkable achievements. *Parasitol. Res.* **2013**, *112*, 1483–1491. [CrossRef] [PubMed]
32. Yang, Y.; Zhou, Y.B.; Song, X.X.; Li, S.Z.; Zhong, B.; Wang, T.P.; Bergquist, R.; Zhou, X.N.; Jiang, Q.W. Integrated control strategy of schistosomiasis in the People's Republic of China: Projects involving agriculture, water conservancy, forestry, sanitation and environmental modification. *Adv. Parasitol.* **2016**, *92*, 237–268.
33. Zhang, Z.; Zhu, R.; Ward, M.P.; Xu, W.; Zhang, L.; Guo, J.; Zhao, F.; Jiang, Q. Long-term impact of the world bank loan project for schistosomiasis control: A comparison of the spatial distribution of schistosomiasis risk in China. *PLoS Negl. Trop. Dis.* **2012**, *6*, e1620. [CrossRef]
34. Zhang, S.Q.; Sun, C.S.; Wang, M.; Lin, D.D.; Zhou, X.N.; Wang, T.P. Epidemiological features and effectiveness of schistosomiasis control programme in lake and marshland region in the People's Republic of China. *Adv. Parasitol.* **2016**, *92*, 39–71.
35. Sun, L.P.; Wang, W.; Hong, Q.B.; Li, S.Z.; Liang, Y.S.; Yang, H.T.; Zhou, X.N. Approaches being used in the national schistosomiasis elimination programme in China: A review. *Infect. Dis. Poverty* **2017**, *6*, 55. [CrossRef] [PubMed]
36. Sun, L.P.; Wang, W.; Zuo, Y.P.; Zhang, Z.Q.; Hong, Q.B.; Yang, G.J.; Zhu, H.R.; Liang, Y.S.; Yang, H.T. An integrated environmental improvement of marshlands: Impact on control and elimination of schistosomiasis in marshland regions along the Yangtze River, China. *Infect. Dis. Poverty* **2017**, *6*, 72. [CrossRef]
37. Wang, X.; Wang, W.; Wang, P. Long-term effectiveness of the integrated schistosomiasis control strategy with emphasis on infectious source control in China: A 10-year evaluation from 2005 to 2014. *Parasitol. Res.* **2017**, *116*, 521–528. [CrossRef] [PubMed]
38. Zhou, X.N.; Jiang, T.W.; Guo, J.G.; Wang, L.Y.; Jiang, Q.W. Project management model and its evolution in schistosomiasis control programme of China. *Chin. J. Schistosomiasis Control* **2010**, *22*, 1–4.
39. Wang, L.D.; Guo, J.G.; Wu, X.H.; Chen, H.G.; Wang, T.P.; Zhu, S.P.; Zhang, Z.H.; Steinmann, P.; Yang, G.J.; Wang, S.P.; et al. China's new strategy to block *Schistosoma japonicum* transmission: Experiences and impact beyond schistosomiasis. *Trop. Med. Int. Health* **2009**, *14*, 1475–1483. [CrossRef] [PubMed]

40. Lu, D.B.; Wang, T.P.; Rudge, J.W.; Donnelly, C.A.; Fang, G.R.; Webster, J.P. Contrasting reservoirs for *Schistosoma japonicum* between marshland and hilly regions in Anhui, China—A two-year longitudinal parasitological survey. *Parasitology* **2010**, *137*, 99–110. [CrossRef]
41. Colley, D.G.; Bustinduy, A.L.; Secor, W.E.; King, C.H. Human schistosomiasis. *Lancet* **2014**, *383*, 2253–2264. [CrossRef]
42. Zhou, Y.B.; Liang, S.; Jiang, Q.W. Factors impacting on progress towards elimination of transmission of schistosomiasis japonica in China. *Parasit. Vectors* **2012**, *5*, 275. [CrossRef]
43. Zhou, Y.B.; Yang, M.X.; Wang, Q.Z.; Zhao, G.M.; Wei, J.G.; Peng, W.X.; Jiang, Q.W. Field comparison of immunodiagnostic and parasitological techniques for the detection of schistosomiasis japonica in the People's Republic of China. *Am. J. Trop. Med. Hyg.* **2007**, *76*, 1138–1143. [CrossRef]
44. Wang, W.; Dai, J.R.; Liang, Y.S. Apropos: Factors impacting on progress towards elimination of transmission of schistosomiasis japonica in China. *Parasit. Vectors* **2014**, *7*, 408. [CrossRef]
45. Lindholz, C.G.; Favero, V.; Verissimo, C.M.; Candido, R.R.F.; de Souza, R.P.; Dos Santos, R.R.; Morassutti, A.L.; Bittencourt, H.R.; Jones, M.K.; St Pierre, T.G.; et al. Study of diagnostic accuracy of Helmintex, Kato-Katz, and POC-CCA methods for diagnosing intestinal schistosomiasis in Candeal, a low intensity transmission area in northeastern Brazil. *PLoS Negl. Trop. Dis.* **2018**, *12*, e0006274. [CrossRef]
46. Oliveira, W.J.; Magalhaes, F.D.C.; Elias, A.M.S.; de Castro, V.N.; Favero, V.; Lindholz, C.G.; Oliveira, A.A.; Barbosa, F.S.; Gil, F.; Gomes, M.A.; et al. Evaluation of diagnostic methods for the detection of intestinal schistosomiasis in endemic areas with low parasite loads: Saline gradient, Helmintex, Kato-Katz and rapid urine test. *PLoS Negl. Trop. Dis.* **2018**, *12*, e0006232. [CrossRef] [PubMed]
47. Steinmann, P.; Keiser, J.; Bos, R.; Tanner, M.; Utzinger, J. Schistosomiasis and water resources development: Systematic review, meta-analysis, and estimates of people at risk. *Lancet Infect. Dis.* **2006**, *6*, 411–425. [CrossRef]

International Journal of
Environmental Research and Public Health

Article

Estimation of the Spatial Suitability of Winter Tourism Destinations Based on Copula Functions

Weiying Cai [1], Hui Di [2] and Xingpeng Liu [2,*]

1 Northeast Asia Leisure Economy Research Center, The Tourism College of Changchun University, Changchun 130607, China; weiycai1982@163.com
2 School of Environment, Northeast Normal University, Changchun 130024, China; dih717@nenu.edu.cn
* Correspondence: liuxp912@nenu.edu.cn; Tel.: +86-0431-8916-5610

Received: 4 December 2018; Accepted: 7 January 2019; Published: 10 January 2019

Abstract: Climate and weather are important factors that determine winter tourism destinations and snow resources and temperature affect the income of the winter tourism industry. Against the background of climate change, abnormal fluctuations in climate elements bring a series of challenges for winter tourism and cause potential losses to the tourism industry. To effectively assess and plan winter tourism destinations, this study establishes the snow abundance and meteorological suitability indices from snow resource and weather conditions to express winter tourism resources, respectively. The coupling relationship of the two indices was used to analyze the spatial suitability of winter tourism destinations based on the copula function. By case analysis, it was found that the Frank copula one is the best fitting function for winter tourism suitability analysis. The Yushu–Jiutai–Yitong–Dongliao line is the boundary of spatial suitability in the study area. The eastern areas of the boundary have great potential for winter tourism and could strive to develop ice-snow projects, whereas the western regions are relatively weak. This study has guiding significance for winter tourism destination development and resource spatial layout.

Keywords: winter tourism; climate change; snow abundance index; meteorological suitability index; copula function; spatial suitability assessment

1. Introduction

Climate and weather are important decision making factors for tourists and also influence the successful operation of tourism businesses, especially winter tourism [1–4]. In recent years, winter outdoor tourism activities have gradually gained popularity in China [5]. Both the government and enterprises have increased investment in winter tourism planning. The planning of winter tourism is affected by multiple factors (e.g., climate condition, terrain, land use, etc.); among them, climate and weather are the most important, and changes in these factors will directly affect tourism industry income. Winter tourism is strongly dependent on climate and weather. Under the background of climate change, changing factors lead to changing spatial suitability for winter tourism. To analyze the impact of weather and climate change on the industry, researchers usually develop a comprehensive index to measure the intensity of weather and climate. Weather and climatic index integrate two or multiple meteorological factors that represent weather and climatic characteristics. The weather index represents short-term weather characteristics, while the climate index represents medium and long-term climate characteristics. With the rapid growth of the economy and the increase in property density, the economic losses caused by extreme weather and climate events were increasing significantly. Several kinds of weather and climate indices were proposed by researchers and widely used in agricultural production, energy consumption, commodities, living consumption, health care, tourism, sports and leisure, transportation, insurance and finance and other fields [6–8].

Vast numbers of references have suggested that the China region is extremely sensitive to climate change [9]. Winter tourism is closely related to climate variations, especially in northeast China regions where resorts are heavily dependent on snow. Increasing temperatures and snow scarce winter seasons pose a big challenge for the winter tourism industry [10]. Increasing winter temperatures, lack of snow, decreasing snow cover and snow depth will result in a shorter skiing season [11], which could lead to smaller number of visitors and reduced revenues, and thus have severe economic impacts on winter tourism destinations [12]. Researches have conducted numerous studies on the impact of climate change on winter tourism. Gajićčapka and Srnec [13] analyzed the time analysis (fluctuations and trends) in different meteorological parameters related to snow (air temperature, total precipitation and air pressure) as well as snow parameters (snow cover frequency, duration and magnitude) themselves. Hoffmann et al. [14] proposed an econometric analysis model with linear regression and count data, which found a positive influence of the awareness of possible climate change effects on the scope of corporate adaptation. Damm et al. [10] analyzed the economic effects of technical snow production under future climate conditions. Dingeldey and Soboll [15] presented an interactive multi-agent scenario assessment model and used it to examine the future impact of climate change on winter tourism in the German and Austrian Alps. Gonseth [16] assessed the sensitivity of winter tourism consumption to changing snow conditions. Bonzanigo et al. [17] explored how to effectively integrate a climate change adaptation perspective with local discourses about sustainability and tourism, an increasing priority for policy-makers in the region and elsewhere. Besides, several studies analyzed the potential impact on snowmaking under climate change in the future [11,18–20]. The purposes of researchers want to accurately assess regional impact, vulnerability, risks and opportunities of climate change on winter tourism [21–29]. Therefore, the impact of climate change should be taken into account in winter tourism destinations planning.

To determine spatial suitability for winter tourism, it is very meaningful to analyze the spatial change in winter tourism resources under climate change. Snow resources and meteorological elements are the basic factors of winter tourism. With climate change, winter tourism in some regions of the world has been seriously affected by rising temperatures [2,30]. Climate change studies in Northeast China have proved that regional mean snow cover duration decreased at -2.7 days per decade and mean maximum snow depth decreased at -0.5 cm per decade, while low-temperature days (≤ -25 °C) had a significant decreasing trend of -3.9 days per decade during 1961–2010 [31]. Rising temperature has led to snowfall reduction and a shorter winter tourism season, which has a certain negative impact on snow and ice projects in winter tourism destinations. It has shortened the development season and increased the costs of their tourism products. Therefore, analyzing the temporal and spatial changes in winter tourism resources and allocation can help decision-makers adjust their tourism management strategies.

Estimation of snow resources is one problem for winter tourism destinations. Snow resources include snow depth and snow cover days as two factors. Because of the significant signal difference between snow and other surface objects, snow cover is easily obtained by optical remote sensing such as MODIS data (e.g., MOD10C1). Compared with snow cover, snow depth data are difficult to obtain because of the poor penetration of optical remote sensing. Therefore, snow depth data are often traditionally obtained by weather station observations. With the development of passive microwave technology, passive microwave remote sensing data (e.g., Scanning Multichannel Microwave Radiometer (SMMR) and Special Sensor Microwave/Imager (SSM/I)) were widely applied to obtain snow depth due to their all-weather good penetrating ability, not being affected by cloud cover and high temporal resolution. The snow depth retrieved from passive microwave remote sensing data has been gradually promoted, and the brightness temperature gradient is the most representative algorithm [32]. Several studies have applied snow depth observation data to analyze the nonlinear relationship between snow depth and the brightness temperature gradient [33,34]. Che et al. [33] calculated the snow depth based on land use classification and the brightness temperature gradient and obtained the snow depth in China.

Besides snow resources, meteorological elements are equally important factors for winter tourism. Therefore, the weather index estimation for outdoor activities is another problem for winter tourism. In winter, outdoor snow sports are closely related to weather conditions, and among all the meteorological elements, air temperature and wind speed affect snow sports directly. Furthermore, snow quality is closely related to air temperature. High air temperature makes snow texture sticky and heavy, influencing sliding speed. Low temperature forms brittle ice that causes weak adhesion and easy skidding. For the influence of weather on outdoor sports, the present studies mainly focus on the degree of human body comfort, and the environmental comfort index was analyzed based on sensible human temperatures [35,36]. However, there were fewer studies on the weather conditions for human comfort in sports. In 2017, the Chinese Meteorological Administration issued a standard for the skiing meteorology index [37]. As one of the industries impacted by climate change, several studies analyzed the sensitivity of winter tourism to snow depth and temperature increase over the last decades [2,32,38,39]. The present researchers analyzed the impacts of climate change on winter skiing tourism areas and indicated that future climate change would shorten the skiing season and sharply reduce skiing visitors, causing economic loss in low altitude and low latitude skiable areas [40]. Several studies found that climate change has a negative impact on the sustainable development of winter tourism, skiing tourism and tourism vulnerability [17,41,42]. In the comprehensive evaluation of winter tourism resources, the present study mainly focuses on the qualitative analysis and characteristic description and lacks a precise quantitative analysis of spatial suitability assessment, which limited its application [43–45]. Under the background of climate change, winter tourism industries will face serious challenges. For government managers, how to rationally determine the winter tourism development zone under the influence of climate change has become an important problem that urgently needs to be solved. This has great significance for the sustainable development of the regional tourism industrial economy.

The winter tourism literature indicates that: (1) meteorological conditions and snow resources are two primary factors for winter tourism suitability assessment; however, very few studies have analyzed the temporal and spatial distribution coupling relationship of the two factors as well as the impacts on tourism planning. (2) The present research is mainly focused on the qualitative description of winter resources, lacking quantitative analysis of spatial suitability distribution, and is not suitable for tourism planning decisions. (3) The impact of climate change on winter tourism is uncertain. To avoid the impact of climate fluctuations on winter tourism, it is of great significance to obtain the spatially suitable development areas. To address these problems, this study proposes the meteorological suitability and snow abundance indices for winter tourism based on meteorological and snow data. Firstly, by analyzing the influence of weather conditions on snow cover and outdoor activities, air temperature, wind speed, relative humidity, and visibility are selected to establish the climate suitability index. Secondly, the depth and duration of snow cover are selected to characterize the spatial abundance of regional snow resources, which is used to establish a snow abundance index. To express the coupling effect of these two indices on winter tourism, the spatial joint probability of the two indices is calculated to express the spatial suitability for winter tourism using the Copula function. Finally, taking Jilin province as the area, this study analyzes temporal and spatial distribution and establishes the suitability of winter tourism resources. This study will provide a new method to analyze the impact of climate change on winter tourism. The proposed method and results will have a guiding significance for ice-snow tourism planning in winter tourism destination countries.

2. Materials and Methods

2.1. Study Area

Northeast China is in a high latitude region affected by cold temperate and temperate continental monsoon climates. This area forms long cold winters and short warm summers. The winter tourism industry occupies the dominant financial position for this area. The study area is Jilin province located in the center of northeast China (40°52′–46°18′ N, 121°38′–131°19′ E) (Figure 1). The administration district area is 1.87×10^5 km^2. The winter in Jilin Province lasts half a year. It has significant potential for winter tourism because of abundant snow resources. Compared with other regions in northeast China, the meteorological conditions of Jilin Province are suited for outdoor activities, with good snow quality for snow sports. Therefore, Jilin Province is a major area for winter tourism, and winter tourism has gradually become a new growth pole to pull economic development into the province. During 2016 and 2017, Jilin province received 6.2 million tourists, and tourism income reached 116 billion CNY (Tourism Commission of Jilin Province data). To deeply plan and develop winter tourism resources, it is very necessary to estimate the spatial suitability for winter tourism in this area.

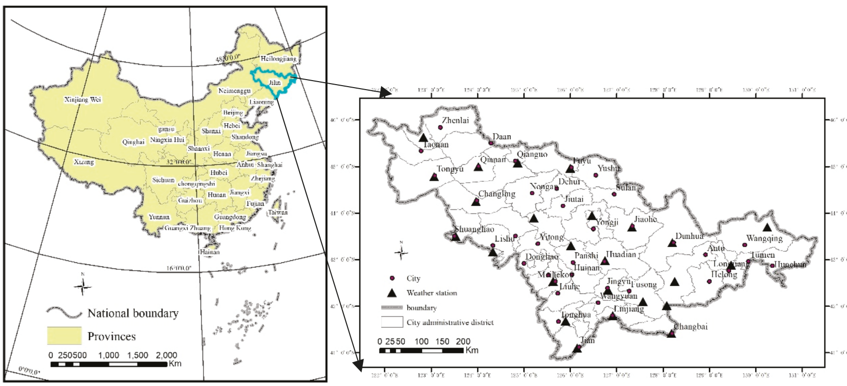


Figure 1. The location of the study area.

2.2. Methods

This study is mainly based on natural conditions (weather and snow) to assess the suitability of regional winter tourism development. To analyze the meteorological conditions and snow resources of Jilin province, the data sources include remote sensing images and local meteorological records. The snow depth data were extracted from the long-term snow depth dataset of China [46]. The original snow data sources are passive microwave images since 1980 obtained from the US National Snow and Ice Data Center (NSIDC). The original long time series dataset of China snow is (1) passive microwave remote sensing SMMR (1980–1987), (2) SSM/I (1987–2007) and (3) SSMI/S (2008–2014). Meteorological data include daily air temperature and wind speed (2 m height) obtained from the Chinese Meteorological Administration (daily data from 1971 to 2016).

2.2.1. Meteorological Suitability Index (MSI)

Air temperature and wind speed are two important factors that directly influence snow quality and outdoor snow sports in winter. When the mean air temperature is about −12 °C, and air relative humidity less than 80%, it will form power snow which is suitable for skiing. The hardness and softness of snow are closely related to air temperature. When the temperature rises, the snow surface will melt gradually under the action of sunlight and the continuous rolling of skis, which cause snow becomes soft; while the extreme cold weather will lead to water condensation and even a thicker ice crystalline layer in the snow, which cause snow to become hard. The present study indicates that when the air temperature is too low (≤−20 °C) tourists feel uncomfortable [37,47,48], and when it is too high (≥2 °C), it influences snow quality and snow cover days [47,49]. Outdoor snow sports are directly restricted by high wind speed. When wind speed reaches a certain level, outdoor activities are limited. When there is wind, the body temperature will be lower. Therefore, when the minimum temperature reaches −16 °C, visitors feel significantly uncomfortable. When the wind-force is less than scale 2 to 3, the skiing is suitable, and when the wind-force is more than scale 5, the skiing will be dangerous. Air relative humidity and visibility also limit outdoor activities [37]. The influence of weather on the spatial and temporal distribution of winter tourism is analyzed by the coupled analysis of air temperature, and wind speed in this study. The classification of a single factor is referred to the meteorological skiing index issued by the Chinese Meteorological Administration [37]. In this study, the fuzzy inference and following steps are used to analyze MSI.

Step 1: Determining the evaluation factors and their grades.

Four factors daily maximum temperature (T), wind speed (W), air relative humidity (H) and visibility (V) were selected to describe meteorological suitability. The factors and their grades are shown in Table 1.

Table 1. The factors and their grades in MSI.

Factors	High	Middle	Low
H(°C)	[−12, −8]	[−16, −12) OR (−8, 2]	>2 OR < −16
W(m/s)	<5.4	[5.4, 10.7]	>5
H(%)	[50, 60]	[30, 50) OR (60, 80]	>80 OR < 30
V(km)	>1	[0.5, 1]	<0.5

Step 2: Constructing the fuzzy membership function of factors

The Gaussian membership function, Z-shaped, and S-shaped membership function were used to establish the fuzzy membership of factors. The symmetric Gaussian function depends on two parameters σ and c as given by:

$$f(x,\sigma,c) = e^{\frac{-(x-c)^2}{2\sigma^2}} \tag{1}$$

Z-shaped and S-shaped membership are spline-based functions. The parameters a and b locate the extremes of the sloped portion of the curve as given by:

$$g(x,a,b) = \begin{cases} 1 & x \leq a \\ 1-2(\frac{x-a}{b-a})^2 & a \leq x \leq \frac{a+b}{2} \\ 2(\frac{x-b}{b-a})^2 & \frac{a+b}{2} \leq x \leq b \\ 0 & x \geq b \end{cases} \tag{2}$$

$$h(x,a,b) = \begin{cases} 0 & x \leq a \\ 2(\frac{x-a}{b-a})^2 & a \leq x \leq \frac{a+b}{2} \\ 1-2(\frac{x-b}{b-a})^2 & \frac{a+b}{2} \leq x \leq b \\ 1 & x \geq b \end{cases} \tag{3}$$

The parameters σ, c, a and b will be established based on the boundary value in Table 1.

Step 3: Determining the weights of factors

Previous studies have shown that air temperature, wind speed, and visibility are all restrictive factors of outdoor activities, and they play an equally important role in winter tourism [37]. The air relative humidity is a non-restrictive factor of winter tourism. Therefore, the weight of this study is assigned as

$$A = [0.3, 0.3, 0.3, 0.1]$$

Step 4: By using the membership formula, the daily fuzzy membership of each meteorological factor of MSI at different levels can be obtained using Equation (4):

$$B = A^{\circ}R \tag{4}$$

where ◦ is fuzzy operator symbols, A is a row matrix of factor weights, R is a matrix of fuzzy sets, which is calculated using Table 2. The MSI is divided into three levels and assigned value 1, 2 and 3. The values 1, 2 and 3 mean high-suitability, medium-suitability, and low-suitability, respectively. The daily MSI is determined according to the principle of maximum membership degree (Equation (5)). To facilitate the calculation of MSI, this study compiled the MSI using MATLAB 2016b (MathWorks, Natick, MA, USA).

$$MSI = max\ (B) \tag{5}$$

Table 2. The membership functions for factors.

Factors	High	Middle	Low
H(°C)	$f(x,\ 2, -10)$	$\max(f(x,2,-14), f(x,2,-3))$	$\max(h(x,-2,\ 2), g(x,-18,-12))$
W(m/s)	$g(x,3,\ 5)$	$f(x,2,\ 7.5)$	$h(x,7,\ 10)$
H(%)	$f(x,5,55)$	$\max(f(x,10,\ 40), f(x,10,70))$	$\max(h(x,75,\ 85), g(x,25,35))$
V(km)	$h(x,0.5,\ 1)$	$f(x,0.1,\ 0.5)$	$g(x,0.1,0.5)$

2.2.2. Snow Abundance Index (SAI)

The abundance of snow resources in a region is mainly expressed by snow thickness and duration. Therefore, daily snow depth and duration are used to calculate the snow abundance index. Winter tourism depends on the amount of snow, only when the snow reaches a certain depth and lasts for a period of time can tourists be attracted. Experience and researchers have found that only when the snow depth reaches more than 10 cm, can greatly reduce the cost of artificial snowfall in the ski resort and get economic benefits [40,50]. It is generally considered that snow sports need at least 30 cm of snow depth [40]. Witmer [51] suggested that a skiable area with at least 30 cm snow depth and 100 days duration per year would be suitable for the skiing industry. If the snowfall is reduced in one year, scenic areas can make snow artificially to satisfy tourists. The snowfall determines the snowmaking cost and tourism income. Therefore, the winter tourism industry needs to obtain the spatial distribution of snow resources. At present, several studies have obtained large-scale spatial snow resources based on remote sensing images and proved their feasibility [32–34]. Based on the snow data obtained from local remote sensing images, the snow abundance index was designed by coupling daily snow depth and duration in this study (Equation (6)). Several studies have provided a method to classify snow depth [52]. In this study, based on previous researches [50–52] and according to the actual situation of snow depth in Jilin Province, the snow depths were divided into <10, 10–20 and >20 three classes:

$$MSI = \begin{cases} 1 & SD \geq 20\ AND\ SC \geq 60 \\ 2 & SD \geq 10\ AND\ SC \geq 30 \\ 3 & \qquad\qquad\qquad others \end{cases} \tag{6}$$

where SD is the snow depth (cm) and SC is the snow cover days (d). The values 1, 2 and 3 mean high-abundance, medium-abundance, and low-abundance, respectively.

2.2.3. Copula Function

Copula functions describe nonlinear relationships among multivariate data and model sample nonlinearly interrelated multivariate data [53]. They are functions that couple the joint distributions to their marginal distributions. Recently, copula methods have been extensively applied for finance, natural sciences and engineering, etc. To analyze the combination of MSI and SAI in space, the joint distribution of two indices is calculated with Copula functions to obtain the suitability degree for winter tourism regions.

Copulas provide a method for measuring the dependence between variables. Sklar [54] described the function relationship between a Copula C and a cumulative distribution function. Suppose there is a bivariate domain, and $F_X(x) = P[X \leq x]$ and $F_Y(y) = P[Y \leq y]$ are cumulative distribution functions of the random variables X and Y. Their joint distribution is $F_{XY}(x, y) = P[X \leq x, Y \leq y]$. Then, on the basis of mathematical theory, there is a unique function C, $F_{XY}(x, y) = C(F_X(x), F_Y(y))$, and C is called a copula function. Thus, if function C was deduced, the joint distribution $F(x, y)$ of variables can be derived from their marginal distributions, $F_X(x)$ and $F_Y(y)$. Furthermore, let the probabilities $u = F_X(x)$ and $v = F_Y(y)$; we can take $x = F_X^{-1}(u)$ and $y = F_Y^{-1}(v)$. Then, $F_{XY}(x, y) = F(F_X^{-1}(u), and\ F_Y^{-1}(v)) = C(u, v)$ is a copula (*Sklar's theorem*) if a two-dimensional function, $C : [0, 1] \times [0, 1] \rightarrow [0, 1]$, meets three conditions:

(1) $C(0, v) = C(u, 0) = 0$;
(2) $C(1, v) = C(u, 1) = 1$;
(3) For all $0 \leq u_1 \leq u_2 \leq 1$ and $0 \leq v_1 \leq v_2 \leq 1$, $C(u_2, v_2) - C(u_2, v_1) - C)u_1, v_2) + C(u_1, v_1) \geq 0$.

There are two categories of widely used copulas: ellipse and Archimedean [55]. Ellipse copulas, such as the Gauss- and *t*-copulas, can be produced by known multivariate distributions. Archimedean copulas are produced from different generators (φ) based on the definition of the Archimedean copula [56]. At present, Archimedean copulas are widely used in actual applications because they can model dependence in arbitrarily high dimensions with only one parameter, governing the strength of dependence. The one-parametric Archimedean Copula is expressed as:

$$C(u, v; \theta) = \varphi^{-1}[\varphi(u, \theta) + \varphi(v, \theta); \theta] \tag{7}$$

where C is a Copula function, φ is the generator function, and φ^{-1} is its pseudoinverse. $\varphi : [0, 1] \times \Theta \rightarrow [0, \infty)$, $\varphi(1) = 0$. θ is a generator function parameter within parameter space Θ. Further important Copula features and the theoretical background can be found in Nelsen [57], who provides a detailed introduction to Copula functions.

Kendall's τ is the rank correlation coefficient. It can be calculated from the available observation samples as:

$$\tau = \begin{pmatrix} N \\ 2 \end{pmatrix}^{-1} \sum_{j=1}^{N} \sum_{i=1}^{j} sign[(x_i - x_j)(y_i - y_j)] \tag{8}$$

where sign = 0 if $[(x_i - x_j)(y_i - y_j)] = 0$, sign = 1 if $[(x_i - x_j)(y_i - y_j)] > 0$, and sign = −1 if $[(x_i - x_j)(y_i - y_j)] < 0$ and $i, j = 1, 2, \cdots, N$ [58]. The corresponding Copula function coefficient θ can be calculated using the functions in the right column of Table 3.

Table 3. Archimedean Copula functions, their generators, and connections to Kendall's τ considered for this study.

Kinds	Copula Function (C)	Generator ($\varphi(t)$)	Range $\theta \in$	Kendall's τ
Frank	$\frac{-1}{\theta} ln[1 + \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}]$	$\ln(\frac{e^{-\theta t}-1}{e^{-\theta}-1})$	$(-\infty, +\infty) \backslash \{0\}$	$r = 1 - \frac{4}{\theta}[1 - \frac{1}{\theta} \int_0^\theta \frac{t}{e^t-1} dt]$
Clayton	$[u^{-\theta} + v^{-\theta} - 1]^{\frac{-1}{\theta}}$	$t^{-\theta} - 1$	$[0, \infty)$	$\tau = \frac{\theta}{\theta+2}$
Gumbel	$exp\{-[(-lnu)^\theta + (-lnv)^\theta]^{\frac{1}{\theta}}\}$	$(-lnt)^\theta$	$[1, \infty)$	$\tau = 1 - \theta^{-1}$

In this study, MSI and SAI are used to describe the suitability degree of winter tourism. The joint probability of the two indices in the space is analyzed based on Copula functions. The root mean square errors (RMSEs) are used to identify the most appropriate Copula function, which is calculated from the theoretical and the empirical joint non-exceedance probabilities [59]. The RMSEs are determined using all of the observed samples and empirical non-exceedance probabilities of events:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[P_c(i) - P_o(i)]^2} \tag{9}$$

where $P_c(i)$ is the joint probability of the i^{th} joint observation value calculated by the Copula function, N is the observation samples, and $P_o(i)$ is the empirical joint nonexceedance probability, calculated as follows [58,60].

$$F_{XY}(x_i, y_i) = P(X \leq x_i, Y \leq y_i) = \frac{\sum_{m=1}^{i}\sum_{l=1}^{i} N_{mj} - 0.44}{N + 0.12} \tag{10}$$

where N_{mj} represents the number of occurrences of (x_m, y_j) with $x_m < x_i$ and $y_j < y_i$, $i = 1, \ldots, N$ and $m, j \in [1, i]$. N is the sample size. In the calculation process, the pairs (x_m, y_j) are arranged in ascending order with respect to x_m.

3. Results

3.1. Changing Temperature and Snowfall under Climate Change

In this study, the winter snowfall, air temperature and wind data are used to assess the suitability of winter tourism in Jilin Province. The snowfall data were mainly used to analyze the influence of snow depth on winter sports, and the air temperature and wind determine the comfort level of outdoor sports. Several research studies show that snowfall and air temperature have changed significantly in Northeast China with climate change [31,61]. As part of Northeast China, the spatial distribution of snow, air temperature, and wind in Jilin province has also changed significantly. Revealing these changes can provide important informational support for winter tourism management in Northeast China. Therefore, in this study, the spatial and temporal distribution of snow, air temperature, and wind in Jilin province were analyzed using meteorological data from 1971 to 2016 (from December to February of the next year).

Figure 2 shows the snowfall depth in Jilin Province varying from 5 to 30 cm, and the snowfall in winter increased by 1.3 mm/10a. Figure 3 shows that the temperature in Jilin Province varied from −16.4 °C to −9.2 °C, and the average temperature increased 0.27 °C/10a in the study area, including 0.25 °C/10a in the west, 0.26 °C/10a in the middle and 0.29 °C/10a in the east. The overall air temperature and snowfall change is beneficial to winter tourism. However, the annual temperature and annual snowfall fluctuated strongly. This will not be a disadvantage to winter tourism.

Climate change has favorable and unfavorable conditions for winter tourism. To avoid the influence of unfavorable factors, we need to distinguish the regions that are more suitable for development of winter tourism space. Therefore, the two indices, namely, MSI and SAI, were used to analyze the suitability of winter tourism space.

MSI evaluated comprehensively the suitability of meteorological environment using fuzzy inference. It is the weather index for suitable winter tourism. By coupling analysis of daily maximum air temperature, mean wind speed, air relative humidity, and visibility in winter, the spatial distribution of annual MSI was obtained in the study area from 1971 to 2016. The evaluation results (class (I)) were verified using the method provided by the literature [38] (class (II)) (Figure 4).

Figure 4 shows that the percentage difference between the two evaluation methods was 27.8%, 37.5% and 37.3% in the three grades of high (value1), medium (value1) and low (value3), respectively. Overall, 65.8% of the evaluation results are coincident. The major reasons are that visibility and relative humidity are considered in this study. Therefore, the degree of suitability predicted in this study is lower than that provided with the literature [38]. But in terms of temperature and wind speed, the evaluation results of the two methods are consistent.

By analyzing the mean value of annual MSI in the space, it is found that the maximum frequencies of high-suitability and medium-suitability degrees in the study area are 27.1% and 26.1%, respectively. The high frequency of high-suitability areas is mainly distributed in the eastern region and gradually decreases to the west, whereas the medium-suitability area is mainly distributed in the central region and gradually decreases to the east and west. It can avoid the inherent subjectivity of objective selection based on experience and make evaluation more scientific and reasonable. Compared with the analytic hierarchy process and risk matrix method, MSI calculated by fuzzy inference is more objective, and the evaluation results are closer to the actual situation.

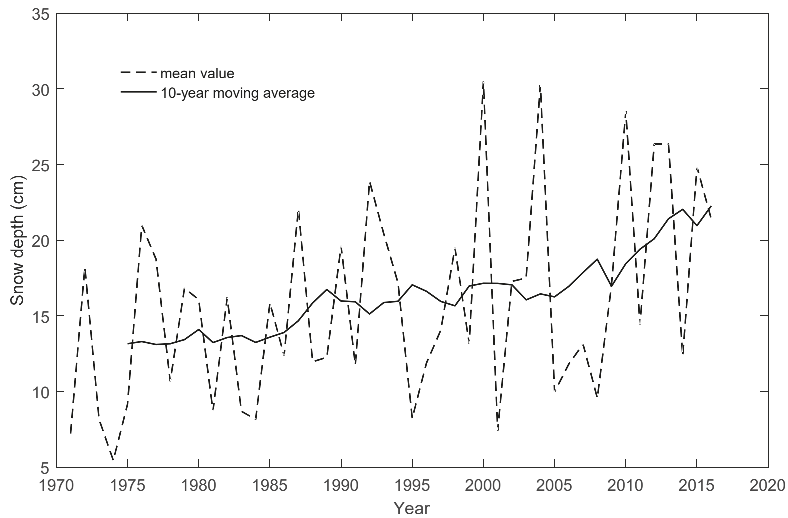


Figure 2. Snow depth change from 1971 to 2016 in Jilin Province.

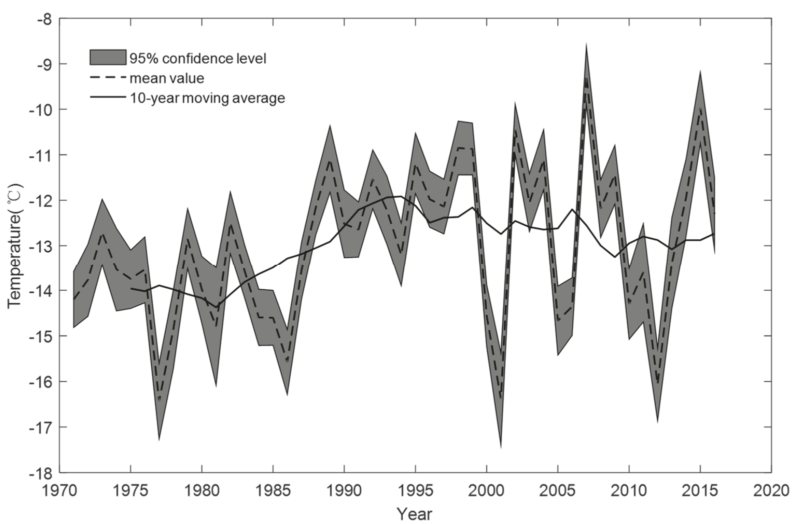


Figure 3. The temperature change from 1971 to 2016 in Jilin Province.

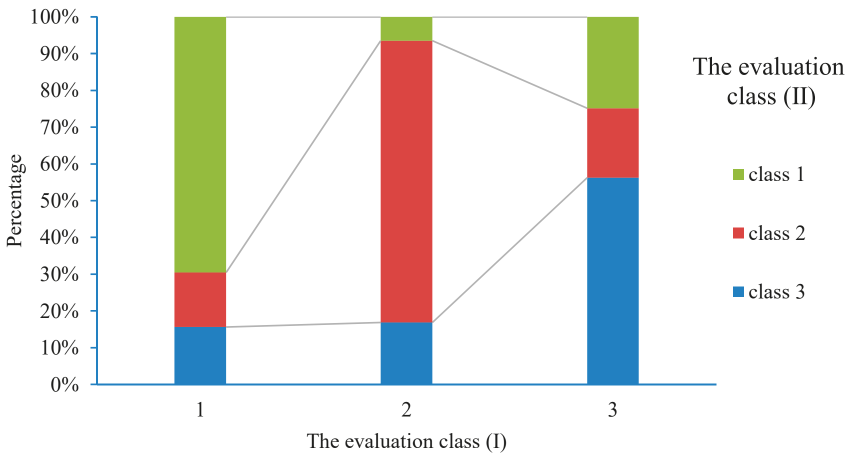


Figure 4. The cross validation of evaluation results in class (I) and class (II).

The SAI indicates the snow richness for outdoor activities in winter. It found an obvious spatial distribution difference between snow depth and duration in eastern, central and western Jilin province. The snow cover in the eastern region lasted for a long time, followed by the central plain area. The western region has the shortest snow cover time. The snow depth and snow cover decrease from southeast to northwest. The snow resource is mainly concentrated in the central-eastern study area. The maximum durations of snow depth exceeding 10 cm and 20 cm were reached at 35 and 79 days, respectively. In the central part of Jilin Province, the 10 cm snow depth lasted for 3–10 days, and the 20 cm snow depth lasted for 10–30 days. In the western part of Jilin, the duration of 10 cm snow is 1–5 days, and 20 cm snow depth is 5–10 days. By analysis of the snow depth contour exceeding 10 and 20 cm (Figure 5), it was found that the areas of 10 cm snow depth increased in the 1980s, 1990s, and 2010s, but the areas of 20 cm snow depth decreased in the same periods. This means that the number of heavy snow days is decreasing and that the number of light snow days increased in these periods.

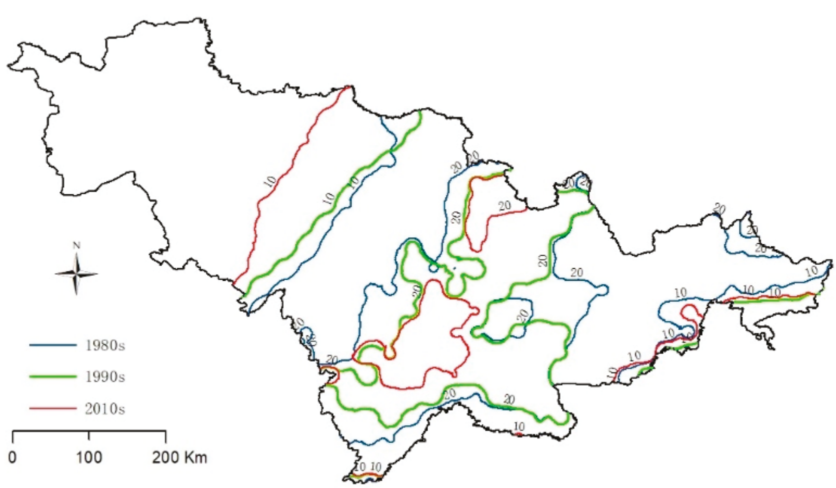


Figure 5. The contour of 10 and 20 cm snow depths in the 1980s, 1990s and 2010s.

3.2. The SAI and SMI Joint Probability

To obtain stable income, a suitable area for winter tourism development should be located where the meteorological factors steadily change annually and the winter tourism resources are suitable for development. Therefore, the coupling of annual MSI and SAI from 1971 to 2014 was used to analyze the suitability of winter tourism. To analyze the spatial coupling relationship between the annual SAI and SMI, the joint probability of the two indices was calculated based on Copula functions, and then the winter tourism suitability degree was established. In this study, the annual SAI and SMI were used to establish the copula function. The RMSEs determined by all of the observed samples identified the most appropriate copula function. By comparing the copula functions in Table 1, the results show that the RMSEs are 0.298 (Clayton), 0.241 (Frank), and 26.264 (Gumbel) at the 0.05 significance level (Figure 6). Therefore, the Frank copula function was selected to calculate the joint probability, and $\theta = -1.471$ was calculated using the study area sample data. Through analysis, it is found that the high-suitability region is located in central and eastern Jilin Province (the joint probability >60%), the low-suitability region is located in west Jilin province (the joint probability <40%), and Yushu–Jiutai–Yitong–Dongliao is the boundary line. In central and eastern Jilin Province, the spatial distribution matching degree of the annual MSI and SAI is high and winter tourism resources are abundant (Figure 6). According to the two indices, the western region is weak for developing winter tourism because of low air temperature and the lack of snow resources.

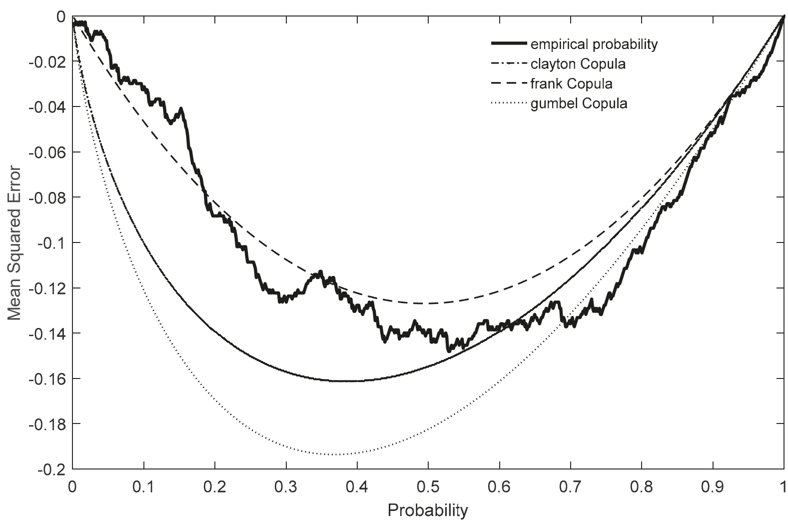


Figure 6. The verification results of copula functions.

Snow-based winter tourism has long been dealing with variability in natural snowfall and seasonal temperatures, which has led to early adaptive interventions and investments [62]. Figure 7 shows that winter tourism suitability in western Jilin is low, mainly due to lower air temperature in the winter (annual air temperature below −13 °C, and annual rainfall below 10 cm), so it is not suitable for development of ice and snow tourism. Due to current limitations on temperatures for snowmaking (usually below −2 °C) [63], western Jilin Province is suitable for artificial snowmaking. However, due to the lack of regional competitiveness, the attractiveness of its scenic spots is not as attractive as that of the central and eastern region. The most suitable areas for development of ice and snow tourism are the central and eastern regions of Jilin Province, mainly because they are rich in snow resources (snow depth is over 15 cm) and annual air temperature (over −10.4 °C) suitable for outdoor activities and long snow retention time (over 150 days), which makes for less restrictions on the development of tourism and for higher winter tourism income. This is the most advantageous area for the development of the ice and snow industry. There are 38 large ski resorts in Jilin province, and all are located in this area. Jilin province has designated 27 key winter tourist attractions, 94% of which are located in the region. These analysis results are consistent with the actual Jilin Province situation.

Copula functions characterize the spatial interdependence of two factors. It could analyze the joint recurrence probability of the two indices in space in this study. The higher the joint recurrence probability is, the better the spatial suitability of winter tourism is, which means that this area was less affected by climate change, and it is suitable for the development of winter tourism. Compared with the traditional comprehensive evaluation method (e.g., weighted overlay), the parameters of this method are obtained by calculating historical data, which can avoid the shortcomings of subjectivity in the weighted overlay.

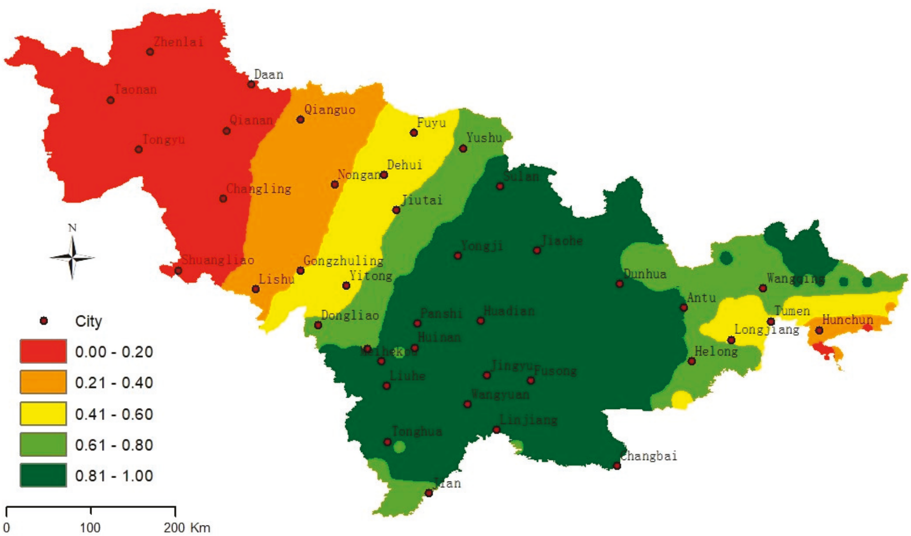


Figure 7. The joint probability spatial distribution of the SAI and MSI in the study area. The suitability degrees were classified as 0.00–0.20 (very low), 0.21–0.40 (low), 0.41–0.60 (medium), 0.61–0.80 (medium-high), and 0.81–1.00 (high).

4. Discussion

To get the spatial distribution of snow resources for winter tourism assessment, remote sensing data were used to calculate snow depth. Snowfall data from meteorological stations were used to verify the accuracy of snow depth obtained from remote sensing data in space. The spatial location of winter tourism scenic spots and skiing areas in Jilin province verifies the assessment results.

4.1. Spatial Snow DEPTH verification

According to the data from meteorological stations, snowfall is increasing gradually from west to east in Jilin province from the boxplot (Figure 8).

The trend obtained from two data sources is identical in the study area. From the variety of snowfall, there is little difference in depth in western Jilin Province, but the snow depth varies greatly in eastern Jilin Province. The reasons are that the western Jilin Province belongs to the arid area, the annual snowfall in this area are 29.3–40 mm in the winter, while this area belongs to a plain area with little snow cover. For the eastern region, this area belongs to a humid area with diverse land use and complex topography, which forms several local microclimates. This area has the abundant precipitation (100–190 mm) in the winter. With the monsoon and the topography, this caused a great change in the winter snow depth.

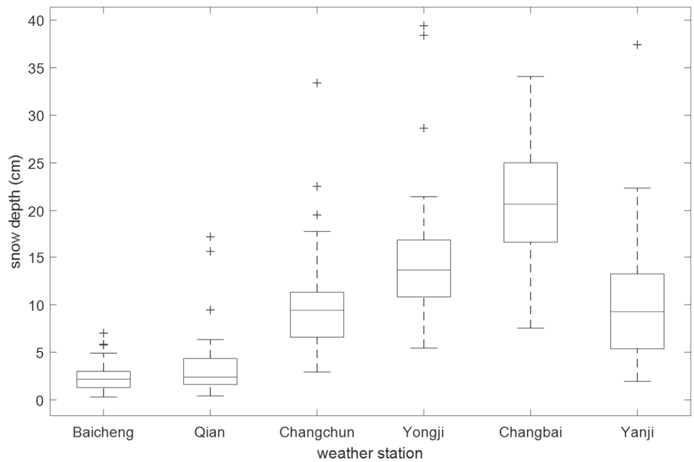


Figure 8. The snowfall from weather stations in the study area.

A total of 186 snowfall records from 1980 to 2014 in six meteorological stations were selected to verify the snow depth results. Based on the ascending observational data of meteorological stations, the tendencies of snow depth in the two sources are consistent in spatial distribution (Figure 9), but there is a certain difference. This is because meteorological observation data mainly record one snowfall, and the snow depth from remote sensing data may be multiple-snowfalls in the same space. Therefore, there will be a certain gap between the two data sources.

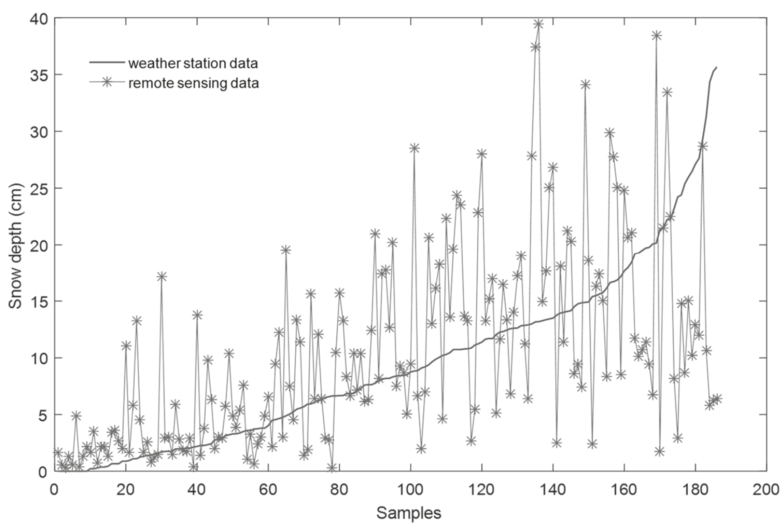


Figure 9. Comparing the two sources of snow resources data.

4.2. Winter Tourism Verification

To verify the reliability of winter tourism resources, this study obtained the spatial location of winter tourism scenic spots and skiing areas in Jilin Province (Figure 10). Their spatial distribution shows that these spots are mainly concentrated on the snow abundant areas. In addition to Changchun, the winter tourist attractions in the eastern region are obviously more than in the west. This shows a decreasing trend from east to west consistent with the spatial distribution suitability degree obtained in this study. Changchun is provincial capital, which located in the central area of Jilin Province. The snow resources in this area are at a medium level. However, due to its developed economy and good tourism infrastructure, it has great attraction for winter tourists. The development of scenic spots on other areas is mainly based on the exploitation of ice and snow resources.

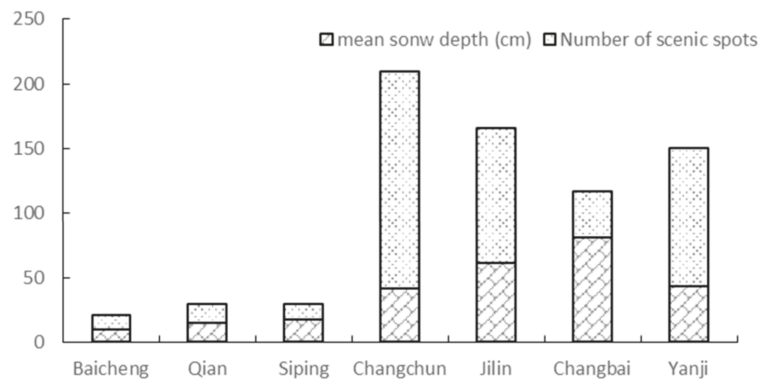


Figure 10. The number of winter tourism scenic spots and skiing areas in Jilin province.

By comparing these locations with the classification results of this study, Cohen's kappa coefficient was used to measure the results (Equation (11)):

$$k = \frac{p_0 - p_c}{1 - p_c} \tag{11}$$

where k is the kappa coefficient, p_0 is the observed agreement among raters, and p_c is the hypothetical probability of chance agreement. By calculating the consistency and coefficient of the actual suitability of winter tourism scenic spots and skiing areas using Equation (7), it is found that k = 0.771 > 0.70, which means that the assessment results are highly consistent with the actual situation.

5. Conclusions

To analyze the spatial suitability distribution of winter tourism, the daily maximum temperature, mean wind speed, relative humidity, and visibility are used to establish the MSI, and snow depth and duration are used to establish the SAI. The coupling relationship of the two indices is used to analyze the suitability degree of winter tourism based on a copula function, and then the distribution of winter tourism resources can be obtained and used in tourism planning.

According to the MSI analysis, the high-suitability area mainly distributed in the eastern region, whereas the medium-suitability area mainly distributed in the central region. The SAI analysis shows that the snow resource mainly concentrates on central-eastern Jilin province, and the snow resources in the western region are less. Through the analysis of snow and weather data from 1980 to 2014, it is found that the *Frank* is the suitable function to analyze the coupling relationship between MSI and SAI with an error of 0.241 at the 0.05 significance level. From the MSI and SAI coupling, it is found that central-eastern Jilin province is suitable for developing tourism with a tendency to decrease from the southeast to the northwest. Yushu–Jiutai–Yitong–Dongliao is the main boundary line of winter tourism resources. From the annual variations of MSI and SAI, there is high suitability in weather conditions and snow resources in the eastern region, which is suitable for the development of winter tourism. On the contrary, it has a few snow resources and large inter-annual fluctuation in the western region, therefore, the construction of long-term ski resort in this area will be very expensive. The low temperature in the west makes it possible to develop the ice-related tourism industry. That is, the eastern area is suitable for developing winter tourism, and the suitability degree of the western region is weak from little snowfall and strong wind.

Previous studies have shown that climate change will bring positive and negative impacts on winter tourism in different regions. There are also large differences in seasons. Climate change causes a lot of losses to the tourism industry in some regions while improving the suitability of tourism in other regions [64]. Combined with the change of 10 cm snow depth contour map (Figure 5) and the average temperature in winter in Jilin Province (Figure 3), it can conclude that the mean temperature in winter in Jilin Province showed an increasing trend (0.29 °C/10a). In terms of snow depth, winter precipitation in Jilin Province showed an increasing trend, and the 10 cm snow depth contour has been moving westward from the 1980s to 2010s. Therefore, combined with these two trends, it can be predicted that the winter tourism suitability in the central -western regions of Jilin Province will increase under the influence of climate change. These conclusions are consistent with the previous research results [65]. Therefore, the suitability boundary of winter tourism will move westward.

In this study, MSI can be used to guide tourists to travel outdoor activities in winter. SAI can be used for regional snow resources assessment and water resources assessment. This study proposed a spatial analysis method for winter tourism suitability that can be used for winter tourism planning. The study results revealed the spatial constraints of meteorological and snow resources on winter tourism. This study will have a potential impact on the attractiveness analysis of winter tourist destinations and the development of winter tourism industry. This study could provide decision support for the tourist choosing winter tourism destinations, the construction of winter ski resorts and the spatial layout of the winter tourism industry. In a further study, the winter tourism spatial suitability and its impact on the industrial economy will be studied by taking into account land use, transport, terrain, and the socioeconomy.

Author Contributions: W.C. analyzed the data and wrote the paper; X.L. provided the research methods in this study; H.D. collected the data.

Funding: This study is supported by the Innovation and Development Fund Project of the Tourism College of Changchun University (2017) Northeast Asia special project (XJZX-03).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Becken, S. *The Importance of Climate and Weather for Tourism*; Technical Report Lincoln University; Lincoln University: Christchurch, New Zealand, 2013.
2. Falk, M. A dynamic panel data analysis of snow depth and winter tourism. *Tourism Manag.* **2010**, *31*, 912–924. [CrossRef]
3. Schwirplies, C.; Ziegler, A. Adaptation of future travel habits to climate change-A micro econometric analysis of tourists from Germany. *Tourism Econ.* **2017**, *23*, 1275–1295. [CrossRef]
4. Falk, M.; Lin, X. The declining dependence of ski lift operators on natural snow conditions. *Tourism Econ.* **2018**, *24*, 662–676. [CrossRef]
5. Wu, B.; Wei, Q.H. White Book on Skiing Industry in China. 2017. Available online: https://www.wenzikong.cc (accessed on 15 August 2018).
6. Scott, D.; Lemieux, C. Weather and climate information for tourism. *Procedia Environ. Sci.* **2010**, *1*, 146–183. [CrossRef]
7. Cools, M.; Moons, E.; Wets, G. Assessing the impact of weather on traffic intensity. *Weather Clim. Soc.* **2010**, *2*, 60–68. [CrossRef]
8. Conradt, S. Flexible weather index-based insurance design. *Clim. Risk Manag.* **2015**, *10*, 106–117. [CrossRef]

9. Zhou, W.; Shi, J.; Haizhen, M.U. Climate change characteristics of winter temperature and precipitation in eastern china. *Res. Sci.* **2010**, *32*, 1088–1096.
10. Damm, A.; Köberl, J.; Prettenthaler, F.; Töglhofer, C. Economic Impacts of Climate Change on Winter Tourism: Challenges for Ski Area Operators. *EGU Gen. Assem. Conf.* **2012**, *14*, 9287.
11. Scott, D.; McBoyle, G.; Mills, B. Climate change and the skiing industry in southern Ontario (Canada): Exploring the importance of snowmaking as a technical adaptation. *Clim. Res.* **2003**, *23*, 171. [CrossRef]
12. Laternser, M.; Schneebeli, M. Long-term snow climate trends of the Swiss Alps (1931–99). *Int. J. Climatol.* **2003**, *23*, 733–750. [CrossRef]
13. Gajićčapka, M.; Srnec, L. Response of snow parameters relevant for winter tourism to climate change. In Proceedings of the International Congress of Biometeorology Harmony Within Nature Proceedings, Tokyo, Japan, 22–26 September 2008.
14. Hoffmann, V.H.; Sprengel, D.C.; Ziegler, A.; Kolb, M.; Abegg, B. Determinants of corporate adaptation to climate change in winter tourism: An econometric analysis. *Glob. Environ. Chang.* **2009**, *19*, 256–264. [CrossRef]
15. Dingeldey, A.; Soboll, A. The future impact of climate change on alpine winter tourism: A high-resolution simulation system in the German and Austrian alps. *J. Sustain. Tourism* **2012**, *20*, 101–120.
16. Gonseth, C. Impact of snow variability on the Swiss winter tourism sector: Implications in an era of climate change. *Clim. Chang.* **2013**, *119*, 307–320. [CrossRef]
17. Bonzanigo, L.; Giupponi, C.; Balbi, S. Sustainable tourism planning and climate change adaptation in the Alps: A case study of winter tourism in mountain communities in the Dolomites. *J. Sustain. Tourism* **2016**, *24*, 637–652. [CrossRef]
18. Steiger, R. The impact of climate change on ski season length and snowmaking requirements in Tyrol. *Aust. Clim. Res.* **2010**, *43*, 251. [CrossRef]
19. Pütz, M.; Gallati, D.; Kytzia, S.; Elsasser, H.; Lardelli, C.; Teich, M. Winter tourism, climate change, and snowmaking in the Swiss Alps: tourists' attitudes and regional economic impacts. *Mountain Res. Dev.* **2011**, *31*, 357–362. [CrossRef]
20. Rixen, C.; Teich, M.; Lardelli, C.; Gallati, D.; Pohl, M.; Pütz, M. Winter tourism and climate change in the Alps: An assessment of resource consumption, snow reliability, and future snowmaking potential. *Mountain Res. Dev.* **2011**, *31*, 229–236. [CrossRef]
21. Koenig, U.; Abegg, B. Impacts of climate change on winter tourism in the Swiss Alps. *J. Sustain. Tourism* **1997**, *5*, 46–58. [CrossRef]
22. Scott, D.; McBoyle, G. Climate change adaptation in the ski industry. *Mitig. Adapt. Strat. Glob. Chang.* **2007**, *12*, 1411. [CrossRef]
23. Scott, D.; Dawson, J.; Jones, B. Climate change vulnerability of the US northeast winter recreation- tourism sector. *Mitig. Adapt. Strat. Glob. Chang.* **2008**, *13*, 577–596. [CrossRef]
24. Winter, C. Tourism and climate change: Risks and opportunities. *Ann. Tourism Res.* **2008**, *35*, 614–616. [CrossRef]
25. Ogrin, M.; Ogrin, D.; Rodman, N.; Močnik, M.; Vengar, R.; Smolej, A. Climate change and the future of winter tourism in slovenia. *Hrvatski Geografski Glas.* **2011**, *73*, 215–228. [CrossRef]
26. Wolfsegger, C.; Gössling, S.; Scott, D.; Gössling, S.; Scott, D. Climate change risk appraisal in the Austrian ski industry. *Tourism Rev. Int.* **2008**, *12*, 13–23. [CrossRef]
27. Ponspons, M.; Johnson, P.A.; Rosascasals, M.; Sureda, B.; Jover, È. Modeling climate change effects on winter ski tourism in Andorra. *Clim. Res.* **2012**, *54*, 197–207. [CrossRef]
28. Pröbstl, U.; Prutsch, A.; Formayer, H.; Landauer, M.; Grabler, K.; Kulnig, A. Climate change in winter sport destinations: Transdisciplinary research for implementing sustainable tourism. *Sustain. Tourism* **2013**, *3*, 165–173.
29. Tranos, E.; Davoudi, S. The regional impact of climate change on winter tourism in Europe. *Tourism Plan. Dev.* **2014**, *11*, 163–178. [CrossRef]
30. Falk, M.; Lin, X. Sensitivity of winter tourism to temperature increases over the last decades. *Econ. Model.* **2018**, *71*, 174–183. [CrossRef]
31. Zhao, C.Y.; Wang, Y.; Zhou, X.Y.; Cui, Y.; Liu, Y.L.; Shi, D.M.; Yu, H.M.; Liu, Y.Y. Changes in climatic factors and extreme climate events in Northeast China during 1961–2010. *Adv. Clim. Chang. Res.* **2013**, *4*, 92–102.

32. Chang, A.T.C.; Foster, J.L.; Hall, D.K. Nimbus-7 SMMR Derived Global Snow Cover Parameters. *Ann. Glaciol.* **2016**, *9*, 39–44. [CrossRef]
33. Che, T.; Li, X.; Jin, R.; Armstrong, R.; Zhang, T.J. Snow depth derived from passive microwave remote-sensing data in China. *Ann. Glaciol.* **2008**, *49*, 145–154. [CrossRef]
34. Zhen, L.I.; Shao, Y.; Zhou, J.; Tian, B.; Chen, Q. Snow-depth retrieval algorithms on MEMLS. *J. Remote. Sens.* **2013**, *17*, 1993–2002.
35. Terjung, W.T. Physiologic climates of the conterminous United States: A bioclimatic classification based on man. *Ann. Assoc. Am. Geogr.* **1966**, *5*, 141–179. [CrossRef]
36. Cai, X. Golden section, the optimization of environmental factors for human body. *Chin. J. Clin. Rehabil.* **2006**, *10*, 227–228.
37. China Meteorological Administration, QXT 386-2017. Available online: http://www.cma.gov.cn/root7/auto13139/201801/t20180117_460483.html (accessed on 10 October 2018).
38. Elsasser, H.; Messerli, P. The vulnerability of the snow industry in the Swiss Alps. *Mountain. Res. Dev.* **2001**, *21*, 335–339. [CrossRef]
39. Elsasser, H.; Bürki, R. Climate change as a threat to tourism in the Alps. *Clim. Res.* **2002**, *20*, 253–257. [CrossRef]
40. Yang, J.; Wan, C. Progress in research on the impacts of global climate change on winter ski tourism. *Adv. Clim. Change Res.* **2010**, *1*, 55–62. [CrossRef]
41. Brouder, P.; Lundmark, L. Climate change in Northern Sweden: Intra-regional perceptions of vulnerability among winter-oriented tourism businesses. *J. Sustain. Tourism* **2011**, *19*, 919–933. [CrossRef]
42. Steiger, R. Scenarios for skiing tourism in Austria: Integrating demographics with an analysis of climate change. *J. Sustain. Tourism* **2012**, *20*, 867–882. [CrossRef]
43. Wang, L. A summery on the ice-snow tourism development research at home and abroad. *Ecol. Econ.* **2010**, *11*, 66–69. (In Chinese)
44. Liu, W.J.; Jiang, M.M. Value and value system construction of the ice snow tourism resources. *China Winter Sports* **2016**, *38*, 90–93. (In Chinese)
45. Cheng, Z.H.; Liu, K.; Sun, J. The comprehensive evaluation of suitability of ice-snow tourism base in China. *Resour. Sci.* **2016**, *38*, 2233–2243. (In Chinese)
46. Che, T.; Dai, L. Long-term snow depth dataset of China. *Cold Arid Reg. Sci. Data Center Lanzhou* **2014**. [CrossRef]
47. Gao, F.; Xie, Y.; Piao, J.H.; Li, T. The relationship of meteorological condition and ski sport. *Jilin Weather* **2017**, *1*, 21–25. (In Chinese)
48. Ming, J.; Chen, X.I.; Meng, M.; Xu, J.-Q. The influence of air temperature variety on skiing in heilongjiang province. *J. Harbin Inst. Phys. Educ.* **2008**, *26*, 107–110. (In Chinese)
49. Richins, H.; Hull, J.S.; Richins, H. *Mountain Tourism: Experiences, Communities, Environments and Sustainable Futures*; CABI Publishing: Oxfordshire, UK, 2016.
50. National Tourism Administration. *Quality Classification of Tourist Skiing Grounds (LB/T037-2014)*; National Tourism Administration: Beijing, China, 2014.
51. Witmer, U. *Erfassung, Bearbeitung und Kartierung von Schneedaten in der Schweiz*; Geographica Bernensia G25, Geographical Instituite, University of Bern: Bern, Switzerland, 1986.
52. Li, X.F.; Zheng, X.M.; Wu, L.L.; Zhao, K.; Jiang, T.; Gu, L.J. Effects of snow cover on ground thermal regime: A case study in Heilongjiang Province of China. *Chin. Geogr. Sci.* **2016**, *26*, 527–538. [CrossRef]
53. Goda, K.; Ren, J.D. Assessment of seismic loss dependence using copula. *Risk. Anal.* **2010**, *30*, 1076–1091. [CrossRef]
54. Sklar, A. Fonctions de Répartition à n Dimensions et Leurs Marges. *Inst. Stat. Univ. Paris* **1959**, *8*, 229–231.
55. Wang, L.F.; Zeng, J.C.; Hong, Y. Estimation of distribution algorithm based on copula theory. In *Exploitation of Linkage Learning in Evolutionary Algorithms. Evolutionary Learning and Optimization*; Springer: Berlin, Germany, 2009; pp. 1057–1063.
56. Nelsen, R.B. *An Introduction to Copulas*, 2nd ed.; Springer Science Business Media Inc.: New York, NY, USA, 2006; ISBN 978-1-4419-2109-3.
57. Nelsen, R.B. *An Introduction to Copulas*; Lecture Notes in Statistics: New York, NY, USA, 1999.
58. Karmakar, S.; Simonovic, S.P. Bivariate flood frequency analysis: Part 2—A copula-based approach with mixed marginal distributions. *J. Flood Risk Manag.* **2009**, *2*, 32–44. [CrossRef]

59. Zhang, L.; Singh, V.P. Trivariate flood frequency analysis using the gumbel–hougaard copula. *J. Hydrol. Eng.* **2007**, *12*, 431–439. [CrossRef]
60. Yue, S.; Ouarda, T.B.; Bobée, B.; Legendre, P.; Bruneau, P. The gumbel mixed model for flood frequency analysis. *J. Hydrol.* **1999**, *226*, 88–100. [CrossRef]
61. Zhao, D.; Zheng, D.; Wu, S. Climate changes in northeastern China during last four decades. *Chin. Geogr. Sci.* **2007**, *17*, 317. [CrossRef]
62. Rutty, M.; Scott, D.; Johnson, P.; Pons, M.; Steiger, R.; Vilella, M. Using ski industry response to climatic variability to assess climate change risk: An analogue study in eastern Canada. *Tourism Manag.* **2017**, *58*, 196–204. [CrossRef]
63. Wobus, C.; Small, E.E.; Hosterman, H.; Mills, D.; Stein, J.; Rissing, M.; Jones, R.; Duckworth, M.; Hall, R.; Kolian, M.; Creason, J.; Martinich, J. Projected climate change impacts on skiing and snowmobiling: A case study of the United States. Glob. *Environ. Chang.* **2017**, *45*, 1–14. [CrossRef]
64. Mayer, M.; Demiroglu, O.C.; Ozcelebi, O. Microclimatic volatility and elasticity of glacier skiing demand. *Sustainability* **2018**, *10*, 3536. [CrossRef]
65. Gao, F.; Wang, T.Y.; Li, D.H.; Wang, Y. Analysis of Winter Climate Change in Jilin Province in Recent 53 Years. *Jilin Weather* **2015**, *3*, 31–34. (In Chinese)

International Journal of
Environmental Research and Public Health

MDPI

Article

The Effectiveness of Laser-Assisted Surgical Excision of Leukoplakias and Hyperkeratosis of Oral Mucosa: A Case Series in A Group of Patients

Dorina Lauritano [1,*,†], Alberta Lucchese [2,†], Federica Gabrione [1], Dario Di Stasio [2], Javier Silvestre Rangil [3] and Francesco Carinci [4]

1 Department of Medicine and Surgery, Centre of Neuroscience of Milan, University of Milano-Bicocca, 20126 Milan, Italy; f.gabrione@campus.unimib.it
2 Multidisciplinary Department of Medical and Dental Specialties, University of Campania- Luigi Vanvitelli, 80138 Naples, Italy; alberta.lucchese@unicampania.it (A.L.); dario.distasio@unicampania.it (D.D.S.)
3 Department of Stomatology, University of Valencia, University Hospital Dr Peset, 46017 Valencia, Spain; javier.silvestre@uv.es
4 Department of Morphology, Surgery and Experimental Medicine, University of Ferrara, 44121 Ferrara, Italy; crc@unife.it
* Correspondence: dorina.lauritano@unimib.it; Tel.: +39-335-679-0163
† Equal contribution for the first authorship.

Received: 24 November 2018; Accepted: 9 January 2019; Published: 13 January 2019

Abstract: *Introduction*: In the different branches of dentistry, the use of laser to solve different clinical situations is increasing due to numerous advantages that have been studied in literature since the 70s. Leucoplakia and hyperkeratosis can benefit from laser-assisted treatment. In most cases biopsy sampling, histological examination and, if no malignant cells are present, the follow-up is needed. However, even if the lesion is free of dysplasia patients often ask to eliminate these white spots that are always a cause of concern. *Aim*: From these numerous requests comes the idea of setting up a laser-assisted protocol as less invasive as possible to be offered to patients. The aim of the study is to find a laser-assisted protocol for the surgical excision of leucoplakia and hyperkeratosis that can both improve the clinical aspect of the lesion and be sustainable for patients. The null hypothesis has been identified in the following statement: the treatment is effective and efficient at the same time; where effectiveness was tested with the following criteria: size of the lesion, tactile perception, discomfort, pain; and efficiency with the following criteria: pain and discomfort perceived during the treatment. *Materials and methods*: To collect all data, a specially designed medical record was used. The diode laser was used with a pulsed mode and the maximum power corresponds to 1.8 W. No anaesthesia was used. Before laser-assisted treatment, the fibre was activated and was used with a contact overflowing. *Results*: Our results show a decrease in the size of the lesion statistically significant. No pain was referred during treatment, except for a slight burning sensation. *Conclusion*: In conclusion we can state that the treatment is both efficient and effective.

Keywords: oral surgery; oral medicine; leucoplakia; hyperkeratosis; laser

1. Introduction

According to the World Health Organization, leucoplakia is defined as a "white patch that cannot be associated either clinically or histopathologically to other diseases" so the diagnosis is by exclusion.

Leucoplakia is the most common oral precancerous alteration and appears as a white, chronic lesion, that cannot be removed with rubbing. Leucoplakia is characterized by an abnormal keratinization of the mucosa. The prevalence in the general population ranges from 0.6% to 5% and is more frequent in those between 40 and 50 years old [1]. The cause is multifactorial and there are several predisposing factors such as: mechanical irritation, dental materials causing galvanic currents, contact with carcinogens such as tobacco and alcohol and gastroesophageal reflux [2,3]. The predisposition to malignant transformation seems associated to a higher CD8+ cells levels in premalignant lesion [4].

The presumed diagnosis is based on the anamnesis and the clinical aspect, while the diagnosis can only be confirmed after both an incisional and excisional biopsy and its subsequent histological examination [5].

Some epidemiological studies have identified the occasional possibility of spontaneous regression of leucoplakias. However, in the vast majority of cases various surgical and non-surgical treatments are proposed such as elimination of the chronic irritative factors, surgical treatment with a scalpel, electrocautery, laser or cryosurgery; conservative treatment is to be considered only if the patient denies consent to surgery or if the areas have a low probability of malignant transformation [6–10].

Hyperkeratoses are benign lesions that usually appear in areas subjected to frictional trauma, for example in adherent gingiva under the prosthetic flanges, the retromolar trigon and the edge of the tongue, where the patient often tends to bite. During intraoral examination, hyperkeratosis tends to have a verruciform or corrugated appearance. Histological examination shows hyperkeratosis, acanthosis, hypergranulosis and inflammation of the stroma.

Although it is benign, some precancerous lesions such as leucoplakia can mimic the characteristics of the hyperkeratosis. The clinical appearance of the lesion can help clinicians in the differential diagnosis between leucoplakia and hyperkeratosis; in fact hyperkeratosis shows less definite margins than the first. Investigating the habits of the patient could be helpful [11,12]. Follow-ups must be on-going even if the histology does not manifest dysplasia, since hyperkeratosis could evolve towards a malignant transformation due to the reoccurrence of the trauma.

It has been many years since laser therapy has been introduced in dentistry and in particular in periodontal field where it succeeded in eliminating pathogenic bacterial niches in inaccessible areas, such as deep pockets, root concavities and furcation areas in a less- traumatic way [13]. One of the most effective laser therapy in decontaminating periodontal pockets and in giving improvement in CAL, PPD and BOP was the photodynamic therapy based on the use of three components: light, oxygen free radicals and photosensitizer [14–17].

Another option in the treatment of chronic periodontitis is the use of a desiccant agent as an adjunct to scaling and root planing (SRP). This protocol resulted in a greater reduction in clinical, microbial and inflammatory mediators compared to SRP alone [18].

The aim of the study is to find a laser-assisted protocol for the surgical excision of leucoplakia and hyperkeratosis that can both improve the clinical aspect of the lesion and be sustainable for patients. The null hypothesis has been identified in the following statement: the treatment is effective and efficient at the same time; where effectiveness was tested with the following criteria: size of the lesion, tactile perception, discomfort and pain; and efficiency with the following criteria: pain and discomfort perceived during the treatment.

2. Materials and Methods

The study group consists of 20 patients affected by leucoplakia and hyperkeratosis selected between October 2017 and March 2018 at the Dr. Peset University Hospital in Valencia and which have been subjected to a laser treatment.

The sample was selected following specific inclusion and exclusion criteria.

2.1. Inclusion Criteria

Execution of a biopsy and following histopathological examination in order to have the certainty to select only leucoplakias and hyperkeratosis,

Possibility of completing the clinical trial.

2.2. Exclusion Criteria

Histopathological presence of dysplasia,

Systemic diseases such as HIV, HBV and HCV infections, celiac disease or physiological conditions such as pregnancy,

Having performed treatment for hyperkeratosis and leucoplakia in the 2 months prior to the clinical trial.

During the first visit clinical data of each patient was collected and the informed consent to the treatment was signed. The patients were also asked about their habits, in order to identify those, which could be detrimental. The study was approved by the Clinical Research Ethics Committee (CEIC) (code 09/093), following the principles of Helsinki for human experimentation.

After collecting this data and having examined the patient both from a systemic and dental point of view, an examination of the lesions was carried out, thus evaluating:

- The site presenting the lesion,
- The histological examination,
- The size of the lesion, in order to have this data, a periodontal probe was used and the greatest distance between the sides of the lesion was measured,
- How long the lesion was present.

As previously stated, the biopsy and its subsequent histological examination were fundamental: the presence of cellular atypia, dysplasia or malignant lesions could be excluded, thus making the patient suitable for receiving laser treatment.

The biopsy was performed at the most significant point of the lesion, in order to also remove a part of the healthy tissue, useful for identification in the subsequent histological examination. First of all a 1.8 mL vial of mepivacaine and adrenaline 1:100,000 (OPTOCAIN, Molteni Dental, Milan, Italy) at the perilesional level were inoculated. After having disinfected the oral mucosa with 0.2% chlorhexidine, the part of tissue to be extracted was delimited with a 4 or 5 mm biopsy punch and subsequently extracted with a scalpel, taking care to remove all the layers of the mucosa. A suture was carried out with simple detached stitches in 4/0 silk, which was removed 7 days later.

The sample was fixed in 10% formalin (BIO-OPTICA, Milan, Italy) and sent to the pathological anatomy laboratory of the Dr. Peset University Hospital in Valencia. Here it was included in paraffin then 5-micron thick sections were made, stained with haematoxylin and eosin and analysed under an optical microscope. Approximately two weeks later the histological examination report was received.

The treatment consisted of four sessions a week and a follow-up visit one month after the last treatment session. *Surgical technique*: the surgical technique provided no anaesthesia and a superficial and fast touch of the lesion, in this way we could observe the removal of the pathological area from the underlying mucosa, which was then removed manually with the aid of anatomical tweezers. Once the white area was removed the underlying mucosa appeared to be clinically normal and healed for second intention.

The protocol stated that at each session before treatment and in the final follow-up session:

- The lesion was measured with periodontal probe and the results were reported in mm,
- The appearance of the lesion was reported,
- The discomfort was measured with an arbitrary numerical scale from 0 to 5 and the type of discomfort that the patient perceived was noted,
- The pain caused by the lesion was measured, with an NRS scale (Numerical Rating Scale) [19],
- The tactile perception was evaluated, that is the sensation of roughness that the patient felt passing the tongue on the area where the lesion was present, with an arbitrary numerical scale from 0 to 5.

During each session some aspects that were presented during the seven days since the previous session were assessed in particular:

- Discomfort measured with an arbitrary numerical scale from 0 to 5.
- Pain caused by the lesion, measured with the NRS scale [19].
- Tactile perception, measured with an arbitrary numerical scale from 0 to 5.

To perform the treatment a diode laser was used which had a wavelength of 940 nm, the fibre had a length of 9 mm and diameter of 300 microns; a pulsed mode was selected with a Ton corresponding to Toff and corresponding to 10^{-3} s, the period corresponds to the summary of Ton and Toff that is 2×10^{-3} s. It is also known that the frequency and the period of every electromagnetic wave are inversely proportional to each other for the equation $\nu = \frac{1}{T}$; reporting the equation in our case shows that

$$v = \frac{1}{2 \cdot 10^{-3}} = \frac{10^3}{2} = 5000 \text{ Hz}. \tag{1}$$

When using the pulsed mode, there is both a peak power and average power where the latter is defined as the percentage of Ton in the period. In this case it stands at 50%. Having set the peak power at 1.8 W, the average constant power during the emission corresponds to 0.9 W; this allows intervention without using anaesthetic devices, unless specifically requested by the patient.

Before laser-assisted treatment the fibre was activated, using the special supplied activator. Both the operator and the patient wore protective glasses. The laser was used with a contact on the mucosa for about 1 min, the tissue was then left to cool for 30 s and the whole sequence was repeated in order to perform a total of four sessions in one visit. After each visit the patient was instructed to avoid too hot and too spicy foods for a few days.

2.3. Statistical Analysis

The data collected during the study were inserted into a database in Excel and the statistical analysis was performed with one of its applications (Stat). The data were standardized. A test unit was used in order to find the best criteria that could best evaluate the laser- treatment. For the analyses there was an intra-examiner agreement. In the first part we analysed the characteristics of the sample, while in the second phase we focused on the data during the treatment in order to obtain information regarding both the efficiency and the sustainability of the treatment.

3. Results

3.1. Sample Characteristics

20 patients were included in this study, of which 9 were affected by hyperkeratosis and 11 were affected by leucoplakia.

The sample consisted of 11 men (55%) with an average age of 55 years old and 9 women (45%) with an average age of 67 years old, the average age of the sample was of 61 years old. The site of onset of lesions in the oral cavity of the different patients was homogenous (Figure 1); the year of onset can be seen in Figure 2.

60% of patients had metals of dental origin such as gold; titanium or mercury derived from amalgam, 65% were carriers of both crowns and mobile prostheses and 45% of individuals had resin restorations on various elements. 35% of patients used alcoholic mouthwash, 25% had malocclusions in particular mono and bilateral cross-bite or third class and that 35% show detrimental habits, in particular episodes of morsicatio.

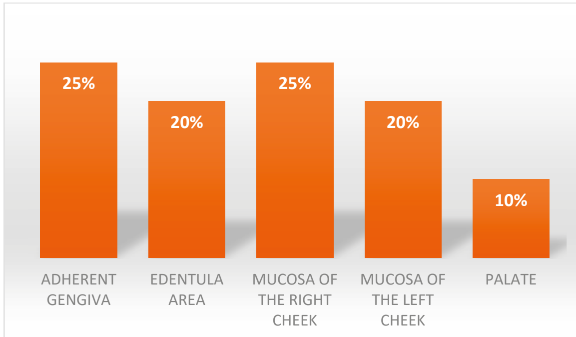


Figure 1. Areas of lesions' presentation.

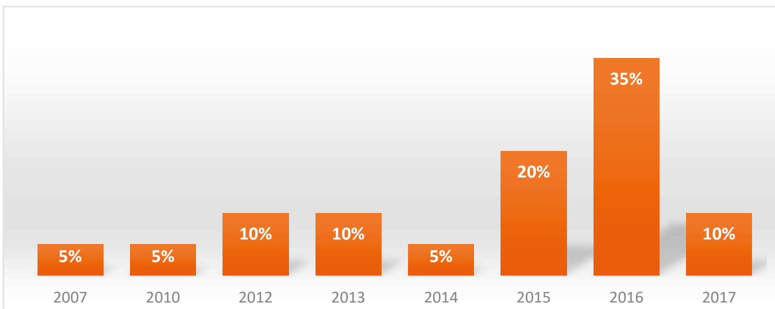


Figure 2. Year of onset of the lesions.

3.2. Analysis of the Lesion Size

The variation of this parameter is evaluated consecutively with one-week intervals between one observation and the next. As the dimension at time T0 was considered as an independent variable and as the variable dependent was dimension at time T1, the coefficient of the regression of this analysis was 0.81. This meant that the lesion dimension between the two time intervals taken into consideration, reduced by an average of 19%. This same reasoning was applied to the independent and dependent variables respectively before T1 and T2, then T2 and T3 and finally T3 and T4. The coefficients of the regression analysis obtained are, in order, 0.39; 0.60; 0.80. The size between the different time intervals is reduced on average by 61%, 40% and 20% respectively. The P value was always inferior to 0.05% (Table 1).

Table 1. Regression analysis, the parameter that has been used was: coefficient, significance and coefficient of determination.

Regression: Lesion Dimension T1		
	Constant	**Lesion Dimension T0**
Coefficient	−8.5327	0.814990
Std error of coef	11.7132	0.03492
t-ratio	−0.7285	23.3347
Significance	4756.97%	0.0000 %
Beta-weight		0.9839
Standard error of regression		42.4721
Coefficient of determination		96.80%
Adjusted coef of determination		96.62%
Number of observation		20
Residual degrees of freedom		18
t-statistic for computing 95% confidence intervals		2.1009

3.3. Discomfort Analysis

The perception of the discomfort caused by the lesion before the laser therapy results, on average, in the scale selected for the measurement of this parameter, approximate to zero. The discomfort perceived during the treatment stands at average values (between 2.5 and 2.0) and tends to decrease less evidently in the short term, while in the medium to long term it follows a decreasing trend from values of 2.6 at T0 to 1.9 at T4. The values show that on average the variable assumes significant values for all observations and that the trend decreased particularly between the second and third week of treatment. The discomfort was always described as a burning sensation (Figure 3).

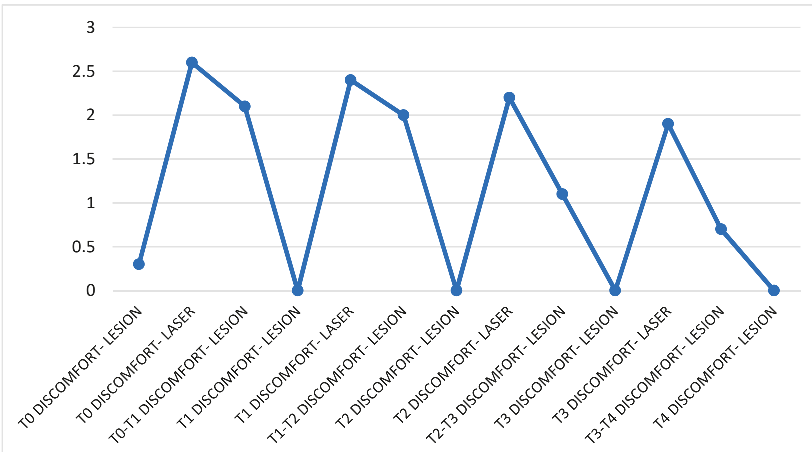


Figure 3. Trend of discomfort perceived throughout the treatment. The data were divided as follow: discomfort felt before the patient was subjected to treatment, during irradiation of the lesion and in the days following the session. The data refer to the entire month of observation.

3.4. Tactile Perception

It remains substantially unchanged between before and during the treatment but the overall tendency of the values is to decrease almost constantly over time, starting at an average of 3.5 at T0 to 0.5 at T4 (Figure 4).

3.5. Pain: Not Perceived

As previously stated in this study both patients with simple hyperkeratotic lesions and patients with leucoplakia were included. In order to highlight if the treatment has a different efficacy between the two lesions, the patients were divided into two groups taking into consideration the nature of the lesions. For each group we analysed the same parameters used in the previous paragraphs and then we compared the results. Regarding discomfort, tactile perception and pain there are not substantial differences neither between the two groups nor between the groups and the general trend previously described. The only difference is in the variation of the size: leucoplakia follows the general trend, in fact the size reduced by 17% between T0 and T1, between T1 and T2 the size reduced by 63%, between T2 and T3 a variation of 39% occurred and finally between T3 and T4 there was a variation of 36%; in all cases the results were statistically significant. With regard to patients with hyperkeratosis, the dimensional variation shows a constant reduction during all the observations, the data showed: between T0 and T1 there was a reduction of 20%, between T1 and T2 the dimension varied by 28% and continued to do so in the following weeks of observation.

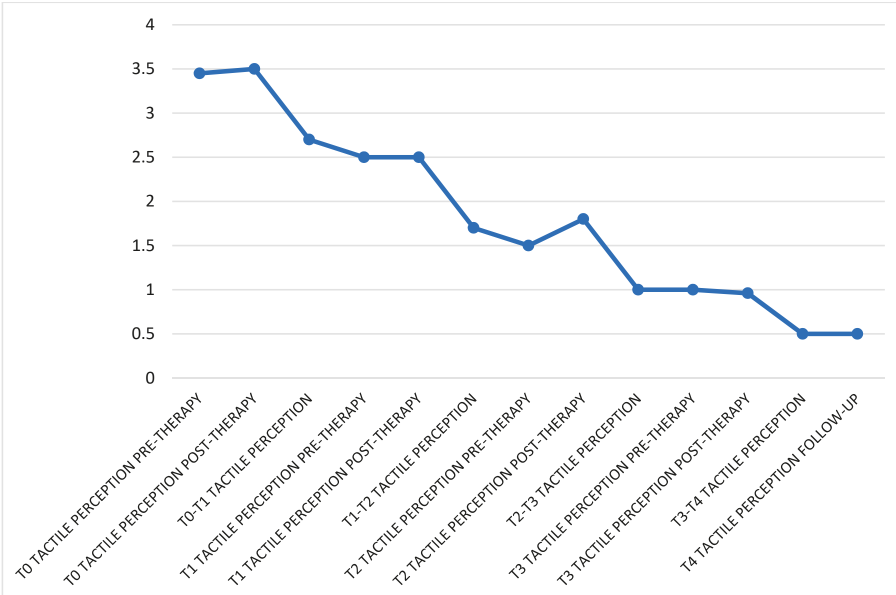


Figure 4. In this graph can be seen the trend of tactile perception perceived throughout the treatment. The data were divided as follow: perception felt before the patient was subjected to treatment, right after the irradiation of the lesion and in the days following the session. The data refer to the entire month of observation.

4. Discussion

Laser-assisted treatment has some advantages compared to traditional surgical techniques.

The main advantage common to all types of laser treatment is photocoagulation of lymphatic, hematic and nerve endings thus giving less intraoperative bleeding, less oedema and post-intervention pain. The placement of sutures is rarely necessary.

The diode laser is not indicated as the main laser for soft tissue surgery but its versatility of use led us to choose it for the study.

The diode laser is a semiconductor laser and exists in different wavelengths: from 980 nm that has greater cutting capacity and fibres that do not need to be activated, up to 810 nm which has a more biostimulatory capacity. This laser therefore has dual functionality [14,17,20]. The protocol used in this study aimed to exploit both the cutting capacity and the biostimulatory capacity; in particular, the latter was always used after the surgical phase. This helped to further reduce postoperative discomfort, oedema and to have a better, faster healing without retractable cicatrized outcomes or functional outcomes. Studies report that the laser beam activity on myofibroblasts result in reduced proliferation which results in minor dysfunctional outcomes and furthermore impacts on vessel proliferation, the synthesis of collagen and the anti-inflammatory capacity. These effects are closely related to the laser settings (fibre, power, exposure time). The best results, in terms of healing, occur with a fluency of about 4 J/cm^2 [21–24].

Some studies refer to the use of lasers to treat leukoplasic and hyperkeratotic lesions but the various studies do not report the same parameters with regard to power settings, exposure time and application distance and fibre diameter [9,10,25–27].

Consequently, in our study it was difficult to compare our parameters with previously reported in the literature, because of a non-homogeneous number of terms of comparison.

Other studies reported the evaluation of the recurrence rate and the rate of malignant transformation [1,7,23,28], however no mention about the efficacy of the treatment in the short term, sustainability or invasiveness was reported.

Our protocol involves the use of the diode laser in pulsed mode with a Ton and a Toff of 110^{-3} s and a frequency of 5000 Hz. These settings were chosen in order to allow, during the emission phase of the ray, a thermal relaxation time corresponding to the Toff, which resulted in diminished overheating of tissues.

The pulsed mode was chosen in order to avoid the harmful overheating of the tissue already mentioned above; this mode allowed us to alternate a lower average power, that corresponded to the percentage of Ton in the period, with a high power peak equal to the maximum power set. In our protocol the percentage of Ton in the period is 50%, thus setting an average power of 0.9 W and a peak power of 1.8 W. With these settings anaesthesia was not required to perform the treatment unlike all the other protocols found in literature.

In order to prevent the treatment from becoming unsustainable for the individual, it was decided to irradiate the tissue for a relatively short time (about 1 min) and then leave 30 s of relaxation. This was a variation of what was found in the literature where the irradiation time was on average over 8 min [8,9,23,29].

To evaluate the effectiveness of laser-assisted treatment, the following parameters and their behaviour were taken into account during the observation period:

- Size of the lesion.
- Tactile perception.
- Discomfort.
- Pain.

The treatment was defined effective if there was a decreasing trend of all the parameters listed above, partially effective if only some parameters reduced or not effective if, compared to the beginning of the observation, the parameters have remained unchanged.

In order to evaluate the efficiency of the treatment and its sustainability for the patient, the pain and discomfort perceived during the irradiation of the lesion were considered. Treatment is considered sustainable for the patient if these parameters show medium-low values and have a decreasing trend during the four weeks of observation, if this does not happen the treatment is considered unsustainable for the patient.

4.1. Effectiveness

Analysis of the data shows that treatment can be defined as effective in 15 patients (75%), partially effective in 5 subjects (25%) and not effective in any of the individuals.

In five patients, belonging to the group where the treatment can be defined as effective, the lesion had completely disappeared in four weeks; while in two individuals the lesion had regressed already during the third week.

In those where the treatment is partially effective, the size of the lesion and the discomfort remained constant. In these cases the lesion did not decrease in size in a clear way but the whitish pigmentation appeared less intense.

4.2. Efficiency

The treatment can therefore be considered efficient for the following reasons: first, the values reported by the subjects are average values and their trend is, even if slightly, decreasing; all patients completed the clinical trial and reported less post-intervention discomfort respect to the initial biopsy performed with a blade.

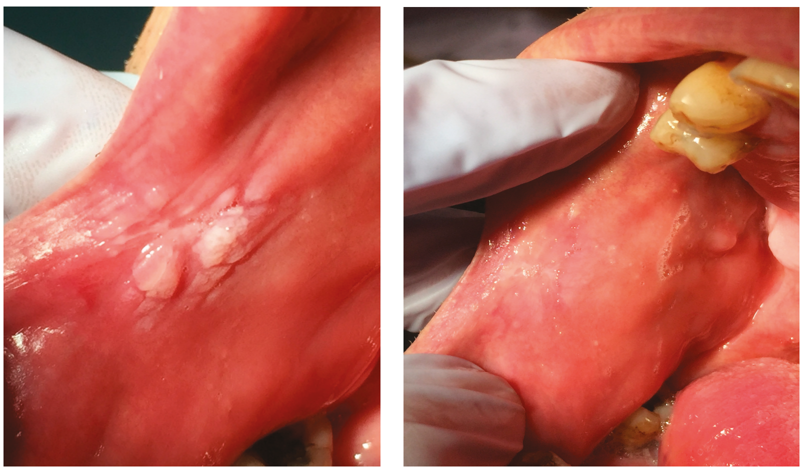
Because no anaesthesia was administered and so neither adrenalin, there were the possibility of having an intraoperative bleeding that could compromise the visibility of the operative field. Thanks to the use of the laser we could provide a bloodless operating field with an excellent visibility of the lesion margins and a smoother treatment, defined the low invasiveness of laser-assisted treatment.

Since the margins were well photocoagulated, no stitches were placed but the area was left to heal by second intention.

After seven days, the complete epithelialization of the area was observed in almost all cases, only in two cases there was still presence of fibrin deposits. None of the patients had an infection of bacterial origin of the wound.

At the end of the treatment there were no retrospective cicatricle outcomes. Regard to lesions that were placed in areas of functional interest, such as the labial commissure, they did not report any functional existences of any kind.

In the following figures the data from lesions treated with our protocol at time T0 and four weeks later, at time T4 (Figures 5a,b, 6a,b and 7a,b)



(**a**) (**b**)

Figure 5. (**a**) Hyperkeratosis during the first visit: the area appears whitish and irregular, (**b**) The same area four weeks after treatment.

(**a**)

Figure 6. *Cont.*

(b)

Figure 6. (**a**): Mucosa of the left cheek of a patient affected by leucoplakia right after the removal of stitches positioned after biopsy sampling. (**b**): Same area three weeks after laser treatment.

(a)

(b)

Figure 7. (**a**) Whitish and irregular area in the gingival mucosa. (**b**) Three weeks after the beginning of the treatment.

5. Conclusions

On the basis of our results, laser-assisted treatment can be considered successful. In the future it will be possible to recruit more patients in order to correlate systemic factors, such as diabetes, anaemia or gastroesophageal reflux, with the effects of the treatment in order to evaluate if those factors influence the result of the therapy. Another aspect to consider could be the rate of recurrence and malignant transformation in a longer follow-up; in this study we follow the patients for only one month after the end of the treatment and no significant differences were found since the last irradiation.

In conclusion we can state that the treatment is a valid alternative to conventional surgery with a scalpel if the lesion is bigger than 8 mm, below this limit it is better to completely eliminate the lesion during a biopsy. Lesions with a maximum size of 2 cm can regress during the four sessions set, obviously the larger the lesion, the more time needed to allow the lesion to regress.

Author Contributions: Conceptualization, D.L.; Data curation, A.L.; Investigation, F.G. and J.S.R.; Supervision, F.C.; Visualization, D.D.S.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bokor-Bratić, M. Prevalence of oral leucoplakia. *Med. Pregl.* **2003**, *56*, 552–555. [CrossRef]
2. Gönen, Z.; Asan, C.Y.; Etöz, O.; Alkan, A. Oral leucoplakia associated with amalgam restorations. *J. Oral Sci.* **2016**, *58*, 445–448. [CrossRef]
3. Kazanowska-Dygdala, M.; Dus, I.; Radwan-oczko, M. The presence of Helicobacter pylori in oral cavities of patients with leucoplakia and oral lichen planus. *J. Appl. Oral Sci.* **2016**, *24*, 18–23. [CrossRef] [PubMed]
4. Chaves, A.L.F.; Silva, A.G.; Maia, F.M.; Lopes, G.F.M.; de Paulo, L.F.B.; Muniz, L.V.; Dos Santos, H.B.; Soares, J.M.A.; Souza, A.A.; de Oliveira Barbosa, L.A.; et al. Reduced CD8+ T cells infiltration can be associated with a malignant transformation in potentially malignant oral epithelial lesions. *Clin. Oral Investig.* **2018**. [CrossRef] [PubMed]
5. Ardiuno, P.G.; Bagan, J.; El-Naggar, A.K.; Carrozzo, M. Urban legends series: Oral Leucoplakia. *Oral Dis.* **2013**, *19*, 642–659. [CrossRef]
6. Kumar, A.; McCaul, L.C.J.; Kerawala, C.J.; Coombed, D.; Godden, D.; Brennan, P.A. How should we manage oral leucoplakia? *Br. J. Oral Maxillofac. Surg.* **2013**, *51*, 377–383. [CrossRef] [PubMed]
7. Monteiro, L.; Barbieri, C.; Martins, M.; Salazar, F.; Vescovi, P.; Meleti, M.; Pacheco, J. Type of surgical treatment and recurrence of oral leucoplakia: A retrospective clinical study. *Oral Med. Pathol.* **2017**, *22*, 520–526.
8. Modegas-Vegara, A.; Hueto-Madrid, J.; Chimenos-Küstner, E.; Bescos-Atin, C. The treatment of oral leucoplakia with the CO_2 laser: A retrospective study of 65 patients. *J. Cranio-Maxillo-Fac. Surg.* **2015**, *43*, 677–681. [CrossRef]
9. Del Corso, G.; Gissi, D.B.; Tarsitano, A.; Costabile, E.; Marchetti, C.; Foschini, L.M. Laser evaporation versus laser excision of oral leucoplakia: A retrospective study with long-term follow-up. *J. Cranio-Maxillo-Fac. Surg.* **2015**, *43*, 763–768. [CrossRef]
10. Chainani-Wu, N.; Madden, E.; Sim, C.; Collins, K.; Silverman, S. Clinical predictors of oral leucoplakia recurrence following CO_2 laser vaporization. *J. Cranio-Maxillo-Fac. Surg.* **2015**, *43*, 1875–1879. [CrossRef]
11. Grammer, R.; Lerman, M.A. Keratosis of unknown significance and leucoplakia: A preliminary study. *Oral Maxillofac. Pathol.* **2014**, *118*, 713–724.
12. Bellato, L.; Martinelli, P.; Lombardi, T. Alveolar ridge keratosis—A retrospective clinicopathological study. *Head Face Med.* **2013**, *9*, 2–5. [CrossRef] [PubMed]
13. Matarese, G.; Ramaglia, L.; Cicciù, M.; Cordasco, G.; Isola, G. The Effects of Diode Laser Therapy as an Adjunct to Scaling and Root Planing in the Treatment of Aggressive Periodontitis: A 1-Year Randomized Controlled Clinical Trial. *Photomed. Laser Surg.* **2017**, *35*, 702–709. [CrossRef] [PubMed]
14. Caccianiga, G.; Rey, G.; Paiusco, A.; Lauritano, D.; Cura, F.; Ormianer, Z.; Carinci, F. Oxygen high level laser therapy is efficient in treatment of chronic periodontitis: A clinical and microbiological study using PCR analysis. *J. Biol. Regul. Homeost. Agents* **2016**, *30* (Suppl. 1), 87–97. [PubMed]
15. Yadwad, K.; Veena, H.R.; Patil, S.; Shivaprasad, B.M. Diode laser therapy in the management of chronic periodontitis—A clinico-microbiological study. *Int. Med. Appl. Sci.* **2017**, *9*, 191–198. [CrossRef] [PubMed]
16. Meimandi, M.; Ardakani, M.; Nejad, A.; Yousefnejad, P.; Saebi, K.; Tayeed, M. The Effect of Photodynamic Therapy in the Treatment of Chronic Periodontitis: A Review of Literature. *J. Lasers Med. Sci.* **2017**, *8* (Suppl. 1), S7–S11. [CrossRef]
17. Roncati, M.; Lauritano, D.; Cura, F.; Carinci, F. Evaluation of light-emitting diode (LED-835 NM) application over human gingival fibroblast: An in vitro study. *J. Biol. Regul. Homeost. Agents* **2016**, *30* (Suppl. 1), 161–167.
18. Isola, G.; Matarese, G.; Williams, R.C.; Siciliano, V.I.; Alibrandi, A.; Cordasco, G.; Ramaglia, L. The effects of a desiccant agent in the treatment of chronic periodontitis: A randomized, controlled clinical trial. *Clin. Oral Investig.* **2018**, *22*, 791–800. [CrossRef]
19. Sirintawat, N.; Sawang, K.; Chaiyasamut, T.; Wongsirichat, N. Pain measurement in oral and maxillofacial surgery. *J. Dent. Anesth. Pain Med.* **2017**, *17*, 253–263. [CrossRef]

20. Usama, R.K.; Onkar, S.; Birangane, R.; Chaudhari, S.; Kulkarni, A. Treatment of oral leucoplakia with diode laser: A pilot study on Indian subjects. *Asian Pac. J. Cancer Prev.* **2015**, *16*, 8383–8386.
21. Kuribayashi, Y.; Tsushima, F.; Morita, K.; Matsumoto, K.; Sakurai, J.; Uesugi, A.; Sato, K.; Oda, S. Long-term outcome of non-surgical treatment in patients with oral leucoplakia. *J. Cranio-Maxillo-Fac. Surg.* **2015**, *51*, 1020–1025.
22. Natekar, M.; Raghuveer, H.; Rayapati, D.; Shobha, E.; Prashanth, N.; Rangan, V. A comparative evaluation: Oral leukoplakia surgical management using diode laser, CO_2 laser and cryosurgery. *Oral Surg.* **2017**, *9*, 779–784. [CrossRef] [PubMed]
23. Brouns, E.R.E.A.; Baart, J.A.; van der Waal, I.; Karagozoglu, K.H.; Aartman, I.H.A. Treatment results of CO_2 laser vaporisation in a cohort of 35 patients with oral leukoplakia. *Oral Dis.* **2013**, *19*, 212–216. [CrossRef] [PubMed]
24. Maloth, K.; Velpula, N.; Kodangal, S.; Sangmesh, M.; Meka, S.U. Photodynamic Therapy—A Non-invasive Treatment Modality for Precancerous Lesions. *J. Lasers Med. Sci.* **2016**, *7*, 30–36. [CrossRef] [PubMed]
25. Van der Hem, P.S.; van der Wal, J.E.; Nauta, J.M.; Roodenburg, J.L.N. The results of CO_2 laser surgery in patients with oral leukoplakia: A 25 year follow up. *Oral Oncol.* **2005**, *41*, 31–47. [CrossRef] [PubMed]
26. Vivek, V.; Balan, A.; Jayasree, R.S.; Sreelatha, K.T.; Gupta, A.K. Three-year follow-up of oral leukoplakia after neodymium: Yttrium aluminum garnet (Nd:YAG) laser surgery. *Laser Med. Sci.* **2008**, *23*, 375–379. [CrossRef] [PubMed]
27. Comacho-Alonso, L.J. Comparison of pain and swelling after removal of oral leukoplakia with CO_2 laser and cold knife: A randomized clinical trial. *Med. Oral* **2013**, *18*, 38–44.
28. Tambuwala, A.; Sangle, A.; Khan, A.; Sayed, A. Excision of Oral Leukoplakia by CO_2 Lasers Versus Traditional Scalpel: A Comparative Study. *J. Maxillofac. Oral Surg.* **2014**, *13*, 320–327. [CrossRef]
29. Romeo, U.; Russo, N.; Palaia, G.; Tenore, G.; Del Vecchio, A. Oral proliferative verrucous leukoplakia treated with the photodynamic therapy: A case report. *Ann. Stomatol.* **2014**, *2*, 77–80. [CrossRef]

International Journal of
Environmental Research and Public Health

MDPI

Article

Hydrophobic Organic Pollutants in Soils and Dusts at Electronic Waste Recycling Sites: Occurrence and Possible Impacts of Polybrominated Diphenyl Ethers

Chimere May Ohajinwa [1,*], Peter M. Van Bodegom [1], Qing Xie [2], Jingwen Chen [2], Martina G. Vijver [1], Oladele O. Osibanjo [3] and Willie J.G.M. Peijnenburg [1,4,*]

1 Institute of Environmental Sciences (CML), Leiden University, P.O. Box 9518, 2300 RA, Leiden, The Netherlands; p.m.van.bodegom@cml.leidenuniv.nl (P.M.V.B.); vijver@cml.leidenuniv.nl (M.G.V.)
2 Key Laboratory of Industrial Ecology and Environmental Engineering (Ministry of Education), School of Environmental Science and Technology, Dalian University of Technology, Dalian 116024, China; qingxie@dlut.edu.cn (Q.X); jwchen@dlut.edu.cn (J.C)
3 Department of Chemistry, University of Ibadan, Ibadan 200282, Nigeria; oosibanjo@yahoo.com
4 Center for Safety of Substances and Products, National Institute of Public Health and the Environment (RIVM), P.O. Box 1, 3720 BA, Bilthoven, The Netherlands
* Correspondence: chimeremay@gmail.com or chimeremayq@yahoo.co.uk (C.M.O); peijnenburg@cml.leidenuniv.nl (W.J.G.M.P.); Tel.: +31(0)-715277434 (W.J.G.M.P.)

Received: 12 December 2018; Accepted: 25 January 2019; Published: 28 January 2019

Abstract: Concerns about the adverse consequences of informal electronic waste (e-waste) recycling is increasing, because e-waste contains some hazardous substances such as polybrominated diphenyl ethers (PBDEs) which is used as flame retardants in electronics. There is dearth of information on the concentrations of PBDEs and the pattern of distribution at the various e-waste recycling sites in Nigeria. This study therefore measured the concentrations of 13 PBDE congeners, in top soils (0–10 cm) and in various dust samples from different e-waste recycling sites (burning, dismantling, repair). PBDE concentrations at e-waste sites were compared with the concentrations in samples from corresponding control sites in three study locations in Nigeria (Lagos, Ibadan, and Aba). There were significant differences in the level of PBDEs congeners between each of the e-waste recycling sites and the corresponding control sites. The levels of PBDEs at the e-waste recycling sites exceeded the levels at the controls sites by a factor of 100 s to 1000 s. In general, PBDE concentrations at the e-waste sites decreased with the intensity of the e-waste recycling activities: burning sites > dismantling sites > repair sites > control sites. Our results suggest that the informal e-waste recycling has negative impacts on the enviroment and human health.

Keywords: electronic waste; informal recycling; PBDEs; soil; dust; Nigeria

1. Introduction

Across the globe, electronic or electrical devices have become indispensable in our daily lives and the use of electronic electrical device is growing at great speed. It is characterized by an increasing number of users and rapid technological advances driven by efficiency, social and economic development. Many people now own multiple personal electronic devices such as information and communication technology (ICT) devices, but the life span of these devices are getting shorter mainly because they become obsolete more quickly compared to the past. In addition, most of these devices are disposed even before they become dysfunctional so as to make space for newer devices with better specifications/functions. Therefore, exponentially growing demand for electronic equipment has led to a rapid increase in the rate of electronic waste (e-waste) generated. e-Waste, also known as Waste Electrical and Electronic Equipment (WEEE), is one of the fastest growing municipal waste streams [1].

In 2016, 44.7 million metric tonnes (Mt) of e-waste were generated globally, and this amount is expected to increase to 52.2 million metric tonnes by 2021 [2].

The concern about e-waste is not only about the volumes generated but also about the unsafe methods used in recycling the electronics, mainly in developing countries. It is reported that 80% of e-waste generated globally is recycled informally or simply dumped at dumpsites or landfils in developing counties in Asia and in Africa [1,2]. The informal/unsafe methods of managing e-waste in developing countries perpetuate due to absence of infrastructure for appropriate waste management, an absence of end-of-life product take-back system, or implementation of extended producer responsibility (EPR) schemes by manufacturers which is enforced in developed countries [1,2]. In addition EPR has been extended beyond the e-waste sector to include packaging waste generally in developed countries [3]. e-waste contains several different substances, some of which are compounds of potential concern which include products of incomplete combustion, reformation products after combustions, elements such as metals-lead, mercury, cadmium, arsenic, beryllium, and flame retardants such as polybrominated diphenyl ethers (PBDEs) [4,5]. These mixtures of different substances, covering both chemicals present in EEE components and mixtures of chemicals released during e-waste processing, may pose significant implications for human health and environmental safety [6,7].

PBDEs are a class of persistent organic pollutants (POPs) which have been used as flame retardants in many consumer products such as electronic equipment, textiles, furniture, automobile seats and other consumer products since the 1970s. PBDEs are also a class of additive brominated flame retardants (BFRs), which are not covalently bound to the products (polymer matrices). In case of a fire, bromine radicals are released as a result of thermal energy. These radicals decrease the flame, and they reduce heat and carbon monoxide production. Because the PBDEs are not permanently bound to the polymer matrices, they are widely dispersed in the environment. In total, PBDE has 209 congeners, which are dependent on the number and position of the bromine atoms on the two-phenyl rings. Approximately 56,418 metric tons of PBDEs were produced globally in 2003 [8]. PBDEs have mostly been produced and used in three commercial groups: pentabromodiphenyl ether (penta-BDE; $C_{12}H_5Br_5O$), octabromodiphenyl ether (octa-BDE; $C_{12}H_2Br_8O$), and decabromodiphenyl ether (deca-BDE; $C_{12}Br_{10}O$) with about 11%, 6%, and 83% of global PBDEs respectively [8,9]. Penta- and octa-BDEs are not single chemicals. Penta-BDE comprises tri-, tetra-, penta-, and hexa-BDE; Octa-BDE comprises hexa-, hepta-, and nona-BDEs; and deca-BDE is a single chemical (BDE 209). The major sources of PBDEs are the manufacturing sector, product application, recycling processes, thermal processes, and wastes disposal reservoirs.

PBDEs are highly persistent in the environment, bioaccumulative in food chain and have a high potential for long-range environmental transport, meaning they can deposit far from their source. These chemicals have been detected in humans and in increasing concentrations in various environmental matrixes, including air, water, soil, sediment, animals and foods in all regions of the world. There is evidence of harmful effects in humans and wildlife, which includes endocrine disruption, immunotoxicity, reproductive toxicity, effects on fetal/child development [10–12], thyroid and neurologic function [13], and cancer [14]. Due to the environmental and health concerns, penta-BDE and octa-BDE have been banned in the European Union and voluntarily phased out in the USA since 2004 [15,16]. Recently, the European Commission restricted use of deca-BDE [17]. PBDEs are listed as persistent organic pollutants (POPs) by the Stockholm Convention [18], while deca-BDE (BDE-209) has been classified as a possible human carcinogen by the United States Environmental Protection Agency [8]. Due to the ban of PBDEs, a number of alternative flame retardants has been introduced, which include: organophosphates esters (OPEs) and a range of other brominated and chlorinated novel flame retardants such as tetrabromobisphenol-A (TBBPA), hexabromocyclododecane (HBCD), bis(2,4,6,-tribromphenoxy) ethane (BTBPE), and several phosphate based compounds, such as triphenyl phosphate [19–22].

Importation of electronics is one major way PBDEs are exported to developing countries such as China [5], India [23], Ghana [24], and Nigeria [25]. When e-wastes are informally recycled using crude methods such as manual dismantling, smelting, and open burning, this leads to incomplete combustion, consequently releasing mixture of hazardous chemicals, including PBDEs, which in turn cause environmental pollution and health problems. PBDE as additive flame retardants are more easily released into the environment than the reactive flame retardants. When released they are attached to particles and transported via various environmental media to distances far from the emission sites. PBDEs enter the environment through multiple pathways, such as emission during manufacturing, from products in use, from combustion, by leaching from landfills, or from recycling of products at the end of their life such as electronics at end-of-life (e-waste) [14]. In the environment, soil and dust are the main environmental receptors of chemical emissions from informal e-waste recycling. Therefore, soils can be secondary sources of emission of PBDEs and soils can contribute to the contamination of air, food and (drinking) water. Hence, soils and dust are the most important environmental media that can reveal the contaminants present in the environment [26]. Moreover, dust is a good indicator for contaminant levels in the atmosphere [27–31].

Data on PBDE concentrations in the environment as a result of informal e-waste recycling in Nigeria is scarce; therefore environmental control on PBDE could be a challenge without adequate information. To gain insight into the PBDE concentrations in the environment as a result of informal e-waste recycling activities we systematically collected soil and dust samples from different selected e-waste recycling sites and from corresponding control sites in three cities where e-waste is recycled in Nigeria. Hence, the objectives of this study were to (1) quantitatively assess the levels of PBDEs in top soils and dust because of different e-waste recycling activities (burning, dismantling, and repair sites); (2) determine the extent to which PBDE concentrations at e-waste sites exceeded the concentrations at control sites; (3) determine the activities that contribute most to the PBDE pollution to the environment; and (4) determine the distribution patterns of the various PBDE congeners. Most importantly, we hope that our findings could be a wake-up call to relevant stakeholders to devise effective interventions to reduce PBDE pollution caused by informal e-waste recycling without impeding the livelihood of the e-waste workers. Our findings are likely to be applicable to other locations or countries where informal e-waste recycling is practiced.

2. Materials and Methods

2.1. Study Locations and Sites

The methods employed in this study have been well detailed in our previous studies [32]. In brief, this study was conducted in three study locations/cities in Nigeria as depicted in the map as A,B,C, which are Ibadan, Aba, and Lagos respectively. The three study locations are some of the largest cities in Nigeria where e-waste is recycled [25]. In each study location, two e-waste recycling areas were selected. In Lagos, the selected sites were Computer village, Ikeja (6.593° N, 3.342° E) and Alaba international market Ojor (6.462° N, 3.191° E). In Ibadan, the selected sites were Ogunpa (7.383° N, 3.887° E) and Queens Cinema areas (7.392° N, 3.883° E). In Aba, the shopping centre (5.105° N, 7.369° E) and Port-Harcourt Road/Cementary (5.104° N, 7.362° E) and Jubilee road/St Michael's Road (5.122° N, 7.379° E) were selected (Figure 1, showing map of the study locations). Alaba international market is the largest market for new and second-hand electronics in West Africa, with approximately ten to fifteen containers arriving daily from Europe and Asia, with each container containing about 400,000 second-hand units [33]. Computer village Ikeja is a popular place where electronics and their parts (new and second hand) can be purchased and repaired. In Ibadan, the selected sites were Ogunpa and Queens Cinema areas. Ogunpa area is known for its activities in scrap/second-hand businesses, including electronics, while Queens cinema is known for sales and repair of both new and second-hand electronics. In Aba, the shopping centre and Port-Harcourt Road/Cementary and Jubilee road/St Michael's Road were selected. The shopping center area is the biggest market for new and used electronics, while the Port-Harcourt road/Cemetery area is known as an area for scrap/second-hand metal businesses in Aba. In Alaba, Lagos, we found only one big e-waste burning site, which is the largest, oldest, and most studied e-waste burning and dismantling site in Nigeria. In Ogunpa, Ibadan and Cemetery area Aba, the burning sites/spots were much smaller but more spread out in small clusters around the areas.

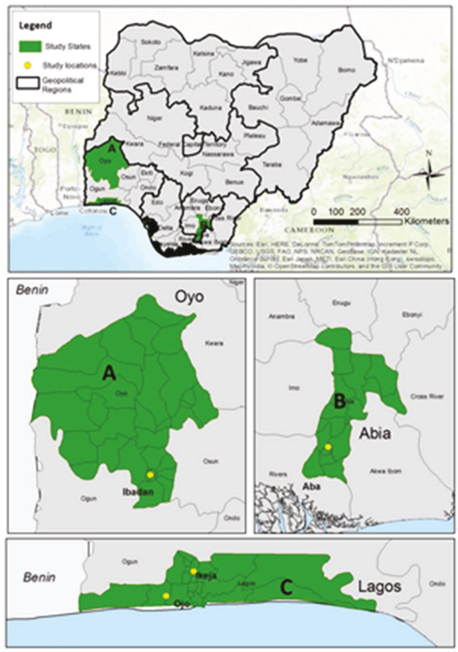


Figure 1. Map of Nigeria showing the study locations.

2.2. Study Design

A comparative cross-sectional study design was adopted to gain insights into the PBDE pollution levels at the e-waste recycling sites compared with non-e-waste sites (control sites) in Nigeria. The study is designed in a way that is a representative of other informal e-waste recycling sites. In each study location, a multi-stage random systematic sampling technique was used to include various groups of e-waste workers and e-waste recycling activities (burning, dismantling, and repair) in the selected areas. Soil and dust samples were collected from the selected sites. The control sites were about 500 m away from the e-waste recycling sites, and consisted of areas with reduced human activity such as play grounds, parks, fields, and a university garden. Three types of e-waste recycling activities sites (burning sites, dismantling sites, and repair sites/shops) were analysed.

Top soils (0–10 cm) and various dust samples were collected from the selected e-waste recycling sites. Type of sample collected on each site depends on how feasible it is to collect the sample; this led to an unbalanced design in soil sample collection. At burning sites, only top soil samples were collected, no dust samples were collected at the burning sites because burning activities take place on bare ground. Direct dust from the electronics, mainly from televisions, computers, printers, and air conditioners were also collected. The locations of the sampling spots were georeferenced using a global positioning system application on a phone. Limitations to the methods of sample collection to ensure representative samples were the absence of unpaved surfaces in some place; therefore samples were collected based on feasibility of sample type collection. This study took place in urban areas where most places are paved, these conditions led to an unbalanced number of each sample type.

However, this study design is a representative of the sites and major activities at informal e-waste recycling sites. Figure 2 presents a schematic flow diagram of the sample collection from the various e-waste sites in the three study locations. In the previous study by Ohajinwa et al. [32] which was on metal pollution at the same sites, a total of 62 samples. While for this study, a total of 56 samples consisting of 16 top soils (0–10 cm), 29 floor dust, five roadside dust, and six direct dust samples (collected from the inside and outside of electronic devices) were analysed. For the locations, 15, 26, and 29 samples were collected from Aba, Ibadan, and Lagos, see Figure 2. The difference in the number of samples was due to loss of samples and non-detection limits set for the PBDE analysis.

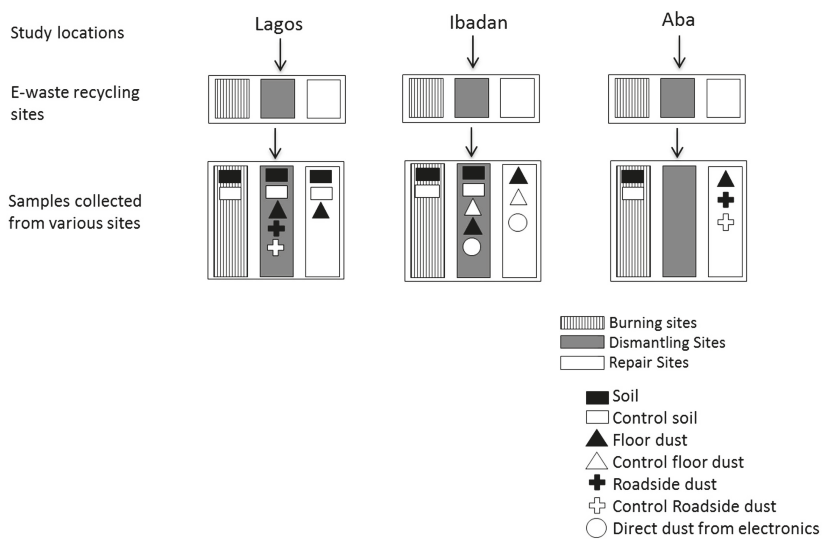


Figure 2. Schematic flow diagram of sample collection in the study locations.

2.3. Sample Collection and Preparation

First, the 10ml amber bottles and aluminium foils were treated in the laboratory. The amber bottles were washed with tap water and laboratory detergent, rinsed with a copious amount of tap water, rinsed with distilled water three times, treated with acetone and with hexane, and then oven-dried at 120 °C for 4 h to ensure no traces of POPs were present. Aluminium foils (for sample wraps on the field) were treated with acetone and hexane, then oven dried at 120 °C to ensure no traces of POPs in the aluminum foil.

On the field, for the soil sampling, each selected site was divided into grids of about 2 m to 10 m wide, depending on the size of the site. Samples were systematically collected from three to six points within each site. The samples were bulked together for the top soil to form a composite representative sample for the specific site. Soil samples were collected using a soil auger, and a soil trowel was used to the transfer soil from the soil auger into aluminum foil (sample wraps). To avoid cross contamination, the soil probe/auger and trowel were decontaminated (cleaned first with a brush and wiped thoroughly with wipes) before each sample collection at each sampling site. Dust samples were collected using fiber dusting brushes to gently sweep the dust and collect it with a dustpan. The soil and dust samples were wrapped in a treated aluminum foil, labelled, and transported to the laboratory. A total of 71 samples (56 samples from the e-waste recycling sites and 15 samples from control sites) were analysed. The total set consisted of 22 top soil (0–10 cm depth) samples, 30 floor dust samples, 13 roadside dust samples, and six direct dust samples. Soil and dust samples were air dried for 7 days, avoiding exposure to sunlight. The samples were homogenized, ground with a mortar and pestle, and sieved through a 1 mm mesh sieve to remove bigger particles. Next, they were transferred into individual 10 mL amber bottles, labelled and stored at −20 °C until shipping to the laboratory for analysis. The samples were collected between May and November 2015.

2.4. Chemicals and Materials

All the solvents used for extraction, purification and analysis were of high-performance liquid chromatography (HPLC) grade (Spectrum Chemical MFG. Corp., New Brunswick, NJ, USA). Silica gel (100–200 mesh) and neutral aluminum oxide (100–200 mesh) were for chromatography purpose (Sinopharm Chemical Reagent Co., Ltd, Shanghai, China), and they were activated before use (i.e., first washed with hexane/dichloromethane (v/v, 1/1) and then baked at 180 °C for 2 h). Acid silica gel (30% w/w) was prepared with activated silica gel and sulphuric acid before use. Anhydrous sodium sulfate (99% purity) and diatomaceous earth (DE, 100% purity) were purchased from Aladdin Ind. Corp. (Shanghai, China) and ThermoFisher Scientific (Waltham, MA, USA) respectively. They were baked at 400 °C for 4 h before use to remove any traces of organic matter.

A standard mixture solution of 14 PBDE congeners (BDE-COC) PBDEs (BDE-17, BDE-28, BDE-71, BDE-47, BDE-66, BDE-100, BDE-99, BDE-85, BDE-154, BDE-153, BDE-138, BDE-183, BDE-190, and BDE-209) and Individual standards of 4 PBDEs (BDE-77, BDE-206, BDE-207, BDE-208) and PCB-209 were purchased from Accu Standard (New Haven, CT, USA), while Isotopically labeled 13C-PCB-208 was purchased from Cambridge Isotope Laboratories (Tewksbury, MA, USA). We used ^{13}C-PCB-208 as the surrogate because, we first used chemical ionization source (CI source) to detect PBDEs with the characteristic ionic fragments, and CI source cannot identify the difference between ^{13}C-labeled PBDEs and unlabeled PBDEs. Secondly, ^{13}C-PCB-208 can be identified by CI source, and its characteristic ionic fragments contain ^{13}C labeled carbon. Also, the physiochemical properties of PCBs and PBDEs are similar with PBDEs.

2.5. Sample Extraction and Cleanup

For the PBDE analysis, from each of the samples, 5 g of homogenized sample was thoroughly mixed with 0.6g DE with a mortar and pestle. Each sample was thereafter spiked with 2ng ^{13}C- labeled PCB-208 and 10 ng PCB-209 standards, and allowed a static equilibration of 5 min in two cycles. The sample was then extracted using Thermo Scientific Dionex ASE 350 accelerated solvent extraction system with *n*-hexane/dichloromethane (v/v, 1/1) at 90 °C, 1500 psi. After extraction, acid washed copper sheets were added to the extracts to remove sulfur present in the samples. The extracts were evaporated to about 10 mL under a gentle stream of N_2, and transferred to a conical centrifuge tube. One mL of concentrated sulfuric acid (98%) was added to the concentrated extracts to carbonize part of the impurities present. The supernatants were transferred to a preconditioned glass cleanup column, which was packed with 0.5 cm neutral aluminum oxide, 3.0 cm neutral silica gel, 3.0 cm acid silica gel, and 1.0 cm anhydrous sodium sulfate from the bottom to the top. The columns were then eluted with 20 mL hexane, and the eluent was evaporated to about 0.3 mL and transferred to a 1.5 mL sample vial. After the internal standard (10 ng BDE-77) had been added to the vial, the volume of the solution was made up to 0.5 mL. Many similar previous studies used BDE-77 as internal standard or surrogate [34,35].

2.6. Sample Analysis

A 6890 GC/5975 MSD system (Agilent, Santa Clara, CA, USA) operated in negative chemical ion source/ selective ion monitoring (NCI/SIM) mode and equipped with a 15 m DB-XLB column (0.25 mm, 0.1 μm film thickness, J&W Sciebtific, Folsom, CA, USA) was used for PBDE separation and quantification. The samples (1 μL) were injected in split less mode. Helium was used as carrier gas at a flow rate of 1.2 mL/min, and the temperature program was set as follows: 90 °C for 2min, increased to 320 °C at 15 °C/min and held for 7 min. The temperature of GC inlet, transfer line, ionization source and quadrupole were set at 290 °C, 300 °C, and 150 °C. The compounds were monitored at m/z 79 and

81 for 3–7 brominated BDEs, m/z 79, 81, 487 and 489 for BDE-206, 207, 208 and 209, m/z 474, 476 for ^{13}C-PCB-208, and m/z 497.6, 499.6 for PCB-209.

Soil pH was measured using a calibrated pH meter (691, Metrohm AG, Herisau, Switzerland) in a weight: volume ratio of 1:10 of soil and tap water, adopting the USEPA method 9054D [36]. Total organic content (TOC) of the soil and dust was determined as the weight loss of dried soil (3 h at 100 °C) at 550 °C for 5 h [37]. Since PBDEs have a great potential to bind to environmental matrices rich in organic carbon [38], measuring the TOC concentration in the soil and dust samples was used to establish whether there was any correlation between the measured PBDEs and TOC.

2.7. Quality Assurance/Quality Control

A meadow soil collected from a cropland in Liaoning, China (123.90° E, 41.38° N) which was tested and demonstrated to be free of most of the studied PBDEs, was used as matrix blank and matrix spike samples. Twenty ng BDE 206, 207, 208, 209 and 4 ng of the other target PBDEs were spiked into 5 g meadow soil to evaluate the method performance. The recoveries for BDE-100, 154, 153, 183, 190, 208, 207, 206 and 209 were 60–107%. For BDE-28, 47, and 99, the recoveries were 32–58%. For all the target compounds, the relative standard deviations of duplicate samples were less than 14%. The recoveries of BDE-17, 66, 71 and 85 were lower than 20%, therefore they were not excluded from the statistically analysis. For the spiked surrogate ^{13}C-PCB-208 and PCB-209, the average recoveries in all samples were 71% and 84%, respectively. The method detection limits (MDLs) values characterized as three times the signal-to-noise ratio were 8–164 pg/g for the target PBDEs (Supplementary Tables). The procedural blanks and solvent blanks were analysed simultaneously with samples to check for interferences and contamination. The reported results of PBDEs in the samples were corrected by recoveries of ^{13}C-PCB-208. Three criteria were also used to ensure the correct identification of the target compounds [39]: (a) The GC (gas chromatography) retention times matched those of the authentic standards within ±0.1 min. (b) the signal-to-noise ratio was greater than 3:1; and (c) the isotopic ratios between the quantitative and confirmation ions were within ±15% of the theoretical values. Both ^{13}C-PCB-208 and PCB-209 were used as surrogate to indicate the stability of the recoveries of each sample. In addition, this measure helps to monitor the recoveries of the target compounds at different concentration level.

2.8. Data Analysis

A two-tailed Pearson correlation coefficient was used to determine the strength of association between PBDE congeners and Total Organic Carbon (TOC). In this contribution, we considered 13 PBDE congeners –BDE-28, BDE-47, BDE-100, BDE-99, BDE-153, BDE-154, BDE-138, BDE-183, BDE-190, BDE-207, BDE-208, BDE-206, BDE-209. The PBDE concentrations were summarized using descriptive statistics (frequencies, median means, and standard deviations). In addition, the pollution status of the different sites were evaluated by calculating the ratio of various PBDE concentrations at the e-waste sites compared to the control sites, which is known as exceedance (Ex) or contamination factor (CF). A CF < 1 indicates low contamination, 1 < CF < 3 moderate contamination, 3 < CF < 6 considerable contamination and CF > 6 indicates a very high contamination level [40,41].

To understand variabilities in the mean concentration distribution of PBDEs in the environmental samples, we evaluated the differences in the mean concentrations of the individual PBDE congeners and the sum of the PBDE congeners ($\sum_{13}$PBDE) by running a series of one-way (using activity site as factor on the log transformed data for soil, floor dust, roadside dust and direct dust separately. Additional series of two-way ANOVAs were run, firstly with activity sites (burning, dismantling, repair, and control sites) and location as explanatory variables; secondly with type of activity/activity site and type of sample (soil and dust) as explanatory variables. Bonferroni post-hoc tests were included to interpret the significant main effects of the ANOVA outputs. Principal component analysis (PCA) was used to evaluate whether variation in PBDE concentrations was similar across PBDE compounds.

A *p*-value of 0.05 was considered statistically significant. All statistical analyses were performed using SPSS version 23 (IBM Statistics20, IBM, Armonk, NY, USA).

3. Results

3.1. Physicochemical Characteristics of the Soil and Dust Samples

The soil texture in Lagos and Aba is sandy loam and in Ibadan is sandy clay [42]. The physiochemical characteristics of the soil and dust samples such as pH, total organic matter content (TOC), are presented on Supplementary Table S1a, b. The pH at the burning sites ranged from 7.2 to 8.4, dismantling sites (1.9 to 9), repair sites (7.9 to 9.24), and control sites (7.4 to 9) while the TOC for the burning sites ranged from 8 to 36.3%, dismantling sites (1.6 to 24%), repair sites (1.14 to 24%), and control sites (0.98 to 5.3%). There were significant differences in the pH ($p = 0.004$) and TOC ($p = 0.000$) levels between all the sites. Also, there were significant differences in the pH ($p = 0.006$) and TOC ($p = 0.00$) between the various e-waste recycling sites.

3.2. PBDE Concentrations at the Various Activity Sites

PBDE concentrations at e-waste sites and control sites are presented in Supplementary Tables S2–S4 and in Figure 3. PBDE congeners were detected in all sampling sites, indicating that PBDEs were widespread pollutants in this research area. The total concentration ($\sum_{13}$PBDE) ranged from 1.702 to 149,770.560 ng/g. The most abundant PBDE congener in all the sites and samples was BDE-209, with concentrations ranging from 0.850 to 147,091.400 ng/g. The maximum $\sum_{13}$PBDE was found in direct dust from TV repair shops in Ibadan. This high variability in PBDE concentrations is a reflection of the activities on the sites. The abundance of the PBDE congeners (considering the median of all the samples) in all the locations are generally in this order: BDE-209 > BDE-207 > BDE-206 > BDE-183 > BDE-208, BDE-99 > BDE-153 > BDE-47 > BDE-190 > BDE-154 > BDE-100 >BDE-28 > BDE-138 >. The PBDE concentrations at the e-waste exceeded the control sites by many folds, see Figure 3 and Supplementary Tables S2–S4 for details of the exceedance (EX) levels. The general pattern of the PBDEs distribution at the e-waste sites showed concentrations in this decreasing order: burning sites > dismantling sites > repair sites > control sites. This shows that burning activities contribute most to the PBDE concentrations in the environmental matrices, as is also clear from the patterns in $\sum_{13}$PBDE (Figure 3). In all the sample types, the highest concentrations were found in Lagos.

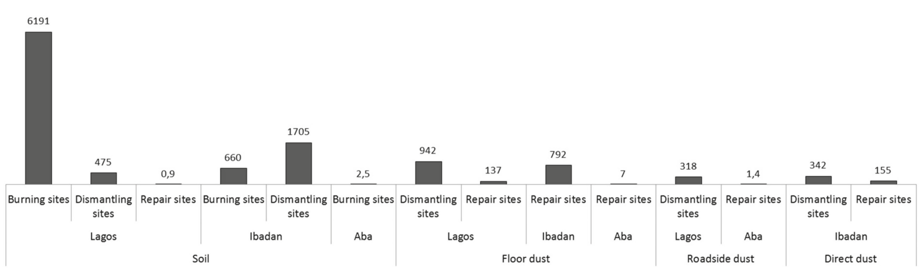


Figure 3. Exceedances of $\sum_{13}$PBDE concentrations in the samples from the e-waste recycling sitescompared to the control sites across locations.

3.3. Assessment of the Top Soil Samples

The one-way ANOVAs showed a significant difference in the concentration of $\sum_{13}$PBDE and all the PBDE congeners, except BDE-28, in top soils between the activity sites, ($p = 0.05$). Post-hoc comparisons indicated the main differences is between control and burning sites, and control and dismantling sites for all the PBDE congeners and $\sum_{13}$PBDE, see Figure 4 and Supplementary Table S5.

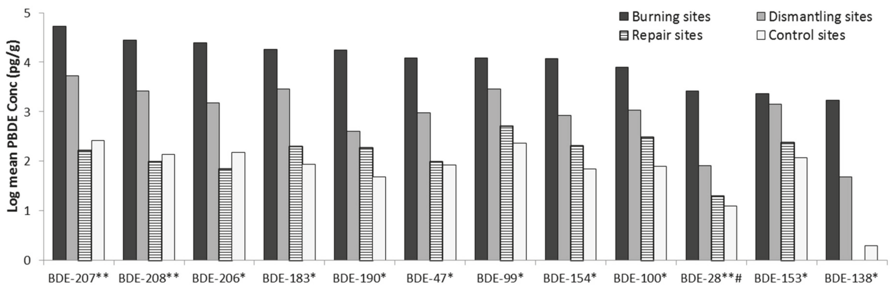


Figure 4. PBDE concentrations in top soils across the sites. The PBDE congeners influenced by activity at the sites are indicated with *, and an additional * for those influenced by location, and # for interaction between activity and location.

While the type of activities on the sites influences the PBDE concentrations on the sites, that effect might differ across locations. To test this assumption, a two-way ANOVA was performed and this ANOVA confirmed a significant difference in all the PBDE concentrations between the activity sites, with the burning sites having the highest concentrations, followed by dismantling sites, then control sites. Forty-one to 100% of the variability in PBDE concentrations was accounted for by the activities at the sites. However, there was no significant difference in the concentration of any of the PBDE congeners between locations, except for BDE-28, 208, and 207. The PBDE concentrations were generally highest in Aba, followed by Lagos, then Ibadan. This excludes BDE-28, which is highest in Lagos. Also, the interaction of activities at the sites and location showed no significant difference on any of the PBDE congeners and $\sum_{13}$PBDE concentrations, except for BDE-28 (see Supplementary Table S6).

3.3.1. Assessment of the Floor Dust Samples

A one-way ANOVA showed a significant difference ($p = 0.05$) in the concentration of some of the PBDE congeners (BDE-99, 47,100, 154, 190, 138) in floor dust between dismantling and repair sites (Figure 5).

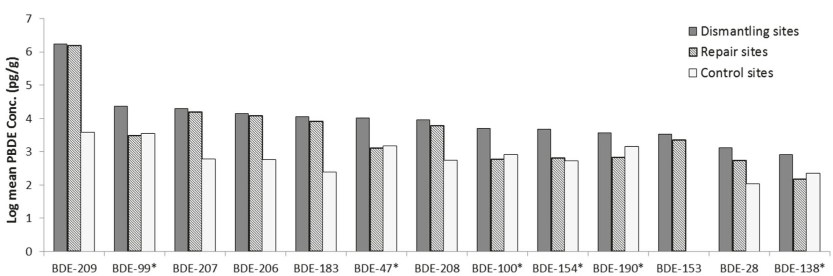


Figure 5. PBDEs concentration in floor dust across the sites. The PBDE congeners influenced by activity at the sites are indicated with *.

The two-way ANOVAs testing individual PBDE congeners and total PBDE ($\sum_{13}$PBDE) concentrations in floor dust from the activity sites (dismantling and repair sites) and across the locations (Lagos, Ibadan, and Aba), showed significant differences in the concentration of some PBDE congeners (except BDE-47, 100, 99, 154, 138, 190) between the activity sites, with dismantling sites having higher concentrations than repair sites. Also, there were significant differences in the concentration of some PBDE congeners (BDE-47, 100, 99) between locations, with Ibadan having generally the highest PBDE concentrations, followed by Lagos, then Aba. The interactions of activities at the sites and location showed no significant difference on any of the PBDE congeners and $\sum_{13}$PBDE concentrations, except for BDE-154 (see Supplementary Table S7).

3.3.2. Assessment of the Roadside Dust Samples

Roadside dust samples were collected only from Lagos and Aba. One-way ANOVA showed a significant difference ($p = 0.05$) in the concentration of almost all the PBDE congeners in roadside dust between dismantling and control sites (Figure 6).

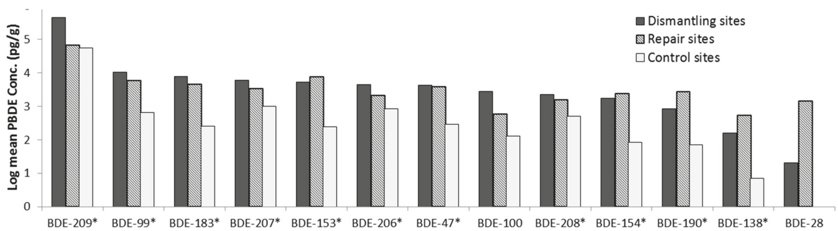


Figure 6. PBDE concentration in roadside dust across the sites. The PBDE congeners influenced by activity at the sites are indicated with *.

3.3.3. Assessment of the Direct Dust Samples

Direct dust samples from electronics were collected from dismantling and repair sites from Ibadan only. A one-way ANOVA showed no significant difference in the concentration of any of the PBDE congeners in direct dust between dismantling and repair sites, except BDE-100, (Figure 7).

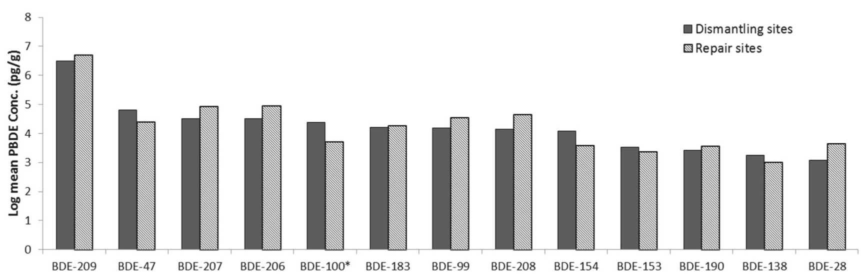


Figure 7. PBDE concentration in direct dust from electronics at the sites. Generally, there were no significant differences in the PBDE concentrations, except for BDE-100.

3.4. Patterns in PBDEs Contamination

The principal component analysis (PCA) using direct oblimin rotation was performed on the correlation matrix of the PBDE concentrations to establish whether the contaminants were actually arising from the same source or not. The analysis revealed one common axis of variation in PBDE concentrations, which accounted for 84% of the total variance. All PBDEs varied in the same direction (Figure 8). All the PBDE congeners had high positive loadings of ≥ 0.803. These findings indicate that PBDEs contamination has one common driver, which might suggest on common source. Pearson correlation confirmed that all PBDE congeners at the e-waste sites strongly correlated positively with each other. There was also correlation between the TOC and all the individual PBDEs and $\sum_{13}$PBDE (Supplementary Table S8).

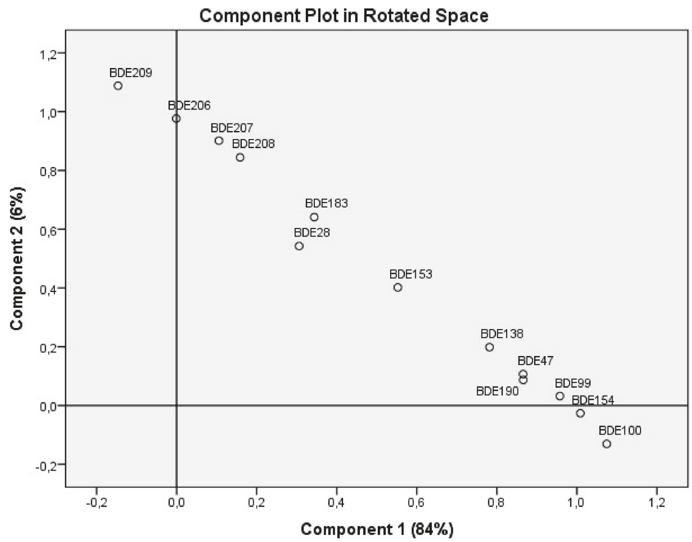


Figure 8. PCA plot of PBDE concentrations.

4. Discussion

In this study we analysed the PBDE concentrations in soil and dust samples from different e-waste activity sites (burning, dismantling, and repair sites) as compared to corresponding control sites in three different cities in Nigeria. The strengths of this study are the analyses of the interrelationships between the e-waste sites and the environmental matrices such as top soils, floor dusts, roadside dust, and direct dust from the electronics in which PBDEs are measured. These distinctions, we did not find in the previous studies.

Most e-waste recycling activities, especially at dismantling and burning sites are carried out outdoors. The recycling activities include storage, washing, cleaning, dismantling, and metal recovery through stripping of wires or open burning. The remains of *e*-waste materials from the recycling activities are dumped outside on the ground. Most repair activities, which involve soldering of various parts, take place indoors but also sometimes outdoors, depending on the settings of the work environment and the weather condition. These activities release large quantities of hazardous substances. Soils and dusts are a major repository for pollutants released into the environment by human activities, and they are important environmental media that can provide information about the level, distribution, and fate of contaminants present in the environment as a result of informal *e*-waste recycling.

Despite the increasing volumes of e-waste generated over the years, collection and recycling of e-waste are still not improved in developing countries [27]. Nigeria imports the largest volume of new and used electronic and electrical equipment in Africa [25]. The amount of e-waste generated in Nigeria has increased from 219 kilotonnes in 2014 [1] to 277 kilotonnes in 2016 [2]. This increase is despite the high weight reduction of electronic devices like computers (PCs). Almost all the e-waste generated are recycled in an unsafe/informal manner [1,25]. This situation in Nigeria is likely to be representative for informal e-waste recycling in countries that lack the resources for safe e-waste recycling such as in India, Brazil, Mexico [43], and Ghana [44] among others.

4.1. Extent of Pollution As A Result of Informal E-waste Recycling

Our findings revealed that open burning of e-waste is the most polluting e-waste recycling activity. This is in accordance with a study by Matsukami et al., [45] which compared burning sites and other e-waste processing sites in Vietnam. PBDEs do not occur naturally in the environment, but traces of PBDEs were found in control sites, indicating deposits of PBDEs in the environment not too far from e-waste recycling sites. These PBDEs might have been transported by wind/air to nearby vicinities, which is in agreement with the observations of decreasing concentrations of PBDEs with increasing distance from e-waste sites [46,47].

In this study we included dust samples, considering that some of the activities (such as repair activities) do not take place on soil (bare ground) most of the time. The added value of determining the PBDE concentrations in the dust samples was to ensure that different types of e-waste activity sites were studied. Dust is one of the main sources of exposure to PBDEs via inhalation or ingestion. Thus it provides information about the level of contaminants in the indoor atmosphere as well as the levels of contaminants to which the workers and the public are exposed. Dust also reflects the characteristics of short and long term activities in the area. Furthermore, the combination of soil and dust samples gives a comprehensive overview of the impact of informal e-waste recycling on the environment. It is likely that there is a cross transference of the PBDEs from the floor dust (indoor) from the shops to the soil, and from the soil from the burning and dismantling sites into the shops (contaminating floor dust). There is also a probable transfer of PBDEs from the e-waste sites to locations farther away from the e-waste recycling sites, which is in agreement with the previous findings [48,49].

We found high concentrations of PBDE congeners at the e-waste sites with the higher molecular weight PBDEs (BDE-209, BDE-153, BDE-183) having the highest concentrations, and BDE-209 being the most abundant. Predominance of BDE-209 in samples is probably due to the fact that the deca-BDE mixture is the predominant PBDE still in use [20,50]. This is in agreement with the results of previous studies in Turkey [51], in five Asian countries [52], in Vietnam by Matsukami et al., (2017) [45], in Ghana by Akortia et al., (2017) [44], and in Nigeria and China [53]. This is similar to the findings of Takigami et al. [54] showing the highest concentrations of PBDEs (BDE-209) in dust from e-waste sites. The maximum concentration of BDE-209 was 147,091 ng/g, as found in dust from a television. However, in a similar study in Ghana, BDE-28 was found to have the highest concentration instead of BDE-209 [44].

The PBDE levels present at the e-waste sites and the control sites reflects the pollution from anthropogenic sources in urban areas. Comparing the mean concentrations of BDE-209 levels in top soils to those reported in previous studies at the same study areas, it is found that BDE-209 concentrations are decreasing slowly in a space of three years by a factor of 1 and 3 at Alaba and Computer village respectively (Table 1).

Table 1. Comparison of PBDE concentrations in soils at e-waste sites with other studies.

Countries	Units	BDE-209	Reference
Computer village Nigeria	ng/g dry wt	583	This study
Alaba international market, Nigeria	ng/g dry wt	7648 ± 8369	
Soil at e-waste recycling site, Ghana	ng/g dry wt	10.6 ± 16.6	[44]
Soils near e-waste recycling site, China	ng/g dry wt	3400 ± 4200	[50]
Soils near e-waste recycling site South Korea	ng/g dry wt	8.8 ± 11	
Soils near e-waste recycling site Vietnam	ng/g dry wt	63	
Guiyu Soils near e-waste recycling site, China	ng/g dry wt	1157 ± 1131	[55]
Computer village dumpsite soil, Nigeria	ng/g dry wt	1820	[51]
Alaba international market dumpsite soil, Nigeria	ng/g dry wt	9800	
Guiyu e-waste dumpsite soil, China	ng/g dry wt	12,130	

It is important to note that the e-waste dismantling and burning sites at Alaba in 2012 is different from the site in 2015 because the site was relocated within Alaba. We suspect that if the same site was retained in Alaba since inception of *e*-waste burning at Alaba, the levels of the PBDE congeners maybe higher than the levels detected. At the same time, we gave it a thought that the decrease may also be because PBDEs use in electronics has been banned. The median BDE-209 concentration at the burning sites is as high as 17,587 ng/g at Alaba, Lagos (Supplementary Tables S2–S4). When compared to Guiyu, China, the BDE-209 levels at Guiyu decreased by a factor of 10 in a space of two years (Table 1).

PBDE levels found at Alaba sites are higher than the levels found in the widely studied area of Guiyu, China in 2014 and 2015. Guiyu, China is known for its notorious intensive unregulated crude e-waste recycling activities. This shows that maybe more notorious e-waste recycling activity maybe going on somewhere in Nigeria, which are yet to be reported. These findings further show that the PBDE concentrations in urban cities (mostly in the slums) in Nigeria are still high and call for concern. This consequently implies that more people in the general population (besides e-waste workers) might be exposed to PBDE. This is more disturbing as majority of the e-waste workers are unaware of the health risks associated with their jobs [48] and do not use any form of PPE [49] or take appropriate cause to protect their health or the environment.

In this study we distinguished between PBDE levels at various e-waste recycling sites (burning, dismantling, and repair sites). These distinctions were not made in the other studies. The different e-waste activities had significantly different mean concentrations of PBDEs. The post-hoc tests revealed that the biggest effect is seen at the burning and control sites being significantly different from the other sites. There was no significant difference in the PBDE concentrations of similar activity sites between the locations, and between sample types for most of the PBDEs. These findings indicate that activities in the vicinity have impact on the level of PBDEs in an area. All PBDE congeners positively correlated with each other. The positive correlations between the PBDE congeners indicate that the PBDEs are likely from the same source with similar emission patterns. This was further confirmed by the PCA, in which 84% of the total variance of all PBDE congeners was accounted for by a common axis.

Although penta-BDE, octa-BDE, and deca-BBE are banned in developed countries where electronics are manufactured, and despite Nigeria having regulations on e-waste management which in turn controls PBDE emissions, PBDEs were found in high concentrations at the e-waste sites and at the control sites, this shows that PBDEs are ubiquitos in Nigerian environment, as stated in other studies which detected PBDE in various environmental matrices [55–58]. There is a possiblity that higher molecular weight BDEs debrominates to lower molecular weight BDEs as stated by Zhang et al. [59]. We found deca-BDE (BDE-209), a high molecular weight BDEs having the highest concentrations at all sites. Therefore deca-BDE may represent important reservoir lower-PBDE congeners. Generally, there were positive correlations between all the PBDE congeners. There were also positive correlations between all the PBDEs and the TOC in top soils from e-waste (Supplementary Table S8). Correlation of TOC with PBDE suggests that PBDE binds to environmental matrix rich in organic carbon. As PBDEs bind strongly to soil particles, they may remain in soil for several years or even decades. Total Organic matter content (TOC) influences the distribution of PBDE in the soil and dust to some extent.

4.2. Implications of High PBDE Concentrations on Health and Environment

As PBDEs do not naturally occur in the environment, there is no doubt that e-waste recycling is a major source of PBDE pollution in Nigeria. When released, they bind strongly (especially congeners with higher content of bromine bind more strongly) to soil, sediment particles, and sewage sludge, in turn making them less mobile in the environment. Therefore, they bioaccumulate and biomagnify in aquatic organisms, fish, and plants, and are eventually transferred up the food chain, ultimately to humans [57]. Moderate to high PBDE congeners are found in air samples closer to the source of pollution, while PBDE congeners with less bromine atoms travel greater distances from their same source [58,59], meaning that people living far away from the source of release may also be at risk of exposure to PBDE. It is assumed that the higher PBDEs may degrade to lower PBDE congeners like tetra-, penta-, and hexa-BDEs in the environment, and that the PBDEs with lesser bromine atoms are more persistent in the atmosphere [60].

The $\sum_{13}$PBDE concentrations found in soils and various dust samples exceeded the Agency for Toxic Substances and Disease Registry (ATSDR) oral Minimal Risk Levels (MRL) of 0.00006 mg/kg/day for lower-brominated PBDEs based on a LOAEL (lowest-observed-adverse-effect level) for endocrine effects in rats. Our values also exceeded The EPA's reference doses (RfDs) for penta, octa-, and decaBDEs are 2×10^{-3}, 3×10^{-3}, and 7×10^{-3} mg/kg/day, respectively [61], suggesting that PBDEs could adversely affect animals and other sensitive species in the environment, and consequently humans in and around the study areas. Humans can be exposed to PBDEs and metals though inhalation, dermal absorption, and consumption of contaminated foods such as fish, meat, and dairy products [62,63]. This is a considerable environmental concern and most likely a health concern. We recommend further toxicological studies on the e-waste workers. It is hoped that the results of this study are a wake-up call on the need for more effective strategies on enforcement of e-waste regulations in Nigeria. We recommend that the enforcement would be effective if the regulations are made through the lens of the informal sector and enforcement agencies collaborates with the informal sector so as not to impade the workers' livelihood. The findings in this study is representative of what might be going on in other places unsafe e-waste recycling is practice, therefore solutions proferred for Nigeria is applicable to other places.

5. Conclusions

Our study showed that PBDE concentrations at the e-waste recycling sites were elevated compared to those detected at the control sites by 100 s to 1000 s times, with BDE-209 being the most abundant in all the samples and at all the sites. There was a significant difference in concentrations of PBDEs at the various e-waste activities sites in this decreasing order: burning sites > dismantling sites > repair sites > control sites. This proves that the type of activities at the sites influences the level of PBDEs, with burning activities having the most effect. This study demonstrates that crude recycling of e-waste contributes significantly to emissions of organic pollutants in the environment. Comparing our results with past studies in the same locations, not much has changed in the the PBDE concentrations considering that PBDE use in electronics has been banned, suggesting that the situation calls for urgent action. Our results suggest that the informal e-waste recycling has negative impacts on the enviroment and consequently on health. There is an urgent need for more effective actions to stop open burning of e-waste and to reverse or stop the environmental deterioratation as a result of informal e-waste recycling. One way to stop unsafe recycling of e-waste is to adopt a bottom-up approach in stopping these unsafe practices by: (1) for the formal institutions to appreciate and work with the informal sectors (2) to the create awearness on the potential health risks of unsafe recycling of e-waste among th e-waste workers.

Supplementary Materials: The following are available online at http://www.mdpi.com/1660-4601/16/3/360/s1, Table S1a: Summary of the Physicochemical parameters (mean of pH and Total Organic matter content) of the samples at the study locations. No standard deviation is given if n=1., Table S1b: Physicochemical parameters (pH and Total Organic Matter Content) of the samples at the study site., Table S2: PBDE median concentrations (ng/g dry weight) across various e-waste sites; and exceedance (Ex) of PBDE concentrations in soil and dusts across various e-waste sites compared to control sites in Lagos., Table S3: PBDE median concentrations (ng/g dry weight) across various e-waste sites; and exceedance(Ex) of PBDE concentrations in soil and dusts across various e-waste sites compared to control sites in Ibadan., Table S4: PBDE median concentrations (ng/g dry weight) across various e-waste sites; and exceedance (Ex) of PBDE concentrations in soil and dusts across various e-waste sites compared to control sites in Aba., Table S5: F-values of the two-Way ANOVA for PBDE concentration in all sample types across the Activity sites (burning. dismantling. repair sites)., Table S6: F-values of the two-Way ANOVA for PBDE concentration in top soils across the locations(Lagos. Ibadan. Aba) and Activity sites (burning. dismantling. repair. and control sites)., Table S7: F-values of the two-Way ANOVA for PBDE concentration in floor dust across the locations(Lagos. Ibadan. Aba) and Activity sites (dismantling and repair sites)., Table S8: Correlations between PBDEs congeners. ∑13PBDE. pH and TOC (Top soils).

Author Contributions: This study is part of the Ph.D. research of C.M.O. C.M.O. conceived the study, carried out the literature review, developed data collection and analyses plan, collected data, analysed the data, drafted the manuscript and provided critical revision to the manuscript. W.J.G.M.P., P.M.V.B., M.V., and O.O. are the supervisors and were involved in study design, data collection plan, revision of the various versions of the manuscript, and made contributions to the manuscript. Q.X. and J.C. carried out the laboratory analysis. P.M.V.B.

was involved in all stages of the write-up and critical revision of the manuscript, and contributed to the data analyses and advanced statistical procedures.

Funding: This research was funded by NUFFIC (Netherlands Universities Foundation for International Cooperation), research grant CF9420/2014 NFP-PhD.14/37.

Acknowledgments: Data for this paper was obtained within a research project supported by the Netherlands Fellowship Program of NUFFIC (Netherlands Universities Foundation for International Cooperation), research grant CF9420/2014 NFP-PhD.14/37. The authors acknowledge the various associations of second-hand electronics dealers, scrap dealer associations in each location for their cooperation and for allowing the soil and dust samples to be collected from the work sites, research assistants for helping out in sample collection, and laboratory assistants for helping out with the data analysis.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Baldé, C.P.; Wang, F.; Kuehr, R.; Huisman, J. *The Global E-Waste Monitor—2014*; United Nations University: Bonn, Germany, 2015; ISBN 978-92-808-4556-3.
2. Baldé, C.P.; Gray, V.; Forti, V.; Kuehr, R.; Stegmann, P. *The Global E-waste Monitor 2017*; United Nations University (UNU): Bonn, Germany; International Telecommunication Union (ITU): Geneva, Switzerland; International Solid Waste Association (ISWA): Vienna, Austria, 2017; ISBN 9789280890532.
3. Cruz, N.F.; Ferreira, S.; Cabral, M.; Simoes, P.; Marques, R.C. Packaging waste recycling in Europe: Is the industry paying for it? *Waste Manag.* **2014**, *34*, 298–308. [CrossRef] [PubMed]
4. United Nations Environmental Programme, Division of Technology I and E, International Environmental Technology. *E-Waste Volume I: Inventory Assessment Manual*; United Nations Environmental Programme: Osaka, Japan, 2007.
5. Widmer, R.; Oswald-Krapf, H.; Sinha-Khetriwal, D.; Schnellmann, M.; Böni, H. Global perspectives on e-waste. *Environ. Impact Assess. Rev.* **2005**, *25*, 436–458. [CrossRef]
6. Brigden, K.; Labunska, I.; Santillo, D.; Johnston, P. *Chemical Contamination at E-Waste Recycling and Disposal Sites in Accra and Korforidua, Ghana*; The Greenpeace Research Laboratories: Exeter, UK, 2008.
7. Asante, K.A.; Agusa, T.; Biney, C.A.; Agyekum, W.A.; Bello, M.; Otsuka, M.; Itai, T.; Takahashi, S.; Tanabe, S. Multi-trace element levels and arsenic speciation in urine of e-waste recycling workers from Agbogbloshie, Accra in Ghana. *Sci. Total Environ.* **2012**, *424*, 63–73. [CrossRef] [PubMed]
8. National Center for Environmental Assessment, Office of Research and Development, United States Environmental Protection Agency (USEPA). *An Exposure Assessment of Polybrominated Diphenyl Ethers*; USEPA: Washington DC, USA, 2010.
9. Kostianoy, A.G.; Boer, De J.; Garrigues, P.; Gu, J.; Jones, K.C.; Knepper, T.; Newton, A.; Sparks, D.L. *The Handbook of Environmental Chemistry (Book 16)*; Damia Barcelo, A.G.K., Ed.; Springer-Verlag: Berlin, German, 2010; ISBN 9783642192685.
10. Agency for Toxic Substances and Disease Registry (ATSDR); Division of Toxicology/Toxicology Information Branch, United States Environmental Protection. *Toxicological Profile for Polybrominated Biphenyls and Polybrominated Diphenyl Ethers*; Agency for Toxic Substances and Disease Registry (ATSDR); Division of Toxicology/Toxicology Information Branch, United States Environmental Protection: Atlanta, GA, USA, 2004.
11. Zheng, X.; Xu, X.; Yekeen, T.A.; Zhang, Y.; Chen, A.; Kim, S.S.; Dietrich, K.N.; Ho, S.M.; Lee, S.A.; Reponen, T.; et al. Ambient air heavy metals in $PM_{2.5}$ and potential human health risk assessment in an informal electronic-waste recycling site of China. *Aerosol. Air Qual. Res.* **2016**, *16*, 388–397. [CrossRef]
12. Berg van den, M.; Denison, M.S.; Birnbaum, L.S.; Devito, M.J.; Fiedler, H.; Falandysz, J.; Rose, M.; Schrenk, D.; Safe, S.; Tohyama, C.; et al. Polybrominated dibenzo-p-dioxins, dibenzofurans, and biphenyls: Inclusion in the toxicity equivalency factor concept for dioxin-like compounds. *Toxicol. Sci.* **2013**, *133*, 197–208. [CrossRef] [PubMed]
13. Koibuchi, N.; Yen, P.M. (Eds.) *Thyroid Hormone Disruption and Neurodevelopment*; Springer: New York, NY, USA, 2016; ISBN 978-1-4939-3735-6.
14. Darnerud, P.O.; Eriksen, G.S.; Johannesson, T.; Larsen, P.B.; Viluksela, M. Polybrominated Diphenyl Ethers: Occurrence, Dietary Exposure, and Toxicology. *Environ. Health Perspect.* **2001**, *109*, 49. [PubMed]

15. Betts, K.S. Unwelcome guest: PBDEs in indoor dust. *Environ. Health Perspect.* **2008**, *116*, 202–208. [CrossRef]
16. La Guardia, M.J.; Hale, R.C.; Harvey, E. Detailed polybrominated diphenyl ether (PBDE) congener composition of the widely used penta-, octa-, and deca-PBDE technical flame-retardant mixtures. *Environ. Sci. Technol.* **2006**, *40*, 6247–6254. [CrossRef]
17. European Union. The European Commission, 2017/227, Amending Annex XVII to Regulation (EC) No 1907/2006 of the European Parliament and of the Council concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) as regards bis (pentabromophenyl) ether. 9 February 2017. *Off. J. Eur. Union* **2017**.
18. United Nations Enviornmental Programme (UNEP). Recommendations of the Persistent Organic Pollutants Review Committee of the Stockholm Convention to amend Annexes A, B or C of the Convention. UNEP/POPS/COP.4/17. In Proceedings of the Stockholm Convention on Persistent Organic Pollutants, Geneva, Switzerland, 4–8 May 2009.
19. United States Environmental Protection Agency (USEPA). *Design for the Environment Furniture Flame Retardancy Partnership: Environmental Profiles of Chemical Flame-Retardant Alternatives for Low-Density Polyurethane Foam*; USEPA: Cincinnati, OH, USA, 2005; Volume 1.
20. Pure Strategies and The Lowell Center for Sustainable Production University of Massachusetts Lowell. *Decabromodiphenylether: An Investigation of Non-Halogen Substitutes in Electronic Enclosure and Textile Applications*; Lowell Center for Sustainable Production, University of Massachusetts: Lowell MA, USA, 2005.
21. Pakalin, S.; Cole, T.; Steinkellner, J.; Nicolas, R.; Tissier, C.; Eisenreich, S. *Review on Production Processes of Decabromodiphenyl Ether (DecaBDE) Used in Polymeric Applications in Electrical and Electronic Equipment, and Assessment of the Availability of Potential Alternatives to Deca BDE*; European Chemicals Bureau Institute of Health and Consumer Protection, Directorate-General, Joint Research Centre ISPRA, European Commission: Italy, Rome, 2007.
22. Castro-Jiménez, J.; González-Gaya, B.; Pizarro, M.; Casal, P.; Pizarro-Álvarez, C.; Dachs, J. Organophosphate ester flame retardants and plasticizers in the global oceanic atmosphere. *Environ. Sci. Technol.* **2016**, *50*, 12831–12839. [CrossRef]
23. Awasthi, A.K.; Li, J. Management of electrical and electronic waste: A comparative evaluation of China and India. *Renew. Sustain. Energy Rev.* **2017**, *76*, 434–447. [CrossRef]
24. Prakash, S.; Manhart, A. *Socio-Economic Assessment and Feasibility Study on Sustainable E-Waste Management in Ghana*; Öko-Institut e.V.: Freiburg, Germany, 2010.
25. Ogungbuyi, O.; Nnorom, I.C.; Osibanjo, O.; Schluep, M. *E-Waste Country Assessment Nigeria*; E-Waste Africa project of the Secretariat of the Basel Convention; Basel Convention Coordinating Centre, Nigeria, UNEP; EMPA: Dübendorf, Switzerland, 2012.
26. Leung, A.O.W.; Duzgoren-Aydin, N.S.; Cheung, K.C.; Wong, M.H. Heavy Metals Concentrations of Surface Dust from e-Waste Recycling and Its Human Health Implications in Southeast China. *Environ. Sci. Technol.* **2008**, *42*, 2674–2680. [CrossRef]
27. Ackah, M. Informal E-waste recycling in developing countries: Review of metal(loid)s pollution, environmental impacts and transport pathways. *Environ. Sci. Pollut. Res.* **2017**, *24*, 24092–24101. [CrossRef] [PubMed]
28. Banerjee, A.D.K. Heavy metal levels and solid phase speciation in street dusts of Delhi, India. *Environ. Pollut.* **2003**, *123*, 95–105. [CrossRef]
29. Wuana, R.A.; Okieimen, F.E. Heavy Metals in Contaminated Soils: A Review of Sources, Chemistry, Risks and Best Available Strategies for Remediation. *ISRN Ecol.* **2011**, *2011*, 402647. [CrossRef]
30. Tang, Z.; Huang, Q.; Yang, Y.; Nie, Z.; Cheng, J.; Yang, J.; Wang, Y.; Chai, M. Polybrominated diphenyl ethers (PBDEs) and heavy metals in road dusts from a plastic waste recycling area in north China: Implications for human health. *Environ. Sci. Pollut. Res.* **2015**, *23*, 625–637. [CrossRef] [PubMed]
31. Lu, X.; Zhang, X.; Li, L.Y.; Chen, H. Assessment of metals pollution and health risk in dust from nursery schools in Xi'an, China. *Environ. Res.* **2014**, *128*, 27–34. [CrossRef] [PubMed]
32. Ohajinwa, C.M.; van Bodegom, P.M.; Vijver, M.G.; Peijnenburg, W.J.G.M. Impact of informal electronic waste recycling on metal concentrations in soils and dusts. *Environ. Res.* **2018**, *164*, 385–394. [CrossRef] [PubMed]
33. Osibanjo, O.; Nnorom, I.C. The challenge of electronic waste (e-waste) management in developing countries. *Waste Manag. Res.* **2007**, *25*, 489–501. [CrossRef]

34. Cai, M.G.; Hong, Q.Q.; Wang, Y.; Luo, X.J.; Chen, S.J.; Cai, M.H.; Qiu, C.R.; Huang, S.Y.; Mai, B.X. Deep-Sea Research II Distribution of polybrominated diphenyl ethers and decabromodiphenylethane in surface sediments from the Bering Sea, Chukchi Sea, and Canada Basin. *Deep Res. Part II* **2012**, *81–84*, 95–101. [CrossRef]
35. Kalachova, K.; Hradkova, P.; Lankova, D.; Hajslova, J.; Pulkrabova, J. Science of the Total Environment Occurrence of brominated fl ame retardants in household and car dust from the Czech Republic. *Sci. Total Environ.* **2012**, *441*, 182–193. [CrossRef] [PubMed]
36. United States Environmental Protection Agency (USEPA). *USEPA Method 9045D: Soil and Waste pH, Part of Test. Methods for Evaluating Solid Waste, Physical/Chemical Methods*; USEPA: Washington, DC, USA, 2004.
37. Pansu, M.; Gautheyrou, J. *Handbook of Soil Analysis, Mineralogical, Organic and Inorganic Methods*; Springer: Berlin, Germany, 2006; ISBN 13.
38. Maya Li Wei-Haas. The Influence of Dissolved Organic Matter on the Fate of Polybrominated Diphenyl Ethers (PBDEs) in the Environment. Ph.D. Thesis, The Ohio State University, Columbus, OH, USA, 2015.
39. Zhu, L.; Hites, R.A. Brominated flame retardants in tree bark from North America. *Environ. Sci. Technol.* **2006**, *40*, 3711–3716. [CrossRef] [PubMed]
40. Hakanson, L. An ecological risk index for aquatic pollution control. a sedimentological approach. *Water Res.* **1980**, *14*, 975–1001. [CrossRef]
41. Islam, S.; Ahmed, K.; Masunaga, S. Potential ecological risk of hazardous elements in different land-use urban soils of Bangladesh. *Sci. Total Environ.* **2015**, *512–513*, 94–102. [CrossRef] [PubMed]
42. Nkwunonwo, U.C.; Okeke, F.I. GIS-based production of digital soil map for Nigeria. *J. Ethiop. Stud. Environ. Vol. Manag.* **2013**, *6*, 498–506. [CrossRef]
43. Lundgren, K. *The Global Impact of E-Waste: Addressing the Challenge*; International Labour Organization: Geneva, Switzerland, 2012; ISBN 9789221268970.
44. Akortia, E.; Olukunle, O.I.; Daso, A.P.; Okonkwo, J.O. Soil concentrations of polybrominated diphenyl ethers and trace metals from an electronic waste dump site in the Greater Accra Region, Ghana: Implications for human exposure. *Ecotoxicol. Environ. Saf.* **2017**, *137*, 247–255. [CrossRef]
45. Matsukami, H.; Suzuki, G.; Someya, M.; Uchida, N.; Tue, N.M.; Tuyen, L.H.; Viet, P.H.; Takahashi, S.; Tanabe, S.; Takigami, H. Concentrations of polybrominated diphenyl ethers and alternative flame retardants in surface soils and river sediments from an electronic waste-processing area in northern Vietnam, 2012–2014. *Chemosphere* **2017**, *167*, 291–299. [CrossRef]
46. Luo, Q.; Wong, M.H.; Wang, Z.; Cai, Z. Polybrominated diphenyl ethers in combusted residues and soils from an open burning site of electronic wastes. *Environ. Earth Sci.* **2012**, *69*, 2633–2641. [CrossRef]
47. Li, Y.; Duan, Y.P.; Huang, F.; Yang, J.; Xiang, N.; Meng, X.Z.; Chen, L. Polybrominated diphenyl ethers in e-waste: Level and transfer in a typical e-waste recycling site in Shanghai, Eastern China. *Waste Manag.* **2014**, *34*, 1059–1065. [CrossRef]
48. Ohajinwa, C.; Van Bodegom, P.; Vijver, M.; Peijnenburg, W. Health Risks Awareness of Electronic Waste Workers in the Informal Sector in Nigeria. *Int. J. Environ. Res. Public Health* **2017**, *14*, 911. [CrossRef]
49. Ohajinwa, C.M.; van Bodegom, P.M.; Vijver, M.G.; Olumide, A.O.; Osibanjo, O.; Peijnenburg, W.J.G.M. Prevalence and injury patterns among electronic waste workers in the informal sector in Nigeria. *Inj. Prev.* **2017**, *24*, 185–192. [CrossRef] [PubMed]
50. Morf, L.S.; Tremp, J.; Gloor, R.; Huber, Y.; Stengele, M.; Zennegg, M. Brominated Flame Retardants in Waste Electrical and Electronic Equipment: Substance Flows in a Recycling Plant. *Environ. Sci. Technol.* **2005**, *39*, 8691–8699. [CrossRef] [PubMed]
51. Cetin, B. Soil concentrations and source apportionment of polybrominated diphenyl ethers (PBDEs) and trace elements around a heavily industrialized area in Kocaeli, Turkey. *Environ. Sci. Pollut. Res.* **2014**, *21*, 8284–8293. [CrossRef] [PubMed]
52. Li, W.-L.; Ma, W.-L.; Jia, H.-L.; Hong, W.-J.; Moon, H.-B.; Nakata, H.; Minh, N.H.; Sinha, R.K.; Chi, K.H.; Kannan, K.; et al. Polybrominated Diphenyl Ethers (PBDEs) in Surface Soils across Five Asian Countries: Levels, Spatial Distribution, and Source Contribution. *Environ. Sci. Technol.* **2016**, *50*, 12779–12788. [CrossRef]
53. Alabi, O.A.; Bakare, A.A.; Xu, X.; Li, B.; Zhang, Y.; Huo, X. Comparative evaluation of environmental contamination and DNA damage induced by electronic-waste in Nigeria and China. *Sci. Total Environ.* **2012**, *423*, 62–72. [CrossRef] [PubMed]

54. Takigami, H.; Suzuki, G.; Hirai, Y.; Sakai, S. Transfer of brominated flame retardants from components into dust inside television cabinets. *Chemosphere* **2008**, *73*, 161–169. [CrossRef] [PubMed]
55. Katima, Z.; Okonkwo, J.O.; Daso, A.P. A review of brominated flame retardants in the environment with emphasis on atmospheric levels, knowledge and information gaps in the African continent. *Atmos. Pollut. Res.* **2017**, *8*, 767–780. [CrossRef]
56. Talsness, C.E. Overview of toxicological aspects of polybrominated diphenyl ethers: A flame-retardant additive in several consumer products. *Environ. Res.* **2008**, *108*, 158–167. [CrossRef]
57. Bocio, A.; Llobet, J.M.; Domingo, J.L.; Corbella, J.; Teixidó, A.; Casas, C. Polybrominated diphenyl ethers (PBDEs) in foodstuffs: Human exposure through the diet. *J. Agric. Food Chem.* **2003**, *51*, 3191–3195. [CrossRef]
58. Strandberg, B.; Dodder, N.G.; Basu, I.; Hites, R.A. Concentrations and spatial variations of polybrominated diphenyl ethers and other organohalogen compounds in Great Lakes air. *Environ. Sci. Technol.* **2001**, *35*, 1078–1083. [CrossRef]
59. Zhang, S.; Xu, X.; Wu, Y.; Ge, J.; Li, W.; Huo, X. Polybrominated diphenyl ethers in residential and agricultural soils from an electronic waste polluted region in South China: Distribution, compositional profile, and sources. *Chemosphere* **2014**, *102*, 55–60. [CrossRef] [PubMed]
60. O'Driscoll, K.; Robinson, J.; Chiang, W.S.; Chen, Y.Y.; Kao, R.C.; Doherty, R. The environmental fate of polybrominated diphenyl ethers (PBDEs) in western Taiwan and coastal waters: Evaluation with a fugacity-based model. *Environ. Sci. Pollut. Res.* **2016**, *23*, 13222–13234. [CrossRef] [PubMed]
61. Pohl, H.R.; McClure, P.; Zaccaria, K.; Llados, F.; Kawa, M.; Citra, M. *Toxicological Profile for Polybrominated Biphenyl and Polybrominated Diphenyl Ethers*; Agency for Toxic Substances and Disease RegistryDivision of Toxicology and Human Health Sciences: Atlanta, GE, USA, 2017.
62. Siddiqi, M.A.; Laessig, R.H.; Reed, K.D. Polybrominated diphenyl ethers (PBDEs): New pollutants-old diseases. *Clin. Med. Res.* **2003**, *1*, 281–290. [CrossRef] [PubMed]
63. Fujimori, T.; Takigami, H.; Agusa, T.; Eguchi, A.; Bekki, K.; Yoshida, A.; Terazono, A.; Ballesteros, F.C. Impact of metals in surface matrices from formal and informal electronic-waste recycling around Metro Manila, the Philippines, and intra-Asian comparison. *J. Hazard. Mater.* **2012**, *221–222*, 139–146. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

The Influence of Physical Fitness on Reasons for Academy Separation in Law Enforcement Recruits

Robert G. Lockie [1,*], Katherine Balfany [1], Ashley M. Bloodgood [1], Matthew R. Moreno [1], Karly A. Cesario [1], Joseph M. Dulla [2], J. Jay Dawes [3] and Robin M. Orr [4]

1 Department of Kinesiology, California State University, Fullerton, Fullerton, CA 92831, USA; kbalfany@csu.fullerton.edu (K.B.); abloodgood17@csu.fullerton.edu (A.M.B.); moreno.matthewr@csu.fullerton.edu (M.R.M.); cesariokarly@csu.fullerton.edu (K.A.C.)
2 Recruit Training Unit, Training Bureau, Los Angeles County Sheriff's Department, Los Angeles, CA 90022, USA; JMDulla@lasd.org
3 Department of Health Sciences, University of Colorado-Colorado Springs, Colorado Springs, CO 80918, USA; jdawes@uccs.edu
4 Tactical Research Unit, Bond University, Robina 4229, QLD, Australia; rorr@bond.edu.au
* Correspondence: rlockie@fullerton.edu; Tel.: +1-657-278-4971

Received: 17 December 2018; Accepted: 27 January 2019; Published: 29 January 2019

Abstract: This study analyzed the effects physical fitness may have on reasons for academy separation in law enforcement recruits. A retrospective analysis was conducted on 401 recruits; 330 recruits graduated (GRAD), and 71 recruits separated at various times during academy. Twenty-eight recruits separated for personal reasons (SEPPR); 18 due to physical training failures (i.e., poor fitness) or injury (SEPFI); and 25 due to academic or scenario failures (SEPAS). Fitness testing occurred prior to academy, and included: Push-ups and sit-ups in 60s; a 75-yard pursuit run (75PR); vertical jump; medicine ball throw; and multistage fitness test (MSFT). A one-way ANOVA with Bonferroni post hoc compared between-group fitness test performance. A multiple stepwise regression calculated whether recruit characteristics or fitness could predict separation. The GRAD group was younger than the SEPAS group ($p < 0.01$), faster in the 75PR than the SEPFI group ($p = 0.02$), and completed more MSFT shuttles than the SEPPR and SEPFI groups ($p = 0.01$). Age predicted GRAD and SEPAS group inclusion; MSFT predicted GRAD, SEPPR, and SEPFI group inclusion. Recruits who had superior high-intensity running capacity (75PR) and aerobic fitness (MSFT) should have a better chance of completing academy. However, this could be influenced by training practices adopted during academy.

Keywords: aerobic capacity; attrition; change-of-direction speed; deputy sheriff; graduation; high-intensity running; police; strength endurance; tactical

1. Introduction

Law enforcement can be a demanding profession that can place high levels of physical [1] and psychological [2,3] stress on those employed in this vocation. The academy period is used by law enforcement academy (LEA) instructors and tactical strength and conditioning facilitators (TSAC-F) to train recruits to tolerate the physical and psychological challenges of policing, while also teaching the necessary procedures and skills required for the job [4–6]. However, not all recruits will graduate from academy. Recruits may separate (i.e., they do not graduate) for a number of different reasons. These reasons may include personal reasons (e.g., they no longer want to work in law enforcement) [3], physical training (PT) session failures (i.e., they do not complete the requisite number of sessions as mandated by the LEA or state) [7], injury [5,8,9], failure in academics or scenario-based training [10–12]. Recruits that separate create a significant financial burden to an agency [5,13]. Thus, it would be pertinent for agencies to understand whether there are certain physical characteristics that influence whether a recruit graduates or separates from academy. If these characteristics could be effectively measured prior to academy, it may provide useful information for an agency to make more cost-effective decisions as to whether they hire certain individuals.

A major component of academy is PT, which should be tailored towards developing recruits such that they can complete the tasks required in law enforcement. Greater aerobic fitness (as measured via number of shuttles in the 20-m multistage fitness test; MSFT) and strength endurance (e.g., push-up and sit-up repetitions) have been correlated with better performance in job-specific tasks [6,12], which highlights the need for fitness in law enforcement populations. Specific to academy graduation, Shusko et al. [13] found that Massachusetts-based recruits in the USA who completed fewer push-ups in 60 s prior to academy, and had a slower 2.4-km run time, were more likely to separate. Orr et al. [9] found that lower-body power measured via vertical jump (VJ) performance was a predictor of injury or illness in Australian police recruits. Accordingly, physical fitness would likely be a factor influencing a recruit's ability to successfully fulfill the requirements of academy and graduate.

A limitation with PT in the academy setting is that agencies may lack the equipment and space to conduct a variety of training practices (i.e., maximal strength training) which could be useful for a law enforcement officer (LEO) [6]. As a result, many academies tend to focus on strength endurance or callisthenic-type exercises conducted in a circuit training fashion [14], in addition to aerobic-focused training (e.g., long, slow distance or formation runs) [15]. These practices are often conducted within a paramilitary 'one-size-fits-all' training model [5,6,8,14,15]. This style of training may lead to an inappropriate application of training load for certain recruits, which could then increase their risk of injury and/or separation [5,16]. The PT practices adopted by agencies may place greater importance on certain physical qualities for recruits (i.e., if running is a focus, then aerobic fitness may be more important for a recruit). This should also be considered when analyzing the physical fitness qualities influencing academy graduation in LEA recruits.

Academy training is used not just to physically develop recruits, but also to challenge them psychologically in order to assess each recruit's tolerance and demeanor under stress [3]. Berg [2] noted that the stress imposed by the verbal commands of training staff is generally designed to test the mental toughness and resilience of recruits, and find any character flaws that may impede being an effective LEO. As an example, a recruit that displays poor emotional stability in the face of adversity during academy may not be the best candidate to become a LEO, where there are major consequences to poor decisions made in the field under stress. Notably, higher physical and psychological stress has been associated with thoughts about quitting the profession in correctional officers [17]. Recruits with lower levels of fitness may not only find the PT more difficult, but could also experience greater psychological stress within the academy environment. Noting the general adaptation of the biological system to stress proposed by Selye [18], the combination of stress imparted by both physical and psychological stressors may lead to system exhaustion and in turn impact on motivation. Similarly, academic stress can have the same effect when combined with physical stress. As an example, in collegiate athletes, the risk of injury was found to higher during periods of high academic stress when compared to periods of lower academic stress [19]. To the author's knowledge, there has been no analysis of the relationship between physical fitness and voluntary decisions made by recruits to separate from a law enforcement training academy, nor regarding any relationships between physical fitness and academic failure during training academy.

Therefore, the purpose of this retrospective study was to analyze the effects physical fitness may have on academy graduation and reasons for separation in LEA recruits. Although there are limitations with conducting retrospective analyses, this is often a necessity in law enforcement research, due to the constraints and external demands placed on these populations. Furthermore, this is very common in the scientific literature [1,4,6,12,13,20–34]. For this study, the recruits who did not graduate were divided into groups according to whether they: Separated for personal reasons; separated due to PT failures (i.e., poor fitness) or injury; or separated due to academic or scenario failures. It was hypothesized that recruits who graduated would display superior physical fitness across the different assessments utilized in this study. This would occur regardless of the reason why a recruit may have separated.

2. Materials and Methods

2.1. Subjects

Retrospective analysis of five academy classes from one agency was conducted. This sample was comprised of 401 recruits (age: 27.30 ± 5.92 years; height: 1.74 ± 0.12 m; body mass: 80.27 ± 14.38 kg), which included 333 males (age: 27.31 ± 5.99 years; height: 1.76 ± 0.12 m; body mass: 83.26 ± 12.66 kg) and 68 females (age: 27.24 ± 5.66 years; height: 1.64 ± 0.07 m; body mass: 66.45 ± 13.89 kg). The five training cohorts started their academy within a calendar year in southern California. Any strength and conditioning or training programs prior to academy were generally completed voluntarily at the individual-level only by recruits [23,27]. Based on the archival nature of this analysis, the institutional ethics committee approved the use of pre-existing data (HSR-17-18-370).

2.2. Procedures

The data in this study were collected by staff working for one LEA, and the procedures have been detailed in the literature [23,27]. While tests of other physical capacities (e.g., flexibility, linear speed, strength, etc.) would have been beneficial to include, this was not possible given the confines of time, equipment, and logistical restrictions provided by the LEA. Nonetheless, the fitness capacities that were assessed within this study are typical of law enforcement recruits in the literature [6,20,23,26–28]. The staff (~20 per testing session) were all trained by a certified TSAC-F who verified the proficiency of the staff members before each session, and all staff followed strict instructions (which will be detailed) to conduct each test. Each recruit's age, height, and body mass were recorded at the start of academy training. Height was measured barefoot using a portable stadiometer (seca, Hamburg, Germany), while body mass was recorded by electronic digital scales (Health o Meter, Neosho, Missouri). As detailed by Lockie et al. [27], all tests were conducted outdoors on concrete or asphalt surfaces at the LEA's training facility on a day scheduled by the staff for the LEA. Testing typically occurred between the hours of 09:00–14:00 depending on recruit availability, and recruits generally did not eat in the 2–3 h prior to their testing session as they were completing non-strenuous activity and employee-specific documentation for the LEA. The weather conditions for testing were typical of the climate of southern California during a calendar year. Although conducting testing outdoors is not ideal, there was no indoor testing facility available for this LEA and these procedures were typical of staff from the LEA [6,27]. Recruits rotated through the assessments in small groups of 3–4 and were permitted to consume water as required during the testing session.

2.3. Push-Ups

Upper-body muscular endurance was assessed via a maximal push-up test where recruits completed as many repetitions as possible in 60 s. The protocol for this assessment followed that of established law enforcement research, where a tester placed a fist on the floor directly under the recruit's chest to ensure they descended to an appropriate depth [1,4,6,22,24,26,27,30,31,33,35]. All female recruits were partnered with a female tester. Recruits started in the standard 'up' position, with the body taut and straight, the hands positioned shoulder-width apart, and the fingers pointed forwards. On the start command, a tester began the stopwatch, and recruits flexed their elbows, lowered themselves until their chests contacted another tester's fist, before extending their elbows to return to the start position. The recruits performed as many push-ups as possible in 60 s using this technique.

2.4. Sit-Ups

Abdominal muscular endurance was assessed via the sit-up test, where recruits completed as many repetitions as possible in 60 s. The sit-up test was conducted according to procedures established in previous law enforcement research [1,4,6,20,22,24,26,27,30,31,36,37]. Recruits laid on their backs with their knees flexed to 90°, heels flat on the ground, and arms crossed across their chest and hands positioned on their opposing shoulders. The feet were held to the ground by a tester. On the start command, recruits raised their shoulders from the ground while keeping their arms crossed, and touched their elbows to their knees. They then descended back down until the shoulder blades contacted the ground. Recruits completed as many repetitions as possible with this technique in the 60-s time period.

2.5. 75-Yard Pursuit Run (75PR)

The 75-yard (68.58-meter) pursuit run (75PR) was designed to simulate a foot pursuit for a LEO [20,27,28], and is shown in Figure 1. Although yards is an Imperial measure, the 75PR is the standard name for this test [38], has been used in the literature [20,27,28], and will be used in this study for clarity. The recruit completed five linear sprints about a square grid (each side was 12.1 m), while completing four, 45° direction changes zig-zagging across the grid. Recruits were also required to step over three barriers that were 2.44 m long and 0.15 m high that simulated road-side curbs during three of the five sprints. Time was recorded via a stopwatch, from the initiation of movement at the start, until the recruit crossed the finish line. Timing via stopwatches is standard practice in LEA testing [1,4,20,22,26–30,37,39]. Furthermore, testers trained in the use of stopwatch timing procedures for running tests can record reliable data [40,41].

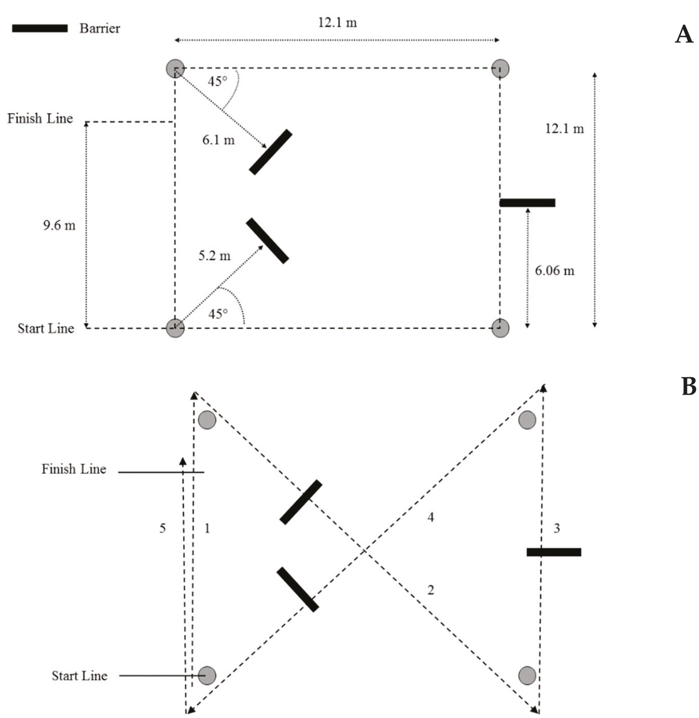


Figure 1. (**A**) The dimensions for the 75-yard pursuit run (75PR) in meters and (**B**) the running direction (numbered in order) for the 75PR.

2.6. Vertical Jump (VJ)

A Vertec apparatus (Perform Better, Rhode Island, USA) was used to measure the VJ, and followed established assessment protocols [4,23,30,37,42,43]. The recruit initially stood side-on to the Vertec (on the recruit's dominant side), reached upward as high as possible, and while keeping their heels on the ground, fully elevated the shoulder to displace as many vanes as possible. The last vane moved became the zero reference. The recruit then jumped as high as possible with no preparatory step, and tapped the highest vane they could with their dominant hand. Height was recorded from highest vane moved. No restrictions were placed on the range of countermovement during the jump. VJ height was calculated in inches by subtracting the standing reach height from the jump height, before being converted to cm [23,30]. Each recruit completed two trials, with a recovery time between trials of approximately 30–60 s, and the best trial used for analysis.

2.7. Medicine Ball Throw (MBT)

The MBT was used to indirectly measure upper-body power, and the procedures were adapted from the literature [23,27,44]. Recruits sat on the ground with their head, shoulders, and lower back against a concrete wall, and projected a 2-kg medicine ball (Champion Barbell, Texas, USA), which was lightly dusted with chalk, as far as possible using a two-handed chest pass. The measurement taken, using a standard tape measure, was the perpendicular distance from the wall to the chalk-marking closest to the wall made by the ball [23,44,45]. Two trials were completed, with a recovery time between trials of approximately 30–60 s, and the best trial was used for analysis.

2.8. 20-m Multistage Fitness Test (MSFT)

The MSFT was used to measure maximal aerobic capacity in the recruits, and followed established procedures [12,27,46]. Recruits were required to run back and forth between two lines spaced exactly 20 m apart, which were indicated by markers. The speed of running for this test was standardized by pre-recorded auditory cues (i.e., beeps) played from an iPad handheld device (Apple Inc., Cupertino, California, USA) connected via Bluetooth to a portable speaker (ION Block Rocker, Cumberland, Rhode Island, USA). The speaker was located in the center of the running area, and positioned in such that it would not interfere with the recruits. The test was terminated when the recruit was unable to reach the lines twice in a row in accordance with the auditory cues. This test was scored according to the final stage the recruit was able to achieve, and the stage was used to calculate the total number of completed shuttles.

2.9. Statistical Analysis

Information as to whether recruits completed academy and graduated, or did not and were separated, were provided by training staff from the LEA. Recruits were then split into four groups based on the information provided by LEA staff: Those that graduated (GRAD), and those that separated for personal reasons (SEPPR), PT failures or injury (SEPFI), or due to academic or scenario failures (SEPAS). Separation due to PT failures and injury were initially two separate groups. However, they were combined into one group, due to less fit recruits being more likely to get injured during academy [47], and because injuries often led to a recruit not completing the required number of PT sessions (which then resulted in academy separation). Sexes were combined within these groups, as all recruits need to attain the same standards to graduate academy, regardless of sex. This approach has been used in previous research [20,22,23,27,28].

Statistical analyses were computed using the Statistics Package for Social Sciences (Version 25.0; IBM Corporation, New York, USA) and Microsoft Excel (Microsoft Office Professional Plus 2016, Microsoft Corporation, Washington, WA, USA). Descriptive statistics (mean $\pm$ standard deviation [SD]) were calculated for each test parameter. A one-way analysis of variance (ANOVA), with Bonferroni post hoc for multiple comparisons, was used to calculate any performance differences in the fitness tests between the four groups. Significance was set at $p < 0.05$ a priori. Similar to previous research [23], effect sizes (d) were also calculated for the between-group comparisons, where the difference between the means was divided by the pooled SD [48]. In accordance with Hopkins [49], a d less than 0.2 was considered a trivial effect; 0.2 to 0.6 a small effect; 0.6 to 1.2 a moderate effect; 1.2 to 2.0 a large effect; 2.0 to 4.0 a very large effect; and 4.0 and above an extremely large effect. Multiple stepwise linear regression was used to determine whether age, height, body mass, or the physical fitness tests predicted graduation or reasons for separation in the recruits. As group inclusion was a categorical variable within SPSS, the data were recoded into dummy variables to provide dichotomous values (1 = group inclusion for either GRAD, SEPPR, SEPFI, or SEPAS; 0 = all other groups). Thus, GRAD, SEPPR, SEPFI, or SEPAS each acted as a dependent variable [6].

3. Results

Across the five classes, 330 recruits graduated (GRAD), while 71 recruits separated at various time points during academy. Of these, 28 recruits were placed in the SEPPR group, 18 in the SEPFI group, and 25 in the SEPAS group. Table 1 displays the descriptive data for all groups, while Table 2 shows the effect size data for the pairwise comparisons. The GRAD group was significantly younger than the SEPAS ($p < 0.01$) group, which had a moderate effect. With regards to the 75PR, the GRAD group was significantly faster than the SEPFI ($p = 0.02$) group, and this had a moderate effect. The GRAD group also completed significantly more MSFT shuttles than the SEPPR and SEPFI groups (both $p = 0.01$ with moderate effects). There were no significant between-group differences for height, body mass, the push-up and sit-up assessments, VJ, or MBT.

Table 1. Descriptive data (mean ± SD) for age, height, body mass, and fitness test performance data for law enforcement academy (LEA) recruits who graduated (GRAD) or separated (SEPPR, SEPFI, and SEPAS) from academy training.

	GRAD (n = 330)	SEPPR (n = 28)	SEPFI (n = 18)	SEPAS (n = 25)
Age (years)	26.67 ± 5.19	29.35 ± 8.02	29.59 ± 6.88	32.70 ± 9.01 *
Height (m)	1.75 ± 0.09	1.74 ± 0.10	1.72 ± 0.07	1.74 ± 0.08
Body Mass (kg)	80.69 ± 14.38	74.57 ± 14.52	79.50 ± 16.27	80.00 ± 12.54
Push-ups (no.)	42.48 ± 15.09	39.48 ± 14.01	34.63 ± 16.44	40.38 ± 12.24
Sit-ups (no.)	36.19 ± 9.04	35.78 ± 8.72	33.13 ± 7.59	34.29 ± 10.39
75PR (s)	16.97 ± 1.32	17.60 ± 1.21	17.94 ± 1.37 *	17.69 ± 1.28
VJ (cm)	53.60 ± 12.53	51.58 ± 13.43	47.94 ± 11.69	53.34 ± 11.83
MBT (m)	5.84 ± 1.22	5.52 ± 1.35	5.73 ± 1.29	5.96 ± 1.01
MSFT shuttles (no.)	52.75 ± 16.69	41.54 ± 10.74 *	39.94 ± 13.03 *	46.08 ± 11.19

* Significantly ($p < 0.05$) different from the GRAD group. GRAD, graduated; SEPPR, separated for personal reasons; SEPFI, separated for physical training failures or injury; SEPAS, separated for academic or scenario failures; 75PR, 75-yard pursuit run; VJ, Vertical Jump; MBT, Medicine Ball Throw; MSFT, multistage fitness test.

Table 2. Pairwise effect size data between LEA recruits who graduated (GRAD) or separated (SEPPR, SEPFI, and SEPAS) from academy training for age, height, body mass, number of push-ups and sit-ups completed in 60 s, time to complete the 75PR, VJ, MBT, and MSFT shuttles.

	GRAD-SEPPR	GRAD-SEPFI	GRAD-SEPAS	SEPPR-SEPFI	SEPPR-SEPAS	SEPFI-SEPAS
Age	0.40	0.48	0.82 *	0.03	0.39	0.39
Height	0.11	0.37	0.12	0.23	<0.01	0.27
Body Mass	0.42	0.08	0.05	0.32	0.40	0.03
Push-ups	0.21	0.50	0.15	0.32	0.07	0.40
Sit-ups	0.05	0.37	0.20	0.32	0.16	0.13
75PR	0.50	0.72 *	0.55	0.26	0.07	0.19
VJ	0.16	0.47	0.02	0.29	0.14	0.46
MBT	0.25	0.09	0.11	0.16	0.37	0.20
MSFT shuttles	0.80 *	0.86 *	0.47	0.13	0.41	0.54

* Moderate effect for the pairwise comparison.

The multiple stepwise linear regression data is shown in Table 3. Age and the MSFT predicted inclusion in the GRAD group, with 9% explained variance. The MSFT predicted inclusion in the SEPPR and SEPFI group, with 1–2% explained variance. Age predicted inclusion in the SEPAS group, with 5% explained variance.

Table 3. Stepwise linear regression analysis for inclusion in each group (GRAD, SEPPR, SEPFI, and SEPAS) and age, height, body mass, push-ups, sit-ups, 75PR, VJ, MBT, and MSFT.

Variables	r	r^2	Adjusted r^2
GRAD			
Age	0.263	0.069	0.067
Age, MSFT	0.310	0.096	0.091
SEPPR			
MSFT	0.129	0.017	0.014
SEPFI			
MSFT	0.143	0.020	0.017
SEPAS			
Age	0.223	0.050	0.047

4. Discussion

This study documented the differences in physical fitness between recruits from one LEA who graduated or separated from academy for a variety of reasons. The results provided some support to the study hypotheses. Recruits that graduated were faster in the 75PR and completed more MSFT shuttles when compared to recruits who separated, due to PT failures or injury. Graduating recruits also completed more MSFT shuttles than recruits who separated for personal reasons. However, there were no differences between any of the groups for performance in the push-up, sit-up, VJ, and MBT assessments. These data suggest that, for this agency, there are specific fitness characteristics that could influence and predict whether a recruit graduates academy. Although the stepwise regression data tended to have very low explained variance, it did tend to support the between-group comparison data. It should however be noted that, while the data exhibited a low albeit significant variance, the training stimulus provided during academy occurs across multiple occasions, rather than a single one-off event. As such, there is essentially a cumulative impact of risk. This phenomenon of differential risk accrual over repeated exposures to events occurring during the training program, is discussed in detail by Pope [50]. As will be also be discussed, this could be influenced by the PT practices adopted by the LEA.

The age, height, and body mass of the recruits in this study were typical of that established in the literature [6,20,23,27,28]. Height and body mass were not significantly different between any of the groups. However, the recruits in the SEPAS group were significantly older than the GRAD group, and age was a predictor of inclusion in this group. Age was also a predictor for inclusion in the GRAD group. Recruits in the SEPAS group either failed academic examinations or the requirements needed in law enforcement-specific training scenarios. Time management [51], and perceived control of time management [52], is important for academic success. Recruits in their late 20's and early 30's may have more outside life influences (e.g., family commitments) that could have influenced the results seen in this study for the SEPAS recruits. Noting the trend toward lower MSFT and slower 75PR results, older officers also tend to have lower levels of physical fitness when compared to their younger counterparts [22,23,31]. This could have influenced the recruits in the SEPAS groups' ability to adequately recover from the stressors of academy, and their subsequent performance in training scenarios. These notions, however, cannot be confirmed by the current research. Nonetheless, outside time commitments and differences in fitness for older recruits could play some role in their ability to complete the academic and scenario training requirements of academy, and could be points to consider for LEA staff.

The SEPFI group performed poorer in the two running assessments (75PR and MSFT) compared to the GRAD group, and the MSFT (albeit with a low explained variance) also predicted inclusion in this group. The 75PR measures high-intensity running and change-of-direction speed [27,28], while the MSFT provides a valid measure of aerobic capacity [53,54]. Running is a heavy focus of PT in the academy period for this LEA [15], so it is understandable that recruits who were superior in running tasks would be in a better position to complete the academy PT requirements and graduate. Furthermore, fitter recruits tend to operate at a lower percentage of their maximal capacity, and as a result can perform certain tasks for longer, fatigue less rapidly, and recover more quickly [55,56]. Shusko et al. [13] also found that police recruits who had a slower 2.4-km run time, which is indicative of lower aerobic fitness, were more likely to separate from academy. Recruits attending academy with poorer levels of fitness, as measured via running tasks, such as the 75PR and MSFT, could be more likely to separate. This could be especially true if a traditional LEA training model, with an emphasis on strength endurance and aerobic fitness, is followed [14,15].

The SEPPR recruits also completed fewer MSFT shuttles than the GRAD group, and the MSFT was a predictor (with a low explained variance) for inclusion in the SEPPR group. A factor that could influence whether a recruit separates for personal reasons is the inappropriate application of PT. As noted, many LEA academies operate via a paramilitary, 'one-size-fits-all' model [5,6,8,14,15], where every recruit is expected to complete the same training regardless of their current fitness or ability. Additionally and as noted, traditional LEA training can often involve a high volume of running [15]. Recruits in the SEPPR group may have found the training adopted by the LEA staff beyond their current physical capacity, which could then have contributed to their voluntary decision to separate. There are certain physiological characteristics that can predispose an individual to be a better runner (e.g., maximal aerobic capacity, lactate threshold, and running economy) [57,58]. While these qualities are trainable [57], there will always be recruits who are better suited to running than others (e.g., they have superior genetics relative to aerobic capacity or running biomechanics) [59,60]. LEA training staff should be wary that they do not lose recruits who are potentially competent LEOs that can perform the relevant job tasks, but are not good distance (e.g., 800 m or longer) runners. Physical fitness can still be improved with lower volumes of running in law enforcement populations, with a concurrent decrease in the risk of injury [5,8]. Given the findings of Trank et al. [61], whereby Naval recruits who ran <40 km (25 miles) where less likely to be injured than those who ran >40 km without negatively affecting physical readiness, future research should investigate moderating the volume and mileage of running during a law enforcement academy, and whether this can influence graduation and separation rates.

It is possible that certain recruits from the SEPPR group resigned for reasons not related to their physical fitness. Further to the PT requirements, a major part of academy is the development of the inherent discipline required in the profession. This can be imposed by the training staff offering very stern commands and directions for expected behaviours [2,3]. Although not measured in this study, there are certain personality types who would be more likely to voluntarily resign for personal reasons. Individuals that get more emotionally upset, and are more tender-minded, carefree, and impulsive, are more likely to voluntarily terminate employment as a police officer [62]. Given the high psychological stress imposed by training officers during academy [2,3], some recruits may have voluntarily resigned due to this stress, as opposed to that from PT. While this is a possibility, it is important to note that, due to the nature of academy, the psychological and physical stressors imposed by staff will be experienced simultaneously. Further, the recruits in the SEPPR group still did perform worse in the MSFT compared to the GRAD group. If a recruit had personality traits that did not fit well with the law enforcement profession, this could be compounded by poorer aerobic fitness, especially if PT had a high volume of running [15]. Future research should investigate the relationships between reasons for separation and personality traits, and other psychological variables (e.g., self-efficacy, motivation) that may influence academy separation.

Strength endurance tests, such as push-ups and sit-ups, predominate in the fitness assessment of law enforcement populations [1,4,6,13,20,22,24,26,27,30,31,33,35–37]. However, these assessments did not differentiate between recruits who graduated or separated for any reason, or predict any reasons for separation. This is contrary to findings from Shusko et al. [13], who found that recruits performing fewer push-ups tended to separate from LEA academies. The VJ and MBT also did not differentiate or predict recruits who graduated or separated. This contrasts with Orr et al. [9], who found poorer VJ performance was a predictor of injury or illness in Australian police recruits. Injury or illness can lead to academy separation for a recruit [5,8,9], especially if they miss a number of PT sessions [7]. However, the use of body weight calisthenics [14] and the volume of running [15] completed during PT at this academy could have limited the impact upper- or lower-body power may have had on recruit graduation rates. These results emphasize that the PT model adopted by LEA training staff will influence those fitness characteristics that could predispose a recruit to graduating or separating from academy.

There are certain limitations to this study that should be noted. This study incorporated a pre-existing fitness testing battery specific to one LEA. Different agencies may use different assessments (e.g., the 2.4-km run instead of the MSFT) [1,12,20,29–31,63]. This could influence the effects fitness may demonstrate with academy graduation rates. The study results could have also been influenced by the nature of the MSFT, where recruits can voluntarily terminate the test. Training staff could not always guarantee that all recruits reached their maximal aerobic capacity. Nonetheless, this test is an established and valid measure of maximal aerobic capacity [54,64,65], and the data from the recruits in this study was similar to that from previous law enforcement research [27]. Due to the sample size, recruits with PT failures and injuries were grouped together. Even though less fit recruits are more likely to get injured because the training load could exceed their capacities [47], it is still possible that certain injuries may have been suffered by recruits with higher levels of fitness. No maximal strength tests were included in the battery used in this study, although this is typical of law enforcement research [6,12,20,22–24,26–28,33]. Future research should investigate whether maximal strength could influence whether a recruit graduates or separates from academy. It would have been beneficial to monitor recruit fitness at multiple time points throughout academy. However, timetables can vary across different academy classes, due to the wide variety of skills and procedures that need to be taught to recruits [4–6], and ensuring the requisite staff are available to instruct. Nonetheless, this should be explored in future research.

5. Conclusions

The results from this study indicated that there were certain fitness characteristics that could influence reasons for academy separation in law enforcement recruits. Recruits that failed academic or training scenarios tended to be older than recruits who graduated. This could be influenced by time commitments external to academy for older recruits, or differences in fitness levels when compared to their younger counterparts that could influence recovery from academy stress. High-intensity running and change-of-direction speed, as measured by the 75PR, and aerobic fitness indicated by the MSFT, were poorer in recruits that separated, due to PT failures or injury. Recruits that separated for personal reasons also had lower aerobic fitness as measured by the MSFT. On the surface, these data suggest that LEA recruits should enhance their high-intensity running and change-of-direction speed (including strength, power, and sprint ability) and aerobic fitness prior to academy to improve their chances of successful graduation. However, these results could also have been influenced by the PT practices adopted by the LEA staff, which involved high volumes of running. Recruits should be aware of the PT training adopted by a LEA, so that they can best prepare for the rigors of that specific academy. Additionally, LEA training staff should consider the total training load they impose during academy to ensure it is not beyond the physical capabilities of some recruits, as this could contribute to injuries or negatively affect a recruit's motivation to graduate.

Author Contributions: Conceptualization, R.G.L.; Methodology, R.G.L., J.J.D., J.M.D., and R.M.O.; Investigation, R.G.L., K.B., A.M.B., M.R.M., K.A.C., J.J.D., J.M.D., and R.M.O.; Data Curation, R.G.L., K.B., A.M.B., M.R.M., and K.A.C.; Writing—Original Draft Preparation, R.G.L., K.B., A.M.B., M.R.M., K.A.C., J.J.D., J.M.D., and R.M.O.; Writing—Review and Editing, R.G.L., K.B., A.M.B., M.R.M., K.A.C., J.J.D., J.M.D., and R.M.O.

Funding: This research received no external funding.

Acknowledgments: The authors would like to thank the training instructors for facilitating this research, and the California State University, Fullerton tactical research team for collating the data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dawes, J.J.; Orr, R.M.; Siekaniec, C.L.; Vanderwoude, A.A.; Pope, R. Associations between anthropometric characteristics and physical performance in male law enforcement officers: A retrospective cohort study. *Ann. Occup. Environ. Med.* **2016**, *28*. [CrossRef]
2. Berg, B.L. First day at the police academy: Stress-Reaction-Training as a screening-out technique. *J. Contemp. Crim. Justice* **1990**, *6*, 89–105. [CrossRef]
3. Lockie, R.G.; Cesario, K.A.; Bloodgood, A.M.; Moreno, M.R. Heart rate response to psychological stress: Importance of stress education for law enforcement recruits. *TSAC Rep.* **2018**, *51*, 4–7.
4. Cocke, C.; Dawes, J.; Orr, R.M. The use of 2 conditioning programs and the fitness characteristics of police academy cadets. *J. Athl. Train.* **2016**, *51*, 887–896. [CrossRef] [PubMed]
5. Orr, R.M.; Ford, K.; Stierli, M. Implementation of an ability-based training program in police force recruits. *J. Strength Cond. Res.* **2016**, *30*, 2781–2787. [CrossRef] [PubMed]
6. Lockie, R.G.; Dawes, J.J.; Balfany, K.; Gonzales, C.E.; Beitzel, M.M.; Dulla, J.M.; Orr, R.M. Physical fitness characteristics that relate to Work Sample Test Battery performance in law enforcement recruits. *Int. J. Environ. Res. Public Health* **2018**, *15*, 2477. [CrossRef] [PubMed]
7. Peace Officer Standards and Training. *Training and Testing: Specifications for Learning Domain #32 Lifetime Fitness*; Commission on Peace Officer Standards and Training: West Sacramento, CA, USA, 2013.
8. Holmes, R.J.; Kornhauser, C.L. Building Injury Free Tactical Athletes: A Repeatable Model for Police Academy Recruit Training. In Proceedings of the TSAC Annual Training, Norfolk, VA, USA, 2–5 April 2018.
9. Orr, R.; Pope, R.; Peterson, S.; Hinton, B.; Stierli, M. Leg power as an indicator of risk of injury or illness in police recruits. *Int. J. Environ. Res. Public Health* **2016**, *13*, 237. [CrossRef]
10. Lockie, R.; Cesario, K.A.; Bloodgood, A.; Moreno, M. Physiological responses to defensive tactics training in correctional populations—Implications for health screening and physical training. *TSAC Rep.* **2018**, *48*, 4–8.
11. Dawes, J.J.; Kornhauser, C.L.; Crespo, D.; Elder, C.L.; Lindsay, K.G.; Holmes, R.J. Does body mass index influence the physiological and perceptual demands associated with defensive tactics training in state patrol officers? *Int. J. Exerc. Sci.* **2018**, *11*, 319–330.
12. Dawes, J.J.; Lindsay, K.; Bero, J.; Elder, C.; Kornhauser, C.; Holmes, R. Physical fitness characteristics of high vs. low performers on an occupationally specific physical agility test for patrol officers. *J. Strength Cond. Res.* **2017**, *31*, 2808–2815. [CrossRef]
13. Shusko, M.; Benedetti, L.; Korre, M.; Eshleman, E.J.; Farioli, A.; Christophi, C.A.; Kales, S.N. Recruit fitness as a predictor of police academy graduation. *Occup. Med.* **2017**, *67*, 555–561. [CrossRef] [PubMed]
14. Moreno, M.R.; Cesario, K.A.; Bloodgood, A.M.; Dulla, J.M.; Lockie, R.G. Heart rate response of a custody assistant class to circuit training during the academy period. In Proceedings of the Southwest American College of Sports Medicine's 37th Annual Meeting, Long Beach, CA, USA, 20–21 October 2017.
15. Cesario, K.A.; Moreno, M.R.; Bloodgood, A.M.; Dulla, J.M.; Lockie, R.G. Heart rate responses of a custody assistant class to a formation run during academy training. In Proceedings of the Southwest American College of Sports Medicine's 37th Annual Meeting, Long Beach, CA, USA, 20–21 October 2017.
16. Pope, R.P.; Herbert, R.; Kirwan, J.D.; Graham, B.J. Predicting attrition in basic military training. *Mil. Med.* **1999**, *164*, 710–714. [CrossRef] [PubMed]
17. Slate, R.N.; Vogel, R.E. Participative management and correctional personnel: A study of the perceived atmosphere for participation in correctional decision making and its impact on employee stress and thoughts about quitting. *J. Crim. Justice* **1997**, *25*, 397–408. [CrossRef]
18. Selye, H. Stress and the General Adaptation Syndrome. *Br. Med. J.* **1950**, *1*, 1383–1392. [CrossRef] [PubMed]

19. Mann, J.B.; Bryant, K.R.; Johnstone, B.; Ivey, P.A.; Sayers, S.P. Effect of physical and academic stress on illness and injury in Division 1 college football players. *J. Strength Cond. Res.* **2016**, *30*, 20–25. [CrossRef] [PubMed]
20. Cesario, K.A.; Dulla, J.M.; Moreno, M.R.; Bloodgood, A.M.; Dawes, J.J.; Lockie, R.G. Relationships between assessments in a physical ability test for law enforcement: Is there redundancy in certain assessments? *Int. J. Exerc. Sci.* **2018**, *11*, 1063–1073.
21. Lockie, R.G.; Dawes, J.J.; Kornhauser, C.L.; Holmes, R.; Orr, R.M. Young officers drive faster, but older officers crash less: Results of a police pursuit driving course. *Police Sci. Aust. N. Z. J. Evid. Based Polic.* **2018**, *3*, 37–41.
22. Lockie, R.G.; Dawes, J.J.; Kornhauser, C.L.; Holmes, R.J. A cross-sectional and retrospective cohort analysis of the effects of age on flexibility, strength endurance, lower-body power, and aerobic fitness in law enforcement officers. *J. Strength Cond. Res.* **2017**. [CrossRef]
23. Lockie, R.G.; Dawes, J.J.; Orr, R.M.; Stierli, M.; Dulla, J.M.; Orjalo, A.J. An analysis of the effects of sex and age on upper- and lower-body power for law enforcement agency recruits prior to academy training. *J. Strength Cond. Res.* **2018**, *32*, 1968–1974. [CrossRef]
24. Lockie, R.G.; Dulla, J.M.; Stierli, M.; Cesario, K.A.; Moreno, M.R.; Bloodgood, A.M.; Orr, R.M.; Dawes, J.J. Associations between body mass and physical fitness assessments in male custody assistants from a law enforcement agency. *J. Aust. Strength Cond.* **2018**, *26*, 43–49.
25. Lockie, R.G.; Fazilat, B.; Dulla, J.M.; Stierli, M.; Orr, R.M.; Dawes, J.J.; Pakdamanian, K. A retrospective and comparative analysis of the physical Fitness of custody assistant classes prior to academy training. *Sport Exerc. Med. Open J.* **2018**, *4*, 44–51. [CrossRef]
26. Lockie, R.G.; Orr, R.M.; Stierli, M.; Cesario, K.A.; Moreno, M.R.; Bloodgood, A.M.; Dulla, J.M.; Dawes, J.J. The physical characteristics by sex and age for custody assistants from a law enforcement agency. *J. Strength Cond. Res.* **2018**. [CrossRef] [PubMed]
27. Lockie, R.G.; Ruvalcaba, T.R.; Stierli, M.; Dulla, J.M.; Dawes, J.J.; Orr, R.M. Waist circumference and waist-to-hip ratio in law enforcement agency recruits: Relationship to performance in physical fitness tests. *J. Strength Cond. Res.* **2018**. [CrossRef] [PubMed]
28. Lockie, R.G.; Stierli, M.; Dawes, J.J.; Cesario, K.A.; Moreno, M.R.; Bloodgood, A.M.; Orr, R.M.; Dulla, J.M. Are there similarities in physical fitness characteristics of successful candidates attending law enforcement training regardless of training cohort? *J. Trainol.* **2018**, *7*, 5–9. [CrossRef]
29. Moreno, M.R.; Lockie, R.G.; Kornhauser, C.L.; Holmes, R.J.; Dawes, J.J. A preliminary analysis of the relationship between the multistage fitness test and 300-m run in law enforcement officers: Implications for fitness assessment. *Int. J. Exerc. Sci.* **2018**, *11*, 730–738.
30. Orr, R.; Dawes, J.J.; Pope, R.; Terry, J. Assessing differences in anthropometric and fitness characteristics between police academy cadets and incumbent officers. *J. Strength Cond. Res.* **2018**. [CrossRef]
31. Dawes, J.J.; Orr, R.M.; Flores, R.R.; Lockie, R.G.; Kornhauser, C.; Holmes, R. A physical fitness profile of state highway patrol officers by gender and age. *Ann. Occup. Environ. Med.* **2017**, *29*. [CrossRef]
32. Dawes, J.J.; Orr, R.M.; Elder, C.L.; Krall, K.; Stierli, M.; Schilling, B. Relationship between selected measures of power and strength and linear running speed amongst Special Weapons and Tactics police officers. *J. Aust. Strength Cond.* **2015**, *23*, 23–28.
33. Dawes, J.J.; Orr, R.M.; Brandt, B.L.; Conroy, R.L.; Pope, R. The effect of age on push-up performance amongst male law enforcement officers. *J. Aust. Strength Cond.* **2016**, *24*, 23–27.
34. Boyce, R.W.; Jones, G.R.; Schendt, K.E.; Lloyd, C.L.; Boone, E.L. Longitudinal changes in strength of police officers with gender comparisons. *J. Strength Cond. Res.* **2009**, *23*, 2411–2418. [CrossRef]
35. Dawes, J.J.; Orr, R.M.; Elder, C.L.; Rockwell, C. Association between body fatness and measures of muscular endurance among part-time SWAT officers. *J. Aust. Strength Cond.* **2014**, *22*, 33–37.
36. Crawley, A.A.; Sherman, R.A.; Crawley, W.R.; Cosio-Lima, L.M. Physical fitness of police academy cadets: Baseline characteristics and changes during a 16-week academy. *J. Strength Cond. Res.* **2016**, *30*, 1416–1424. [CrossRef]
37. Beck, A.Q.; Clasey, J.L.; Yates, J.W.; Koebke, N.C.; Palmer, T.G.; Abel, M.G. Relationship of physical fitness measures vs. occupational physical ability in campus law enforcement officers. *J. Strength Cond. Res.* **2015**, *29*, 2340–2350. [CrossRef]
38. Los Angeles County Sheriff's Department. Physical Ability Test. Available online: http://lasdcareers.org/physical-ability-test/ (accessed on 9 March 2017).

39. Rossomanno, C.I.; Herrick, J.E.; Kirk, S.M.; Kirk, E.P. A 6-month supervised employer-based minimal exercise program for police officers improves fitness. *J. Strength Cond. Res.* **2012**, *26*, 2338–2344. [CrossRef]
40. Hetzler, R.K.; Stickley, C.D.; Lundquist, K.M.; Kimura, I.F. Reliability and accuracy of handheld stopwatches compared with electronic timing in measuring sprint performance. *J. Strength Cond. Res.* **2008**, *22*, 1969–1976. [CrossRef] [PubMed]
41. Mann, J.B.; Ivey, P.J.; Brechue, W.F.; Mayhew, J.L. Validity and reliability of hand and electronic timing for 40-yd sprint in college football players. *J. Strength Cond. Res.* **2015**, *29*, 1509–1514. [CrossRef]
42. Lockie, R.G.; Schultz, A.B.; Jordan, C.A.; Callaghan, S.J.; Jeffriess, M.D.; Luczo, T.M. Can selected functional movement screen assessments be used to identify movement deficiencies that could affect multidirectional speed and jump performance? *J. Strength Cond. Res.* **2015**, *29*, 195–205. [CrossRef]
43. Lockie, R.G.; Schultz, A.B.; Callaghan, S.J.; Jordan, C.A.; Luczo, T.M.; Jeffriess, M.D. A preliminary investigation into the relationship between functional movement screen scores and athletic physical performance in female team sport athletes. *Biol. Sport* **2015**, *32*, 41–51. [CrossRef] [PubMed]
44. Lockie, R.G.; Moreno, M.R.; Bloodgood, A.M.; Cesario, K.A. Practical assessments of power for law enforcement populations. *TSAC Rep.* **2018**, *49*, 6–12.
45. Lockie, R.G.; Callaghan, S.J.; Jordan, C.A.; Luczo, T.M.; Jeffriess, M.D. Does the trunk stability push-up provide an effective measure of upper-body function specific to male team sport athletes? *J. Athl. Enhanc.* **2013**, *2*. [CrossRef]
46. Robinson, J.; Roberts, A.; Irving, S.; Orr, R.M. Aerobic fitness is of greater importance than strength and power in the load carriage performance of specialist police. *Int. J. Exerc. Sci.* **2018**, *11*, 987–998.
47. Molloy, J.M.; Feltwell, D.N.; Scott, S.J.; Niebuhr, D.W. Physical training injuries and interventions for military recruits. *Mil. Med.* **2012**, *177*, 553–558. [CrossRef]
48. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Lawrence Earlbaum Associates: Hillsdale, NJ, USA, 1988; p. 567.
49. Hopkins, W.G. How to interpret changes in an athletic performance test. *Sportscience* **2004**, *8*, 1–7.
50. Pope, R. Prediction and Prevention of Lower Limb Injuries and Attrition in Army Recruits. Ph.D. Thesis, Charles Sturt University, Bathurst, Gambia, 2002.
51. Britton, B.K.; Tesser, A. Effects of time-management practices on college grades. *J. Educ. Psychol.* **1991**, *83*, 405–410. [CrossRef]
52. Macan, T.H.; Shahani, C.; Dipboye, R.L.; Phillips, A.P. College students' time management: Correlations with academic performance and stress. *J. Educ. Psychol.* **1990**, *82*, 760–768. [CrossRef]
53. Paradisis, G.P.; Zacharogiannis, E.; Mandila, D.; Smirtiotou, A.; Argeitaki, P.; Cooke, C.B. Multi-stage 20-m shuttle run fitness test, maximal oxygen uptake and velocity at maximal oxygen uptake. *J. Hum. Kinet.* **2014**, *41*, 81–87. [CrossRef] [PubMed]
54. Ramsbottom, R.; Brewer, J.; Williams, C. A progressive shuttle run test to estimate maximal oxygen uptake. *Br. J. Sports Med.* **1988**, *22*, 141–144. [CrossRef] [PubMed]
55. Knapik, J.J. The importance of physical fitness for injury prevention: Part 1. *J. Spec. Oper. Med.* **2015**, *15*, 123–127.
56. Orr, R.M.; Caust, E.; Hinton, B.; Pope, R. Selecting the best of the best: Associations between anthropometric and fitness assessment results and success in police specialist selection. *Int. J. Exerc. Sci.* **2018**, *11*, 785–796.
57. Midgley, A.W.; McNaughton, L.R.; Jones, A.M. Training to enhance the physiological determinants of long-distance running performance. *Sports Med.* **2007**, *37*, 857–880. [CrossRef]
58. Brandon, L.J. Physiological factors associated with middle distance running performance. *Sports Med.* **1995**, *19*, 268–277. [CrossRef] [PubMed]
59. Kong, P.W.; de Heer, H. Anthropometric, gait and strength characteristics of Kenyan distance runners. *J. Sports Sci. Med.* **2008**, *7*, 499–504. [PubMed]
60. Williams, C.J.; Williams, M.G.; Eynon, N.; Ashton, K.J.; Little, J.P.; Wisloff, U.; Coombes, J.S. Genes to predict VO(2max) trainability: A systematic review. *BMC Genom.* **2017**, *18*, 831. [CrossRef] [PubMed]
61. Trank, T.V.; Ryman, D.H.; Minagawa, R.Y.; Trone, D.W.; Shaffer, R.A. Running mileage, movement mileage, and fitness in male U.S. Navy recruits. *Med. Sci. Sports Exerc.* **2001**, *33*, 1033–1038. [CrossRef] [PubMed]
62. Drew, J.; Carless, S.A.; Thompson, B.M. Predicting turnover of police officers using the sixteen personality factor questionnaire. *J. Crim. Justice* **2008**, *36*, 326–331. [CrossRef]

63. Brewer, J.; Buckle, P.; Castle, P. The use of work place physiological measurements to establish the minimum fitness standards required for entry into the United Kingdom Police Service. *J. Athl. Enhanc.* **2013**, *2*. [CrossRef]
64. Léger, L.A.; Mercier, D.; Gadoury, C.; Lambert, J. The multistage 20 metre shuttle run test for aerobic fitness. *J. Sports Sci.* **1988**, *6*, 93–101. [CrossRef] [PubMed]
65. Mayorga-Vega, D.; Aguilar-Soto, P.; Viciana, J. Criterion-related validity of the 20-m shuttle run test for estimating cardiorespiratory fitness: A meta-analysis. *J. Sports Sci. Med.* **2015**, *14*, 536–547. [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

Roadside Exposure and Inflammation Biomarkers among a Cohort of Traffic Police in Kathmandu, Nepal

Kabindra M. Shakya [1,*], Richard E. Peltier [2], Yimin Zhang [3] and Basu D. Pandey [4]

1 Villanova University, Department of Geography and the Environment, Villanova, PA 19085, USA
2 University of Massachusetts, Department of Environmental Health Science, Amherst, MA 01003, USA; rpeltier@schoolph.umass.edu
3 Villanova University, Department of Mathematics and Statistics, Villanova, PA 19085, USA; yimin.z@villanova.edu
4 Kathmandu and Everest International Clinic and Research Center, Sukraraj Tropical and Infectious Disease Hospital, Kathmandu 9045, Nepal; drbasupandey@gmail.com
* Correspondence: kabindra.shakya@villanova.edu; Tel.: +1-610-519-3590; Fax: +1-610-519-3338

Received: 3 January 2019; Accepted: 25 January 2019; Published: 29 January 2019

Abstract: Air pollution is a major environmental problem in the Kathmandu Valley. Specifically, roadside and traffic-related air pollution exposure levels were found at very high levels exceeding Nepal air quality standards for daily $PM_{2.5}$. In an exposure study involving traffic police officers, we collected 78 blood samples in a highly polluted spring season (16 February 2014–4 April 2014) and 63 blood samples in the less polluted summer season (20 July 2014–22 August 2014). Fourteen biomarkers, i.e., C-reactive protein (CRP), serum amyloid A (SAA), intracellular adhesion molecule (ICAM-1), vascular cell adhesion molecule (VCAM-1), interferon gamma (IFN-γ), interleukins (IL1-β, IL-2, IL-4, IL-6, IL-8, IL-10, IL-12, IL-13), and tumor necrosis factor (TNF-α) were analyzed in collected blood samples using proinflammatory panel 1 kits and vascular injury panel 2 kits. All the inflammatory biomarker levels were higher in the summer season than in the spring season, while particulate levels were higher in the spring season than in the summer season. We did not find significant association between 24-hour average $PM_{2.5}$ or black carbon (BC) exposure levels with most of analyzed biomarkers for the traffic volunteers working and residing near busy roads in Kathmandu, Nepal, during 2014. Inflammation and vascular injury marker concentrations were generally higher in females, suggesting the important role of gender in inflammation biomarkers. Because of the small sample size of female subjects, further investigation with a larger sample size is required to confirm the role of gender in inflammation biomarkers.

Keywords: roadside exposure; air pollution; inflammation biomarker; Nepal; $PM_{2.5}$

1. Introduction

Air pollution is an important environmental health challenge across the world. Many studies have reported adverse health effects associated with roadside traffic exposures [1–3]. Air pollution has been linked with various health effects such as emergency room visits, childhood obesity [4], reduced lung function [3] preterm birth [5], autism [6], kidney disease [7], dementia [8], cardiovascular and respiratory illnesses, and overall mortality [9]. Studies have identified a large number of illnesses that are associated with air pollution exposure [10].

Exposure to air pollutants, such as fine particles, causes pulmonary inflammation and results in illnesses such as atherothrombosis [11–13]. Cytokines, a group of peptides and proteins, are related to inflammatory response to particulate exposure [14]. Several inflammatory biomarkers have been found to be associated with particulate matter (PM) exposure [15–17]. Cytokines such as interleukin-6 (IL-6) have been shown to be significantly correlated with traffic-related air pollution [18]. Traffic-related exposure increases the risk of cardiovascular disease and populations exposed to traffic pollutants are linked to elevated inflammation biomarkers and blood pressure [19]. IL-8, IL-1β, and C-reactive protein (CRP) are found to increase after ozone exposure [20]. Occupational or environmental PM exposure increases the levels of intracellular adhesion molecule (ICAM-1) and CRP [21]. Occupational exposure among people working as taxi drivers are related with inflammation biomarkers [22]. A study among trucking industry workers has found positive association of sICAM-1 but no significant association of IL-6 and hs-CRP with occupational particulate exposure [23]. Most animal models to date have shown that increasing doses of $PM_{2.5}$ enhance inflammation biomarker concentrations [24].

However, inconsistent results on specific biomarkers related to PM exposure have also been observed [25–31] and no statistical difference has been observed in cytokines levels after laboratory exposure to concentrated ambient particles (CAPs) [32,33]. There was no significant association between CRP, IL-6, and tumor necrosis factor (TNF-α) and CAPs exposure in a study in UK, US, and Canada [32–34] while a study in the Netherlands reported the positive association between $PM_{2.5}$ or PM_{10} with CRP [35]. Enhanced levels of inflammatory markers, such as CRP, IL-6, and TNF-α, suggest the future risk of cardiovascular diseases [25]. Individual responses to exposure may also vary and may be related with genetic predisposition [36] and other factors. Many studies are done in a laboratory in a controlled setting using animals or cell assays. Other ambient exposure studies are reported from mainly developed countries. In a review of published studies of systemic inflammation markers in humans, Møller et al. [26] reported a total of 25 studies from Europe, 23 studies from North America and 11 studies from Asia. Studies from Asia were mainly from developed or rapidly developing nations (e.g., China, Singapore, Taiwan, Iran, India, and Israel). Among rural Indian women, higher levels of IL-6, IL-8, and TNF-α in sputum were found from the households using biomass fuel than the ones using a cleaner fuel, i.e., liquefied petroleum gas [37]. There are overall fewer studies in understanding inflammatory responses for population exposed to traffic-related pollutants [30]. This study aims to assess the biomarker levels among the population that are routinely exposed to high particle levels in a developing country.

Kathmandu is the capital city of Nepal, and is undergoing rapid development. This has led to quickly increasing population, vehicles, and urbanization, thus leading to increased emission of air pollutants in the Kathmandu Valley. Because the major air pollution source in the valley is traffic, human exposures near busy roads are expected to be higher than those in other locations in the valley [38] and traffic police are likely to be high-risk groups for adverse effects [29]. To investigate the roadside exposure of air pollution and related health effects, a major field campaign was conducted in 2014. Roadside and residential particulate pollution, anthropogenic gases, and respiratory health effects associated with roadside exposure from the same study have been published elsewhere [3,39–41]. In summary, roadside exposures to $PM_{2.5}$ and black carbon (BC) were related to reduction of lung function among traffic police. Though $PM_{2.5}$ concentrations were greatly reduced during summer compared to those in spring, components such as BC and several elements were not much lower during summer compared to those in spring, indicating the important contribution of vehicular emissions in both seasons. Several studies [42,43] have reported the high level of particulate levels in the Kathmandu Valley. High particulate levels have also been reported from the other valley in Nepal [44]. The main objective of the current work is to assess the seasonal changes in inflammation biomarkers among traffic workers and analyze the association of biomarker concentrations with air pollution exposure. To our knowledge, this is the first study from Nepal analyzing a comprehensive suite of inflammation biomarkers to assess the effect of air pollution exposure. This study provides baseline data to compare the biomarker levels among the population at different environmental conditions in future studies.

Populations who work on roads such as taxi drivers are routinely exposed to high air pollution levels and are considered as high-risk groups [22]. Traffic police can also represent other roadside occupational exposures [45]. Higher chromosomal aberration frequencies in lymphocytes in Turkey [46] and higher biomarkers of inflammation/infection in Pakistan [47] are found in traffic police compared to those in control population. Monitoring biomarkers in such cohorts helps to investigate the effects of occupational exposure to pollution. Traffic police in the Kathmandu Valley work on roads and direct the flow of traffic because the studied area had no functioning traffic signals. These workers spend several hours per day in traffic and such exposures have been found to decrease the lung function after the occupational exposure [3]. Therefore, traffic police officers were selected to evaluate the occupational exposure of particulate pollution and the biomarker levels. We hypothesized that particulate levels will be associated with levels of inflammation biomarkers among traffic police officers in the Kathmandu Valley.

2. Method

2.1. Study Population

A total of 53 traffic volunteers were recruited for this study. Prior to the study, approval for the study was taken from the institutional review board at the University of Massachusetts, Amherst and Nepal Health Research Council, Nepal. Permission to carry out the study was also completed from the government of Nepal. A total of 36 traffic volunteers participated in air pollution exposure study in the spring season, and 30 traffic volunteers participated in the summer season, though not all subjects consented to provide blood samples, where 33 volunteers permitted blood samples in spring, and 29 volunteers permitted blood samples in summer. These samples were from 29 men and 4 women in spring, and 25 men and 4 women in summer. From the subjects who volunteered to provide blood samples, there were 13 smokers and 14 non-smokers in the spring season and there were 9 smokers and 21 non-smokers in the summer season. All the subjects were prescreened for their medical history of asthma, serious respiratory or heart diseases, tuberculosis, and diabetes, and these were used as exclusion criteria for participation in this study. Descriptive statistics of subject demographics are given in Table 1.

Table 1. Description of study subjects and air pollution during spring (n = 33) and summer (n = 29).

Parameter	Spring		Summer		
	Mean	SD	Mean	SD	p-value [4]
a. Description of subjects [1]					
BMI (kg/m^2)	32	5	32	3	
Age	28	5	28	5	
Employment (years)	5	3			
b. Air pollution measurements					
$PM_{2.5}$ (μg/m^3) [2]	123.51	38.66	45.21	24.09	0.000
BC (μgC/m^3) [2]	18.80	7.68	16.46	7.52	0.124
Passive sampling [3]					
Ozone (μg/m^3)	14.02	9.57	16.6	6.52	0.633
Sulfur dioxide (μg/m^3)	6.75	0.35	25.46	10.99	0.098
Nitrogen dioxide (μg/m^3)	103.94	15.65	102.02	51.21	0.939
Nitric oxide (μg/m^3)	134.03	41.64	126.33	94.30	0.903

[1] Based on subject samples in spring (n = 33) and summer (n = 29); [2] 24-hour average; [3] Passive sampling data are the one-week mean concentration from five sites [39]. [4] p-values are the results of two-sided independent t-tests.

2.2. Sampling

This study was conducted in two phases: spring (16 February to 4 April 2014) and summer (20 July to 22 August 2014). Six sites were selected for the study: Kalanki, Balaju, Chabahil, Koteswor, Thapathali, and Jawalakhel. These sites were selected because they were the busiest traffic intersections in the valley. Detailed descriptions of the sampling have previously been published [3,40]. Briefly, at each site, personal exposure of particulate pollution was monitored for six traffic volunteers for up to six days (Sunday to Friday). Sunday is a working day in Nepal. While working on the road, each of the six traffic volunteers carried a bag containing air pollution monitors. These six traffic volunteers worked around the vicinity of each of the six sites. We followed thirty-six traffic volunteers from six sites for six weeks during spring and thirty traffic volunteers from five sites during summer. Volunteers were requested to wear N-95 masks for half of the week (3 days) as an intervention component of the study, where personal exposure to ambient pollutants was dramatically decreased. At each site, all volunteers were requested to wear masks either on the first half of the week (Sunday to Tuesday) or on the second half of the week (Wednesday to Friday). This was done to avoid confounding effects between the use of mask and day of the week. In general, all volunteers have a similar number of working hours (8–10 h), with two working shifts: morning and afternoon. The average temperature during sampling periods in spring and summer was 14.8 °C and 23.6 °C, respectively; relative humidity in spring and summer was 73.2% and 88.0%, respectively; total precipitations in spring and summer were 50.47 and 266.6 mm, respectively [40].

2.3. Blood Sample Collection and Biomarker Analysis

Blood samples were collected by a professional phlebotomist working at a local hospital. Three blood samples per subject were taken in the beginning, middle, and end of the week separately. In total, there were 78 blood samples from 33 traffic volunteers in the spring season and 63 blood samples from 29 traffic volunteers in the summer season. Blood samples were centrifuged for 15 min within 24 h of sample collection. After centrifugation, serum samples were stored in a freezer at −20 °C. Standard deep freezers at –80 °C were not available, and are, in fact, rare in Nepal. Samples were then transported under refrigeration to our laboratory at the University of Massachusetts, and then stored at −80 °C. No sample quality checks were performed in this set of samples to assess sample degradation.

V-PLEX assay kits (Mesoscale Discovery, Rockville, MD, USA) were used for biomarker analysis. Proinflammatory Panel 1 (human) kits were used for analyzing 10 cytokines: interferon gamma (IFN-γ), interleukins (IL-1β, IL-2, IL-4, IL-6, IL-8, IL-10, IL-12p70, and IL-13), and tumor necrosis factor (TNF-α) (Table 2). Vascular injury panel 2 (human) kits were used for analyzing serum amyloid A (SAA), CRP, vascular cell adhesion molecule (VCAM-1), and intercellular adhesion molecule (ICAM-1). Protocols from the respective kits were followed to analyze the biomarkers on Discovery Workbench (Mesoscale Discovery, Rockville, MD, USA).

Table 2. Comparisons of biomarker concentrations for traffic volunteers in Kathmandu between spring and summer, 2014.

Biomarkers	Independent Samples						Dependent Samples				
	Summer		Spring		Independent *t*-Test		Summer–Spring		*t*-Test		Wilcox Test
	Mean	Std. Dev.	Mean	Std. Dev.	*t* stat.	*p*-value	Mean Diff.	Std. Dev. Diff.	*t* stat	*p*-value	*p*-value
CRP [1]	4.11	3.26	2.47	2.59	1.91	0.064	1.98	2.71	2.53	0.028 *	0.042 *
SAA [1]	2.16	2.29	0.81	1.11	2.59	0.014 *	1.49	1.76	2.94	0.013 *	0.016 *
ICAM-1 [1]	0.83	0.41	0.64	0.21	2.07	0.046 *	0.15	0.53	0.96	0.356	0.380
VCAM-1 [1]	0.76	0.39	0.59	0.21	1.77	0.086	0.18	0.51	1.22	0.249	0.176
IFN-γ [2]	25.53	7.89	17.76	7.73	2.56	0.014 *	5.58	8.89	2.17	0.052	0.064
IL-1β [2]	3.03	1.94	1.46	0.28	3.91	0.001 ***	0.59	0.53	3.85	0.003 **	0.001 ***
IL-2 [2]	7.42	2.41	3.53	1.35	6.89	0.000 ***	4.96	3.16	5.43	0.000 ***	0.001 ***
IL-4 [2]	0.37	0.11	0.19	0.08	6.34	0.000 ***	0.18	0.09	6.84	0.000 ***	0.001 ***
IL-6 [2]	2.26	0.82	0.89	0.40	7.33	0.000 ***	1.21	0.71	5.93	0.004 **	0.001 ***
IL-8 [2]	64.29	63.54	12.98	4.76	3.78	0.001 **	78.93	74.56	3.67	0.001 ***	0.001 ***
IL-10 [2]	2.31	0.59	1.28	0.62	5.83	0.000 ***	1.13	0.81	4.85	0.008 **	0.002 **
IL-12 [2]	2.18	0.51	1.42	0.43	5.56	0.000 ***	0.61	0.65	3.25	0.003 **	0.012 *
IL-13 [2]	1.49	0.59	0.56	0.34	6.64	0.000 ***	0.62	0.57	3.78	0.008 **	0.005 **
TNF-α [2]	5.16	2.09	3.23	1.19	3.90	0.000 ***	1.79	2.15	2.88	0.015 *	0.001 **

Concentrations: [1] μg/mL; [2] pg/mL; level of significance: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. All the tests were two-sided and conducted on the data of each individual's average biomarker concentrations. The independent *t*-test performed a two-sample *t*-test, assuming independence of the volunteer samples in the two seasons. Dependent samples only counted the subjects with biomarker measurements in both seasons. Both a parametric *t*-test and a nonparametric Wilcox test were performed on the dependent samples.

2.4. Air Pollution Exposure

Real-time exposure levels of $PM_{2.5}$ and BC were measured from the individual subjects for 5–6 days using devices that were carried by participants. A nephelometer (pDR-1500, Thermo Scientific Inc., MA, USA) and a microaethalometer (AE51, AethLabs, CA, USA) were used to measure $PM_{2.5}$ and BC levels, respectively. Particulate samples were collected by polyflourotertaethylene (PFTE) filters for 24 h on nephelometers for post hoc analyses. The filters were analyzed for water-soluble ions by ion chromatography and for elements by X-ray fluorescence spectroscopy. Results for the air pollution and chemical components were published elsewhere [40]. Sampling lines used an inlet affixed to the breathing zone of a participant.

2.5. Data Analysis

Figure 1 shows the summary statistics of biomarker results in the spring and summer seasons of 2014. For each biomarker, outliers were identified by falling outside three standard deviations of the means for spring and summer separately. The means and standard deviations were calculated by assigning equal weights to each data entry (an individual could have multiple data entries if they consented to provide blood samples at multiple days of the week). These outliers were hence removed and the resultant data were used for all subsequent analysis. As a result, no data were removed for IL-1β and IL-12; <1% of data were removed for IFN-γ, IL-2, IL-4, IL-6, IL-13, TNF-α, and ICAM-1; <2% of data were removed for IL-8, and IL-10; 4% data were removed for VCAM-1; 5% of data were removed for SAA, and 6% of data were removed for CRP.

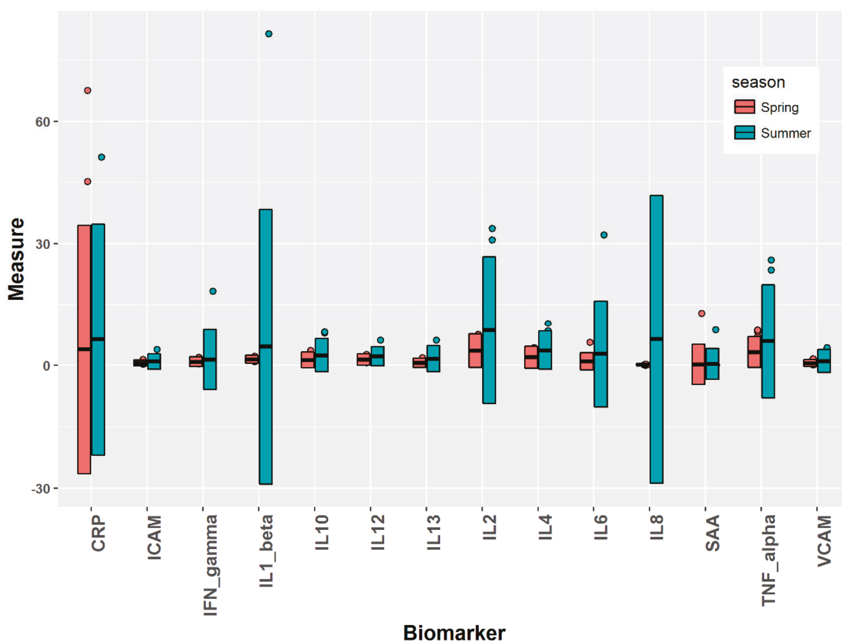


Figure 1. Concentrations of biomarkers during spring and summer seasons, 2014. CRP, SAA, ICAM-1, and VCAM-1 are shown in μg/mL; IFN-γ, IL1-β, IL-2, Il-4, IL-6, IL-8, IL-10, IL-12, IL-13, and TNF-α are given in pg/mL. In the figure, the central line across the box represents the mean, and the upper and lower boundaries correspond to three standard deviations above and below the mean, respectively. The outliers falling outside three standards of the mean are marked with circles. For ease of illustration, four biomarkers are rescaled: IL-4 upscaled by a factor of 10; SAA and IFN-γ downscaled by a factor of 1/20; IL-8 downscaled by a factor of 1/100.

To further investigate variations in biomarker measurements, linear mixed regression models were built, with each type of biomarkers as the response variable and the following factors as potential explanatory variables, such as season, $PM_{2.5}$, BC, gender, smoking, and the use of masks. The linear mixed model also made it happen to include the dependency structure of biomarkers from the same subjects who were followed in both seasons. Figure S1 (Supplementary Materials) illustrates the overall modelling process.

First, a linear mixed model can be formulated as follows:

$$y_{ij} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_6 x_6 + \eta_i + \epsilon_{ij}, \tag{1}$$

$$\eta_i \sim N\left(0,\ \sigma_\eta^2\right), \quad \epsilon_{ij} \sim N\left(0,\ \mathrm{var}\left(\epsilon_{ij}\right)\right), \tag{2}$$

where y_{ij} is the level of one biomarker from subject i at jth measurement and $x_1, \ldots, x_6$ are the explanatory variables, including $PM_{2.5}$, BC, season (spring or summer), smoking (Y or N), gender (F or M) and mask (Y or N), with corresponding coefficients $\beta_1, \ldots, \beta_6$. The variables x_1 to x_6 are treated as fixed effects and η_i denotes the random effect among subjects. Hence, the dependency between the biomarkers measured from the same subjects is included in the model by sharing the same term η_i. The term ϵ_{ij} is the random error in each measurement and its variance is assumed to be also associated with the explanatory variables, following the power-of-X dispersion function [48] specified below:

$$\mathrm{var}\left(\epsilon_{ij}\right) = \sigma_\varepsilon^2 \exp(\gamma_0 + \gamma_1 x_1 + \gamma_2 x_2 + \ldots + \gamma_6 x_6), \tag{3}$$

where σ_ε is an unknown parameter and γ_k $(k = 0,\ 1,\ \ldots, 6)$ are unknown parameters called the dispersion effects parameters.

Second, model diagnostics were performed on the full model to investigate the need for transformation on the biomarker response variables and to assess outliers. Box-Cox transformations were applied on the biomarker variables, if necessary, to correct non-normality. To examine assumptions with the transformed model, scaled residuals, defined through the Cholesky decomposition of the variance-covariance matrix, were obtained in place of raw residuals, as they tended to be uncorrelated with the constant mean zero. To further detect outliers and potentially influential data points, restricted likelihood distance [49] was used as an overall influence measure in addition to scaled residuals. Specifically, a data point with either a restricted likelihood distance or a scaled residual more than three was identified as an outlier and removed from the dataset. Once an outlier was deleted, the model on the updated dataset was refitted and an outlier was removed one at a time until no more outliers were detected in the residual analysis.

Third, on the transformed full model, two stepdown variable selection procedures were employed partially based on p-values of the effects. Specifically, using the complete fixed effects under consideration, a stepdown selection on the dispersion effects only was performed, until the lowest Bayesian information criterion (BIC) [50] was achieved. Then keeping these selected dispersion effects, another stepdown selection on the main effects only was performed and BIC was used to select the final model. Note that the last model from the second stepdown selection did not necessarily have the lowest BIC. Thus, sometimes, some insignificant effects were kept in the final model to attain a lower BIC, or some significant effects might be excluded for the same reason.

Finally, the adequacy of each final linear mixed model was checked by a residual analysis. As shown in Figure S1 (Supplementary Materials), outliers were identified and removed in the same manner until the model was free of outliers. The final results of the models for each biomarker could be found in Table 3.

Table 3. Effects estimate from linear mixed models.

Effects	Categories	Estimate	Std. Error	DF	*t* Value	Pr > | *t* |
CRP [1]						
Season	Summer	0.3462	0.0616	53	1.90	0.0630
$PM_{2.5}$		0.0010	0.0005	53	5.62	<0.0001
SAA [2]						
Season	Summer	1.4828	0.2742	59	5.41	<0.0001
$PM_{2.5}$		0.0029	0.0026	59	1.13	0.2637
Gender	Female	0.9548	0.3167	59	3.01	0.0038
VCAM-1 [2]						
Season	Summer	0.3154	0.0908	59	3.47	0.0010
ICAM-1 [2]						
Season	Summer	0.3679	0.0831	70	4.43	<0.0001
IL-1β [2]						
Season	Summer	0.3829	0.1124	67	3.41	0.0011
$PM_{2.5}$		−0.0018	0.0007	67	−2.45	0.0171
IL-2 [2]						
Season	Summer	0.8941	0.06547	72	13.59	<0.0001
Gender	Female	0.5940	0.1334	72	4.45	<0.0001
IL-4						
Season	Summer	0.1812	0.0154	66	11.77	<0.0001
Gender	Female	0.2077	0.01931	66	10.76	<0.0001
IL-6 [2]						
Season	Summer	0.9458	0.08211	71	11.52	<0.0001
Gender	Female	0.6877	0.1463	71	4.70	<0.0001
IL-8 [2]						
Season	Summer	1.1651	0.1411	59	8.26	<0.0001
IL-10						
Season	Summer	1.2331	0.1417	57	8.71	<0.0001
$PM_{2.5}$		0.0028	0.0009	57	3.07	0.0033
BC		-2×10^{-5}	4.44×10^{-6}	57	−3.55	0.0008
Mask	No	−0.1199	0.0477	57	−2.52	0.0147
Gender	Female	1.1191	0.2337	57	4.79	<0.0001
IL-12						
Season	Summer	0.6483	0.1048	62	6.19	<0.0001
$PM_{2.5}$		−0.0021	8.56×10^{-4}	62	−2.42	0.0184
Mask	No	−0.2006	0.0506	62	−3.97	0.0002
Gender	Female	0.8294	0.1720	62	4.82	<0.0001
IL-13						
Season	Summer	0.6104	0.1014	58	6.02	<0.0001
$PM_{2.5}$		−0.0019	0.0005	58	−4.08	<0.0001
Mask	No	−0.0963	0.0297	58	−3.24	0.0020
Gender	Female	0.9112	0.0576	58	15.81	<0.0001
IFN-γ [3]						
$PM_{2.5}$		−0.0044	0.0014	68	−3.22	0.0019
Gender	Female	0.9682	0.2299	68	4.21	<0.0001
TNF-α						
Season	Summer	1.4738	0.2158	65	6.83	<0.0001
BC		-5.00×10^{-5}	1.60×10^{-5}	65	−3.23	0.0020
Smoker	No	−0.7800	0.3269	65	−2.39	0.0200

[1] The 4th root of the variable was used in the model. [2] The log of the variable was used in the model. [3] The square root of the variable was used in the model.

3. Results

3.1. Air Pollution Exposure over Seasons

Table 1 compares particulate air pollution exposure between spring and summer. Independent two-sample t tests were conducted on $PM_{2.5}$, BC and other measurements from passive sampling.

From Table 1, $PM_{2.5}$ concentrations were significantly lower in the summer season than in the spring season, a generally expected finding given the generally cleaner conditions in summer at this location. While the mean $PM_{2.5}$ concentration was larger by a factor of 2.7 during spring compared to that in summer, the mean BC concentration was larger by a factor of 1.1 during spring compared to that in summer. Measurements were made from busy roads and roadsides, and the BC sources at these sites were related mainly with vehicular exhaust. Because traffic activities were not expected to be different during two seasons, the insignificant difference in BC concentrations suggested the importance of traffic-related PM sources in both seasons. The increased level of $PM_{2.5}$ from summer to spring was due to the higher regional pollution in spring than in summer. There were also additional sources such as seasonal operation of brick kilns and refuse burning during the spring season. In contrast to particle pollution, nitrogen oxides, sulfur dioxide, and ozone measurements obtained from passive sampling were not significantly different between the two seasons. Although the $PM_{2.5}$ level was shown to be associated with the season, they did not present highly overlapping power in explaining the biomarker response variables in the linear mixed models. It was shown that in the full model with the six independent variables, the variance inflation factors (VIFs) of $PM_{2.5}$ and season, which measure the dependency among variables, were both below 4, where the multicollinearity issue usually arises when any of the VIFs exceeds 5 or 10. More model results about the effect of $PM_{2.5}$ and season will be discussed in Section 3.3.

3.2. Biomarkers over Seasons

The comparison results of biomarker concentrations between spring and summer of 2014 are shown in Table 2. Three statistical tests were considered. The independent *t*-test performed a two-sample *t*-test assuming independence of the volunteer samples in the two seasons. However, there were repeated subjects who participated in the samples from both seasons. Unfortunately, we were not able to follow all subjects during both seasons. In fact, during the second phase of study in summer, more than half of subjects had been transferred to other locations. Only 13 subjects were repeated from the respective sites in both seasons. Out of these 13 repeated subjects, serum biomarker data were available for only 12 subjects. Hence, dependent samples only counted these 12 subjects. Then, a parametric two-sample *t*-test and a nonparametric Wilcox test were both performed on the dependent samples. All these three tests using independent or dependent samples were two-sided and conducted on the data of each individual's average biomarker concentrations. From Table 2, out of 14 biomarkers, 10 of them showed significant increases, at the significance level of 0.05, in summer than in spring (consistent across all three tests), whereas VCAM-1 exhibited no significant differences in any of the three tests. Three biomarkers, CRP, ICAM-1, and IFN-γ, showed significant differences only partially from one or two of the tests. Of these ten biomarkers with consistently significant results, eight of them were very significant with *p*-values below 0.001 while assuming independent samples and five of them were very significant using dependent samples. The top three biomarkers with the overall most significant differences were IL-2, IL-4, and IL-6.

3.3. Effects of Air Pollution and Various Factors on Biomarkers

In the previous section, test results appeared to show that season had a strong impact on biomarker concentrations. However, those results did not take into account the potential effects of other factors, along with season, on biomarkers. Hence, linear mixed regression models were run to investigate the effect of six variables simultaneously on biomarker levels: $PM_{2.5}$, BC, season, smoking habit, wearing mask, and gender. Table 3 lists the estimated coefficients of these variables as fixed effects in the linear mixed models for each biomarker. The results showed that $PM_{2.5}$ exposure was positively associated with three biomarkers, i.e., CRP, SAA, and IL-10, and negatively associated with four biomarkers, i.e., IL-1β, IL-12, IL-13, and IFN-γ, while holding all the other factors in the models constant. BC did not have an effect on most of biomarkers except two biomarkers, where it was negatively correlated with two biomarkers, i.e., Il-10 and TNF-α, but the effect was very small (Table 3). As expected, season had a strong effect on biomarker levels and it was the only factor that showed a significant effect for all the biomarkers except IFN-γ. Consistently, the concentration levels across all these biomarkers tended to be higher in summer than in spring.

Subject gender also had an effect on eight biomarkers: SAA, IL-2, IL-4, IL-6, IL-10, IL-12, IL-13, and IFN-γ. For these eight biomarkers, females tended to have higher concentration levels than males. Note that there were 7 female subjects out of 53 subjects. The effects of season and gender are illustrated in Figure 2 in the model for IL-6. In addition, it was found that wearing a protective facemask had a positive effect on IL-10, IL-12, and IL-13. The smoking habit was positively associated only with TNF-α with a p-value of 0.02.

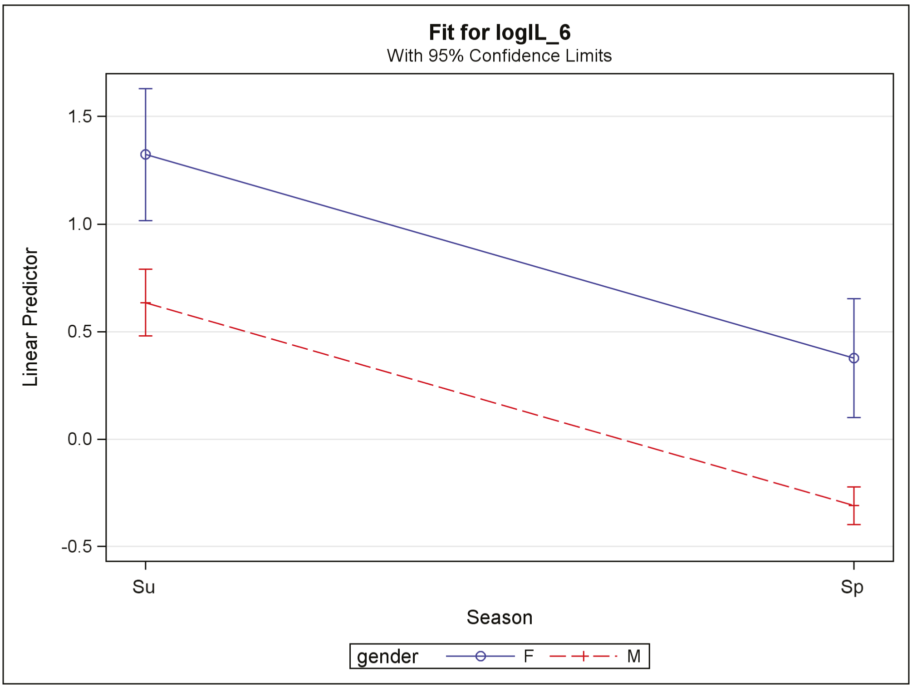


Figure 2. Predicted log (IL-6) concentration (with confidence intervals) versus season (Su: summer; Sp: spring) and gender (F: female; M: male).

Since the $PM_{2.5}$ concentration did not exhibit a significant effect on more than half of the biomarkers from the linear mixed models, we considered to further investigate the $PM_{2.5}$ chemical composition (elemental concentrations) and their relations with season on biomarkers. The results were plotted in Figure 3. It can be seen that the contributions of six elements—aluminum (Al), silica (Si), potassium (K), iron (Fe), nickel (Ni), and zinc (Zn)—were enhanced in summer than in spring. This showed that the variation among biomarker levels may be related to particulate composition more than the total particulate mass alone. Some of these elements such as Fe, Ni, and Zn are also related with traffic emissions [51], suggesting also the changes in traffic patterns in two seasons. However, there was no significant difference in BC concentrations between two seasons. Due to a large amount of missing values in element contributions, they were not included as explanatory variables in the linear mixed models of this study, but they could be valuable factors for biomarkers to be investigated in future studies. Besides these six elements, other elements were also measured on collected filters by X-ray fluorescence spectroscopy, but they were not enhanced during summer than during spring [40].

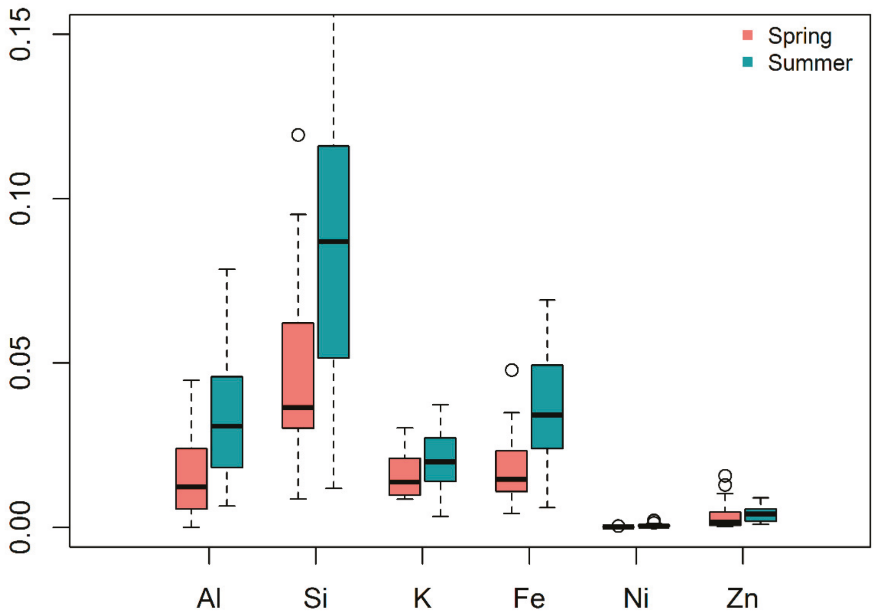


Figure 3. Normalized concentrations of elements (µg/m^3 of elements normalized by µg/m^3 of $PM_{2.5}$) during spring and summer seasons, 2014.

4. Discussion

4.1. Air Pollution Exposure

Exposure to $PM_{2.5}$ was generally higher in the spring season than in summer in the Kathmandu Valley (Table 1), but the elemental species contribution of $PM_{2.5}$ was found to be enhanced during summer than during spring (Figure 3). It was mainly $PM_{2.5}$ that saw the reduction in summer. BC levels were not significantly different between two seasons, suggesting only minor differences in traffic-related PM sources during two seasons. BC, emitted from incomplete combustion of fossil fuels and biomass, is used as a tracer of combustion sources and has been used to investigate traffic pollution in urban areas [52–54]. All pollution measurements were taken from roadside at the busiest intersections, where traffic-related activities were the major source of $PM_{2.5}$. There were also other sources such as dust resuspension, construction-related activities, refuse burning, brick-kiln, and these activities were minimized during summer than in spring. Meteorology might play an important role as well, and summer monsoonal conditions likely had an important determinative effect on $PM_{2.5}$, a pollutant with a multitude of sources by decreasing concentrations through atmospheric washout, whereas BC was less impacted by washout. Temperature and total precipitation were lower in the spring season compared to in summer. Thus, a reduced atmospheric boundary layer height and reduced wind, lower temperature and low or no precipitation may have impacted the removal of $PM_{2.5}$ during spring. Among the trace gases, ozone is slightly increased, but not significantly, during summer than in spring as it is formed through a photochemical process.

4.2. Biomarkers

On average, biomarker levels were higher during summer than in spring, despite having lower $PM_{2.5}$ levels. Statistical tests, after adjusting for factors such as smoker, age, and gender in the models, yielded mixed association between biomarkers and $PM_{2.5}$. Some biomarkers (CRP, SAA, and IL-10) were positively associated while other biomarkers (IL-1β, IL-12, IL-13, and IFN-γ) were negatively associated with $PM_{2.5}$. No significant association was found between $PM_{2.5}$ and seven biomarkers (VCAM-1, ICAM-1, IL-2, IL-4, IL-6, IL-8, and TNF-α). Previous studies have reported mixed findings for the association of various biomarkers with $PM_{2.5}$ exposures. A cross-sectional study among the healthy residents living in traffic congested areas in Thailand did not show significant association of $PM_{2.5}$ with IL-8 [27]. In the Greater Boston Area, Alexeeff et al. [28] found significant association between BC and sICAM-1, but no significant association between BC and VCAM-1. Positive association between $PM_{2.5}$ exposure and hs-CRP was found among traffic policemen in China [29] but negative association was found among workers at truck terminals in Northeastern US [30]. Lower levels of TNF-α were observed among adolescents living in a city with high $PM_{2.5}$ levels than in less polluted city in Bulgaria, and they attributed this to inhibition of cytokine production by particulates [31]. Results from this study suggested that $PM_{2.5}$ mass alone was not the sole important factor in affecting the biomarker levels on the studied subjects. There may be several other factors such as age, lifestyles, past environmental exposures, and ethnicity, which may have contributed to the variation of biomarkers. Among these factors, the following potential effects are discussed in subsequent paragraphs, such as (1) chemical composition, stress, and weather, (2) long-term occupational exposure, and (3) body fat and genetics.

Firstly, the higher levels of inflammatory biomarkers during the cleaner summer season might be due to relative fractions of PM chemical components rather than only PM mass concentration, which is consistent with other findings [14] that suggest that PM speciation might be more important than the PM concentration in determining biomarker changes in humans. For example, Carter et al. [55] attributed the increase in cytokines in human airway epithelia cells to the metals found in particles. Brucker et al. [56] found metals in blood were positively correlated with pro-inflammatory cytokines in taxi drivers. A similar effect may be playing a role in inducing inflammation and vascular injury markers in this cohort, even in the presence of the reduced bulk $PM_{2.5}$ concentration. Alternatively, the pattern of biomarkers could also be attributed to other factors, for example season- and weather-related stress. Humidity and temperature was higher during summer than during spring. The sampling period in summer coincided with the monsoon season in Nepal. All the traffic volunteers were working on these busy roads to direct the traffic and prevent and control traffic jams at these busy intersections. There were no operating traffic lights at the sampling locations. We met two times a day (early morning before they go to work; late afternoon after they return from work) with each of traffic volunteers and we visibly observed higher stress during summer than during spring.

Secondly, it is also possible that inflammation biomarkers were already enhanced in these traffic volunteers because of their high occupational exposure and thus the inflammation markers that we measured were immune to short-term influence from $PM_{2.5}$ personal exposure in our cohort study. Ying and Rajagopalan [57] suggest that the lack of association between short-term effects of PM exposure and inflammation biomarker does not necessarily mean that there is no effect from long-term exposure or there is no effect on other cytokine pathways. People working on roads such as taxi drivers have been observed with elevated levels of inflammation biomarkers compared to a control population [22].

Thirdly, it is also possible that the new recruits in summer may have other conditions causing high inflammation concentrations. Persons with excess body fat may have high inflammation biomarkers such as CRP and IL-6 [58,59]. However, mean BMIs were similar during two seasons (Table 1) and are not likely important here. Bind et al. [36] found genetics to play a role in chronic inflammation from air pollution exposure, though we lacked genetic information from this cohort. Certain population may have greater biological susceptibility or sociodemographic vulnerability [60]. Rückerl et al. [61,62] observed high association of inflammation biomarkers with air pollution in population with genetic susceptibility.

The study was conducted from six sites inside the Kathmandu Valley: Thapathali, Koteswor, Jawalakhel, Chabahil, Balaju, and Kalanki. The biomarker concentrations were not distributed evenly among the six sites studied. The highest biomarker concentrations were found in Thapathali, one of the cleaner locations in this study. Spatial variations among personal particulate exposure levels at six sites during two seasons are given in Table S1 (Supplementary Materials) and details are given in Shakya et al. [3,40]. As was shown from our model, gender had a significant effect on a large number of biomarkers (eight out of fourteen biomarkers), where female officers had higher concentrations of biomarkers compared to their male counterparts. Gender is known to play an important role in the degree of inflammation [63]. Other studies have shown higher biomarker levels in general in females than in males, e.g., CRP levels in the Dallas Heart Study [64]. Burnout, depression, and anxiety also affect differently inflammation biomarkers depending on gender [65]. Though female traffic volunteers were living together at a dormitory in the same house located at the sample sites, if they were still using biomass for cooking at their home, these dirtier indoor environment may contribute to higher inflammation biomarker levels [37]. That will also have more influence on females than males because females in Nepal are most likely to be responsible for cooking activities. Of all the six sites, Thapathali was the only site with all female traffic officers, all of whom were non-smokers, and consequently the highest biomarker concentrations were found in Thapathali. Because of allocation of females at only one site, we cannot discard the possibility of confounding effect of sites on gender. However, all the sites were located not very far from each other; distances among the six sites ranged from 3 to 11 km. The limitation of findings on gender is reiterated here due to very small female sample size. Further studies are needed before coming to conclusion regarding the role of gender on inflammation biomarker levels.

The lack of distinct association between $PM_{2.5}$ and inflammatory biomarkers, and adhesion molecules in this study does not exclude the possibility of chronic effects on the pulmonary inflammation and the cardiovascular system. Conclusions from the present study point to the complexity of explanatory variables, and limitations of sample size and short duration of study (i.e., 5–6 days per participant).

4.3. Limitations

This study has several limitations. Serum samples were stored in a freezer (−20 °C) at Kathmandu and were stored inside a freezer (−80 °C) only after arriving at the University of Massachusetts, Amherst. No quality checks were conducted to assess sample degradation due to storage at higher temperature. The study was conducted from the subjects who have unusually high levels of daily occupational exposure, and therefore likely reflects results specifically to individuals who have extreme exposures. The study was also complicated by not having the same individuals followed during both seasons. There are several variables that were untested in this study and these variables may be very important in affecting the health of the subjects. Perspiration, stress, and exhaustion might be high during hot summers than in cooler spring conditions. Psychological stress may also elevate inflammatory biomarkers such as IL-6 [59], though stress induced by busy traffic was not likely much different between spring and summer in this cohort as traffic patterns were generally unchanged. There was also an issue of compliance on wearing masks in summer because comfort levels in wearing masks were lower during summer than in spring. Respiratory allergies and illnesses which were not included in this study may also be playing an important role. Additionally, average 24-hour $PM_{2.5}$ used for daily exposure may not be the right indicator. Alternatively, a higher percentile closer to the peak occupational exposure may more accurately reflect the highly polluted environments of day-time traffic. There can be various other factors that were not considered in the current regression model such as age, lifestyles and years on the job. Though the subjects were living together at the dormitory at the sampling site during the study and their residential exposures were considered in this study, the prior residential history may still have played a role in their health. For example, female traffic police might have been using biomass for cooking activities in the past, and such history was not noted and considered in the current study.

5. Conclusions

Though air pollution exposure levels were higher in spring than in summer, measured biomarkers levels were higher during summer than in spring. The results showed that cytokine expression (biomarker levels) did not show dependency only with personal $PM_{2.5}$ exposure levels among traffic police and there may be other unmeasured factors such as genetics and stress. In general, short-term daily $PM_{2.5}$ exposure levels solely affected half of measured biomarker concentrations in a highly polluted environment in the Kathmandu Valley. Particularly, $PM_{2.5}$ had a mixed effect on these measured biomarker concentrations (positive association with 3 biomarkers and negative association with 4 biomarkers). Among the six variables tested for the statistical effect on individual biomarkers, season and gender were the most important variables affecting biomarker concentrations among traffic volunteers. This study has several limitations and further study is recommended to continue the investigation on potential factors that impact inflammation biomarkers.

Supplementary Materials: The following are available online at http://www.mdpi.com/1660-4601/16/3/377/s1. The strategy diagrams for building linear mixed models and variation of air pollution among six sites are given in supplementary information. Figure S1. Strategy diagram for building linear mixed models. Table S1. Variation of personal air pollution exposure at six sites between two seasons.

Author Contributions: Conceptualization, K.M.S. and R.E.P.; data collection and analysis, K.M.S.; Statistical analysis and software, Y.Z.; resources, K.M.S., R.E.P., and B.D.P.; writing—original draft preparation, K.M.S.; writing—review and editing, Y.Z. and R.E.P.

Funding: This study was supported by IASS in Potsdam, Germany, and the International Centre for Integrated Mountain Development (ICIMOD), Lalitpur, Nepal.

Acknowledgments: We acknowledge the Ministry of Home Affairs, the Metropolitan Traffic Police Division, and the traffic police volunteers for their cooperation in this research. We appreciate the help from the Nepal Health Research Council (NHRC) for the permission to conduct this study, and K. Aryal from NHRC for his cooperation and support on coordination with various government agencies. We thank the excellent support from our field research assistants A. Shahi, B. Gaire and J. Balami, and a local phelobotomist B. Ghimire. We also thank the help from B. Khadayat and R. R. Pokhrel. We also appreciate the help from M. Rupakheti at the Institute for Advanced Sustainability Studies (IASS) in Potsdam and R. Maskey at Tribhuvan University, Nepal.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Brook, R.D. You are what you breathe: Evidence linking air pollution and blood pressure. *Curr. Hypertens. Rep.* **2005**, *7*, 427–434. [CrossRef]
2. Brugge, D.; Durant, J.L.; Rioux, C. Near-highway pollutants in motor vehicle exhaust: A review of epidemiologic evidence of cardiac and pulmonary health risks. *Environ. Health* **2007**, *6*, 23. [CrossRef]
3. Shakya, K.M.; Rupakheti, M.; Aryal, K.; Peltier, R.E. Respiratory effects of high levels of particulate exposure in a cohort of traffic police in Kathmandu, Nepal. *J. Occup. Environ. Med.* **2016**, *58*, e218–e225. [CrossRef] [PubMed]
4. Jerrett, M.; McConnell, R.; Chang, C.C.R.; Wolch, J.; Reynolds, K.; Lurmann, F.; Gilliland, F.; Berhane, K. Automobile traffic around the home and attained body mass index: A longitudinal cohort study of children aged 10–18 years. *Prev. Med.* **2015**, *50*, S50–S58. [CrossRef] [PubMed]
5. Malley, C.S.; Kuylenstierna, J.C.I.; Vallack, H.W.; Henze, D.K.; Blencowe, H.; Ashmore, M.R. Preterm birth associated with maternal fine particulate matter exposure: A global, regional and national assessment. *Environ. Int.* **2017**, *101*, 173–182. [CrossRef] [PubMed]
6. Weisskopf, M.G.; Kioumourtzoglou, M.A.; Roberts, A.L. Air pollution and autism spectrum disorders: Causal or confounded? *Curr. Environ. Health Rep.* **2015**, *2*, 430–439. [CrossRef]
7. Bowe, B.; Xie, Y.; Li, T.; Yan, Y.; Xian, H.; Al-Aly, Z. Associations of ambient coarse particulate matter, nitrogen dioxide, and carbon monoxide with the risk of kidney disease: A cohort study. *Lancet Planet Health* **2017**, *1*, 267–276. [CrossRef]
8. Power, M.C.; Adar, S.D.; Yanosky, J.D.; Weuve, J. Exposure to air pollution as a potential contributor to cognitive function, cognitive decline, brain imaging, and dementia: A systematic review of epidemiologic research. *Neurotoxicology* **2016**, *56*, 235–253. [CrossRef] [PubMed]

9. Dockery, D.W.; Pope, C.A.; Xu, X.; Spengler, J.D.; Ware, J.H.; Fay, M.E.; Ferris, B.G., Jr.; Speizer, F.E. An association between air pollution and mortality in six U.S. cities. *N. Engl. J. Med.* **1993**, *329*, 1753–1759. [CrossRef]
10. Beil, L. The list of diseases linked to air pollution is growing. *Sci. News* **2017**, *192*, 18.
11. Ghio, A.J.; Kim, C.; Devlin, R.B. Concentrated ambient air particles induce mild pulmonary inflammation in healthy human volunteers. *Am. J. Respir. Crit. Care Med.* **2000**, *162*, 981–988. [CrossRef] [PubMed]
12. Pope, C.A., III; Hansen, M.L.; Long, R.W.; Nielsen, K.R.; Eatough, N.L.; Wilson, W.E.; Eatough, D.J. Ambient particulate air pollution, heart rate variability, and blood markers of inflammation in a panel of elderly subjects. *Environ. Health Perspect.* **2004**, *112*, 339–345. [CrossRef] [PubMed]
13. Kido, T.; Tamagawa, E.; Bai, N.; Suda, K.; Yang, H.H.; Li, Y.; Chiang, G.; Yatera, K.; Mukae, H.; Sin, D.D.; Van Eeden, S.F. Particulate matter induces translocation of IL-6 from the lung to the systemic circulation. *Am. J. Respir. Cell Mol. Biol.* **2011**, *44*, 197–204. [CrossRef] [PubMed]
14. Alfaro-Moreno, E.; Garcia-Cuellar, C.; De-Vizcaya-Ruiz, A.; Rojas-Bracho, L.; Osornio-Vargas, A. Cellular mechanisms behind particulate matter air pollution related health effects. In *Air Pollution: Health and Environmental Impacts*; Taylor & Francis Group: Boca Raton, FL, USA, 2010.
15. Takizawa, H.; Abe, S.; Ohtoshi, T.; Kawasaki, S.; Takami, K.; Desaki, M.; Sugawara, I.; Hashimoto, S.; Azuma, A.; Nakahara, K.; Kudoh, S. Diesel exhaust particles up-regulate expression of intercellular adhesion molecules-1 (ICAM-1) in human bronchial epithelial cells. *Clin. Exp. Immunol.* **2000**, *120*, 356–362. [CrossRef]
16. Ishii, H.; Fujii, T.; Hogg, J.C.; Hayashi, S.; Mukae, H.; Vincent, R.; van Eeden, S.F. Contribution of IL-1 beta and TNF-alpha to the initiation of the peripheral lung response to atmospheric particulates (PM10). *Am. J. Physiol. Lung Cell Mol. Physiol.* **2004**, *287*, L176–L183. [CrossRef] [PubMed]
17. Fujimaki, H.; Kurokawa, S.; Yamamoto, S.; Satoh, M. Distinct requirements for interleukin-6 in airway inflammation induced by diesel exhaust in mice. *Immunopharmacol. Immunotoxicol.* **2006**, *28*, 703–714. [CrossRef]
18. Zhang, X.; Staimer, N.; Gillen, D.L.; Tjoa, T.; Schauer, J.J.; Shafer, M.M.; Hasheminassab, S.; Pakbin, P.; Vaziri, N.D.; Sioutas, C.; Delfino, R.J. Associations of oxidative stress and inflammatory biomarkers with chemically-characterized air pollutant exposures in an elderly cohort. *Environ. Res.* **2016**, *150*, 306–319. [CrossRef]
19. Rioux, C.L.; Tucker, K.L.; Mwamburi, M.; Gute, D.M.; Cohen, S.A.; Brugge, D. Residential traffic exposure, pulse pressure, and C-reactive protein: Consistency and contrast among exposure characterization methods. *Environ. Health Perspect.* **2010**, *118*, 803–811. [CrossRef] [PubMed]
20. Devlin, R.B.; Duncan, K.E.; Jardim, M.; Schmitt, M.T.; Rappold, A.G.; Diaz-Sanchez, D. Controlled exposure of healthy young volunteers to ozone causes cardiovascular effects. *Circulation* **2012**, *126*, 104–111. [CrossRef]
21. Plummer, L.E.; Pinkerton, K.E.; Madl, A.K.; Wilson, D.W. Effects of nanoparticles on the pulmonary vasculature. In *Cardiovascular Effects of Inhaled Ultrafine and Nanosized Particles*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2011.
22. Brucker, N.; Moro, A.M.; Charão, M.F.; Durgante, J.; Freitas, F.; Baierle, M.; Nascimento, S.; Gauer, B.; Bulcão, R.P.; Bubols, G.B.; et al. Biomarkers of occupational exposure to air pollution, inflammation and oxidative damage in taxi drivers. *Sci. Total Environ.* **2013**, *463–464*, 884–893. [CrossRef] [PubMed]
23. Chiu, Y.-H.M.; Garshick, E.; Hart, J.E.; Spiegelman, D.; Dockery, D.W.; Smith, T.J.; Laden, F. Occupational vehicle-related particulate exposure and inflammatory markers in trucking industry workers. *Environ. Res.* **2016**, *148*, 310–317. [CrossRef] [PubMed]
24. Sun, Q.; Yue, P.; Deiuliis, J.A.; Lumeng, C.N.; Kampfrath, T.; Mikolaj, M.B.; Cai, Y. Ambient air pollution exaggerates adipose inflammation and insulin resistance in a mouse model of diet-induced obesity. *Circulation* **2009**, *119*, 538–546. [CrossRef]
25. Kodavanti, U.P.; Chen, L.-C.; Costa, D.L. Experimental studies in animals. In *Cardiovascular Effects of Inhaled Ultrafine and Nanosized Particles*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2011.
26. Møller, P.; Danielsen, P.H.; Karottki, G.D.; Jantzen, K.; Roursgaard, M.; Klingberg, H.; Jensen, D.M.; Chrotophersen, D.V.; Hemmingsen, J.G.; Cao, Y.; et al. Oxidative stress and inflammation generated DNA damage by exposure to air pollution particles. *Mutat. Res.* **2014**, *762*, 133–166. [CrossRef] [PubMed]
27. Vattanasit, U.; Navasumri, P.; Khadka, M.B.; Kanitwithayanun, J.; Promvijit, J.; Autrup, H.; Ruchirawat, M. Oxidative DNA damage and inflammatory responses in cultured human cells and in humans exposed to traffic-related particles. *Int. J. Hyg. Environ. Health* **2014**, *217*, 23–33. [CrossRef] [PubMed]

28. Alexeeff, S.E.; Coull, B.A.; Gryparis, A.; Suh, H.; Sparrow, D.; Vokonas, P.S.; Schwartz, J. Medium-term exposure to traffic-related air pollution and markers of inflammation and endothelial function. *Environ. Res. Perspect.* **2011**, *119*, 481–486. [CrossRef]
29. Zhao, J.; Gao, Z.; Tian, Z.; Xie, Y.; Xin, F.; Jiang, R.; Kan, H.; Song, W. The biological effects of individual-level PM2.5 exposure on systemic immunity and inflammatory response in traffic policemen. *Occup. Environ. Med.* **2013**, *70*, 426–431. [CrossRef]
30. Neophytou, A.M.; Hart, J.E.; Cavallari, J.M.; Smith, T.J.; Dockery, D.W.; Coull, B.A.; Garshick, E.; Laden, F. Traffic-related exposures and biomarkers of systemic inflammation, endothelial activation and oxidative stress: A panel study in the US trucking industry. *Environ. Health* **2013**, *12*, 105. [CrossRef]
31. Dobreva, Z.G.; Kostadinova, G.S.; Popov, B.N.; Petkov, G.S.; Stanilova, S.A. Proinflammatory and anti-inflammatory cytokines in adolescents from Southeast Bulgarian cities with different levels of air pollution. *Toxicol. Ind. Health* **2015**, *31*, 1210–1217. [CrossRef]
32. Ghio, A.J.; Hall, A.; Bassett, M.A.; Cascio, W.E.; Devlin, R.B. Exposure to concentrated ambient air particles alters hematologic indices in humans. *Inhal. Toxicol.* **2003**, *15*, 1465–1478. [CrossRef]
33. Mills, N.L.; Robinson, S.D.; Fokkens, P.H.; Leseman, D.L.; Miller, M.R.; Anderson, D.; Freney, E.J.; Heal, M.R.; Donovan, R.J.; Blomberg, A.; et al. Exposure to concentrated ambient particles does not affect vascular function in patients with coronary heart disease. *Environ. Health Perspect.* **2008**, *116*, 709–715. [CrossRef]
34. Brook, R.D.; Urch, B.; Dvonch, J.T.; Bard, R.L.; Speck, M.; Keeler, G.; Morishita, M.; Marsik, F.J.; Kamal, A.S.; Kaciroti, N.; et al. Insights into the mechanisms and emdiators of the effects of air pollution exposure on blood pressure and vascular function in healthy humans. *Hypertension* **2009**, *54*, 659–667. [CrossRef]
35. Strak, M.; Hoek, G.; Godri, K.J.; Gosens, I.; Mudway, I.S.; van Oerle, R.; Spronk, H.M.; Cassee, F.R.; Lebret, E.; Kelly, F.J.; et al. Composition of PM affects acute vascular inflammatory and coagulative markers—The RAPTES Project. *PLoS ONE* **2013**, *8*, e58944. [CrossRef] [PubMed]
36. Bind, M.A.; Coull, B.; Suh, H.; Wright, R.; Baccarelli, A.; Vokonas, P.; Schwartz, J. A novel genetic score approach using instruments to investigate interactions between pathways and environment: Application to air pollution. *PLoS ONE* **2014**, *9*, e96000. [CrossRef] [PubMed]
37. Dutta, A.; Roychoudhury, S.; Chowdhury, S.; Ray, M.R. Changes in sputum cytology, airway inflammation and oxidative stress due to chronic inhalation of biomass smoke during cooking in premenopausal rural Indian women. *Int. J. Hyg. Environ. Health* **2013**, *216*, 301–318. [CrossRef]
38. Aryal, R.K.; Lee, B.-K.; Karki, R.; Gurung, A.; Kandasamy, J.; Pathak, B.K.; Sharma, S.; Giri, N. Seasonal PM10 dynamics in Kathmandu Valley. *Atmos. Environ.* **2008**, *42*, 8623–8633. [CrossRef]
39. Kiros, F.; Shakya, K.M.; Rupakheti, M.; Regmi, R.P.; Maharjan, R.; Byanju, R.M.; Naja, M.; Mahata, K.; Kathayat, B.; Peltier, R.E. Variability of anthropogenic gases: Nitrogen oxides, sulfur dioxide, ozone and ammonia at Kathmandu Valley, Nepal. *Aerosol Air Qual. Res.* **2016**, *16*, 3088–3101. [CrossRef]
40. Shakya, K.M.; Rupakheti, M.; Shahi, A.; Maskey, R.; Pradhan, B.; Panday, A.; Puppala, S.P.; Lawrence, M.; Peltier, R.E. Near-road sampling of PM2.5, BC, and fine-particle chemical components in Kathmandu Valley, Nepal. *Atmos. Chem. Phys.* **2017**, *17*, 6503–6516. [CrossRef]
41. Shakya, K.M.; Peltier, R.E.; Shrestha, H.; Byanju, R.M. Measurements of TSP, PM10, PM2.5, BC, and PM chemical composition from an urban residential location in Nepal. *Atmos. Pollut. Res.* **2017**, *8*, 1123–1131. [CrossRef]
42. Chen, P.; Kang, S.; Li, C.; Rupakheti, M.; Yan, F.; Li, Q.; Ji, Z.; Zhang, Q.; Luo, W.; Sillanpaa, M. Characteristics and sources of polycylic aromatic hydrocarbons in atmospheric aerosols in the Kathmandu Valley, Nepal. *Sci. Total Environ.* **2015**, *538*, 86–92. [CrossRef]
43. Kim, B.M.; Park, J.-S.; Kim, S.-W.; Kim, H.; Jeon, H.; Cho, C.; Kim, J.-H.; Hong, S.; Rupakheti, M.; Panday, A.K.; et al. Source apportionment of PM_{10} mass and particulate carbon in the Kathmandu Valley, Nepal. *Atmos. Environ.* **2015**, *123*, 190–199. [CrossRef]
44. Tripathee, L.; Kang, S.; Rupakheti, D.; Zhang, Q.; Huang, J.; Sillanpaa, M. Water-soluble ionic composition of aerosols at urban location in the foothills of Himalaya, Pokhara Valley, Nepal. *Atmosphere* **2016**, *7*, 102. [CrossRef]
45. Crebelli, R.; Tomei, F.; Zijno, A.; Ghittori, S.; Imbriani, M.; Gamberale, D.; Martini, A.; Carere, A. Exposure to benzene in urban workers: Environmental and biological monitoring of traffic police in Rome. *Occup. Environ. Med.* **2001**, *58*, 165–171. [CrossRef] [PubMed]
46. Burgaz, S.; Demircigil, G.C.; Karahalil, B.; Karakaya, A.E. Chromosomal damage in peripheral blood lymphocytes of traffic policemen and taxi drivers exposed to urban air pollution. *Chemosphere* **2002**, *47*, 57–64. [CrossRef]

47. Kamal, A.; Cincinelli, A.; Martellini, T.; Malik, R.N. Linking mobile source-PAHs and biological effects in traffic police officers and drivers in Rawalpindi (Pakistan). *Ecotoxicol. Environ. Saf.* **2016**, *127*, 135–143. [CrossRef]
48. Carroll, R.J.; Ruppert, D. *Transformation and Weighting in Regression*; Chapman and Hall: London, UK, 1988.
49. Cook, R.D.; Weisberg, S. *Residuals and Influence in Regression*; Chapman and Hall: New York, NY, USA, 1982.
50. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **1978**, *6*, 461–464. [CrossRef]
51. Lawrence, S.; Sokhi, R.; Ravindra, K.; Mao, H.; Prain, H.D.; Bull, I.D. Source apportionments of traffic emissions of particulate matter using tunnel measurements. *Atmos. Environ.* **2013**, *77*, 548–557. [CrossRef]
52. Invernizzi, G.; Ruprecht, A.; Mazza, R.; De Marco, C.; Mocnik, G.; Sioutas, C.; Westerdahl, D. Measurement of black carbon concentration as an indicator of air quality benefits of traffic restriction policies within the ecopass zone in Milan, Italy. *Atmos. Environ.* **2011**, *45*, 3522–3527. [CrossRef]
53. Targino, A.C.; Gibson, M.D.; Krecl, P.; Rodrigues, M.V.C.; dos Santos, M.M.; Correa, M.P. Hotspots of black carbon and PM2.5 in an urban area and relationships to traffic characteristics. *Environ. Pollut.* **2016**, *218*, 475–486. [CrossRef]
54. Liu, M.; Peng, X.; Meng, Z.; Zhou, T.; Long, L.; She, Q. Spatial characteristics and determinants of in-traffic black carbon in Shanghai, China: Combination of mobile monitoring and land use regression model. *Sci. Total Environ.* **2019**, *658*, 51–61. [CrossRef]
55. Carter, J.D.; Ghio, A.J.; Samet, J.M.; Devlin, R.B. Cytokine production by human airway epithelial cells after exposure to an air pollution particle is metal-dependent. *Toxicol. Appl. Pharmacol.* **1997**, *146*, 180–188. [CrossRef]
56. Brucker, N.; Moro, A.; Charão, M.; Bubols, G.; Nascimento, S.; Goethel, G.; Barth, A.; Prohmann, A.C.; Rocha, R.; Moresco, R.; et al. Relationship between blood metals and inflammation in taxi drivers. *Clin. Chim. Acta* **2015**, *444*, 176–181. [CrossRef]
57. Ying, Z.; Rajagopalan, S. Hyptertension and vascular toxicity of PM. In *Environmental Cardiology*; The Royal Science of Chemistry: Cambridge, UK, 2011.
58. Visser, M.; Bouter, L.M.; McQuillan, G.M.; Wener, M.H.; Harris, T.B. Elevated C-reactive protein levels in overweight and obese adults. *JAMA* **1999**, *282*, 2131–2135. [CrossRef]
59. Yudkin, J.S.; Kumari, M.; Humphries, S.E.; Mohamed-Ali, V. Inflammation, obesity, stress and coronary heart diseas: Is interleukin-6 the link? *Atherosclerosis* **2000**, *148*, 209–214. [CrossRef]
60. Wu, W.; Jin, Y.; Carlsten, C. Inflammatory health effects of indoor and outdoor particulate matter. *J. Allergy Clin. Immunol.* **2018**, *141*, 833–844. [CrossRef] [PubMed]
61. Rückerl, R.; Hampel, R.; Breitner, S.; Cyrys, J.; Kraus, U.; Carter, J.; Dailey, L.; Devlin, R.B.; Diaz-Sanchez, D.; Koenig, W.; et al. Associations between ambient air pollution and blood markers of inflammation and coagulation/fibrinolysis in susceptible populations. *Environ. Int.* **2014**, *70*, 32–49. [CrossRef]
62. Rückerl, R.; Schneider, A.; Hampel, R.; Breitner, S.; Cyrys, J.; Kraus, U.; Gu, J.; Soentgen, J.; Koenig, W.; Peters, A. Association of novel metrics of particulate matter with vascular markers of inflammation and coagulation in susceptible populations–results from a panel study. *Environ. Res.* **2016**, *150*, 337–347. [CrossRef] [PubMed]
63. Zeltser, D.; Rogowski, O.; Berliner, S.; Mardi, T.; Justo, D.; Serov, J.; Rozenblat, M.; Avitzour, D.; Shapira, I. Sex differences in the expression of haemoorheological determinants in individuals with atherothrombotic risk factors and in apparently healthy people. *Heart* **2004**, *90*, 277–281. [CrossRef] [PubMed]
64. Khera, A.; McGuire, D.K.; Murphy, S.A.; Stanek, H.G.; Das, S.R.; Vongpatanasin, W.; Wians, F.H., Jr.; Grundy, S.M.; de Lemos, J.A. Race and gender differences in C-reactive protein levels. *J. Am. Coll. Cardiol.* **2005**, *46*, 464–469. [CrossRef] [PubMed]
65. Toker, S.; Shirom, A.; Shapira, I.; Berliner, S.; Melamed, S. The association between burnout, depression, anxiety, and inflammation biomarkers: C-reactive protein and fibrinogen in men and women. *J. Occup. Health Psychol.* **2005**, *10*, 344–362. [CrossRef]

International Journal of
Environmental Research and Public Health

Article

Incineration Kinetic Analysis of Upstream Oily Sludge and Sectionalized Modeling in Differential/Integral Method

Yanqing Zhang [1], Xiaohui Wang [2], Yuanfeng Qi [1,*] and Fei Xi [1]

1 School of Environmental and Municipal Engineering, Qingdao University of Technology, Qingdao 266033, China; zyq_luck@163.com (Y.Z.); vadaxi@163.com (F.X.)
2 Technical Test Center of Shengli Oil Field, Dongying 257001, China; wxhui583@126.com
* Correspondence: qiyuanfeng@qut.edu.cn; Tel.: +86-5328507126

Received: 27 December 2018; Accepted: 25 January 2019; Published: 29 January 2019

Abstract: As the most significant solid residue generated in the oil production industry, upstream oily sludge was regarded as hazardous waste in China due to its toxicity and ignitability, and to date, the incineration process has been considered the most efficient method in practice. Due to the complicated components of oily sludge, a kinetic model of the incineration process was difficult to build, and is still absent in engineering use. In this study, multiple non-isothermal thermogravimetric analysis (TGA) and differential scanning calorimetry (DSC) analysis were applied for the kinetic analysis of upstream oily sludge in air conditions. A viewpoint regarding the rules to sectionalize the reaction stages was raised, and a differential integral method to obtain the incineration kinetic model was provided. The results showed that four stages that were divided based on the weight-loss regions in the TGA curves and the endothermic/exothermic sections in the DSC curves were suitable to obtain an incineration kinetic model of oily sludge. The integral method was beneficial for obtaining the average activation energy of each stage, and the differential method was suitable for gaining the nth-order reaction rate equation and the pre-exponential factor before the operating temperature became lower than 635.968 °C. The average activation energies of stages one, two, three, and four were 60.87 KJ/mol, 78.11 KJ/mol, 98.82 KJ/mol, and 15.96 KJ/mol, respectively. The nth-order reaction rate equations and pre-exponential factors of stages one, two, and three were 0.82, 3.50, and 2.50, and $e^{13.32}$min^{-1}, $e^{19.69}$min^{-1}, and $e^{21.00}$min^{-1}, respectively.

Keywords: oily sludge; incineration; kinetic; modeling; differential/integral method

1. Introduction

Upstream oily sludge is the most significant solid waste generated in the oil production industry, and is mainly discharged from the crude oil storage process [1,2]. Ordinarily, crude oil is housed in oil tanks prior to being refined to petroleum products, and the heavier species are separated and settled at the bottom of the storage tanks [3–5]. The solid sediments are the major components of upstream oily sludge, which contains a high concentration of complex petroleum hydrocarbons (PHCs, e.g., asphaltenes, resins, and tar), fine solids, and heavy metals [6,7]. On account of the toxicity and ignitability characteristics, which represented a significant adverse effect to ecosystem and human health, both upstream and downstream oily sludge have been regarded as hazardous waste in China since 2008 [8–10]. A variety of oil recovery and/or sludge disposal methods have been studied for the treatment of upstream oily sludge, such as thermal treatment (incineration or pyrolysis) [11–13], solidification [14], solvent extraction [15,16], photocatalysis [17], ultrasonic treatment [18], and biodegradation [18–21]. In China, the incineration process was identified as the most efficient method for the disposal of upstream oily sludge, and has been successfully designed, established, and commercialized in the last few years. However, the other mentioned methods have been rarely applied in practice for failing to reach a compromised balance between satisfying the strict environmental regulations and maintaining a reasonable operating cost [22,23]. Various incinerators such as circulating fluidized bed combustion, rotary kiln, and chain boiler combustion were adopted in the industrial application and operated with a combustion temperature between 730–1200 °C [1,12,23]. Furthermore, excess air and auxiliary fuels were indispensable for the incineration process. The incineration product was directly affected by a variety of factors, including the pretreatment method, operating temperature, residence time, feedstock quality, and addition of auxiliary fuels [24].

Most of the current studies focused on the thermal co-treatment of upstream oily sludge with auxiliary solid waste and/or the by-products that exist in gaseous phases and solid residue [22,25,26]. Generally, thermal analysis occupied the pivotal position throughout the thermal treatment of solid waste, and was frequently studied in the dehydration, carbonization, and incineration of industrial waste such as red mud, sewage sludge, and antibiotic residues [27–29]. The thermogravimetric analysis (TGA) test was a representative non-isothermal method for thermal kinetics analysis that was sensitive enough to exhibit the weight loss of the reactant with the operating temperature/time. It was usually applied for the thermal decomposition of certain reactants in air or nitrogen conditions. Ordinarily, the reaction kinetics of thermal decomposition was represented by the nth-order reaction rate equation [30–33] (see Equations (1)–(3)). Furthermore, the Arrhenius equation (Equation (4)) was commonly utilized to describe the reaction rate constant.

$$d\alpha/dt = K(T)f(\alpha) \tag{1}$$

$$\alpha = (W_i - W_t/W_i - W_e) \times 100\%,\ \alpha \in (0\% - 100\%) \tag{2}$$

$$f(\alpha) = (1-\alpha)^n \tag{3}$$

$$K(T) = Aexp(-E_a/RT) \tag{4}$$

where t and T are the operating time and temperature; α is the conversion ratios of the reactant; and $d\alpha/dt$ is the relationship between the instantaneous conversion ratio and the operating time. In Equation (2), W_i, W_n, and W_e are the initial weight, weight at a certain time, and the final weight of the reaction, respectively. In Equations (1) and (4), $K(T)$ is the reaction rate constant. In Equation (3), $f(\alpha)$ is the nth-order reaction rate equation, and n is the reaction order. Meanwhile, in Equation (4), A is the pre-exponential factor; E_a is the activation energy; and R is the gas constant.

The main objective for thermal kinetics analysis was to obtain the basic three elements [33], i.e., the activation energy (E_a), the pre-exponential factor (A), and the representation of the nth-order reaction rate equation ($f(\alpha)$). Obviously, it was imprecise to distinguish the mass signal versus time or temperature in a single isothermal or non-isothermal thermogravimetric test. Thus, multiple non-isothermal thermogravimetric analyses were often applied for the thermal kinetic studies. If the relationships between the operating temperature and reaction time were in the form of Equation (5), and meanwhile, the heating rate was constant in a certain TGA test, Equation (1) could be re-written as Equation (6):

$$T = \beta t + T_o \Rightarrow dT/dt = \beta \tag{5}$$

$$d\alpha/dT = A/\beta \times exp(-E_a/RT)f(\alpha) \tag{6}$$

where T and T_o are the operating temperature and initial temperature, respectively; β is the heating rate, $d\alpha/dT$ is the relationships between the instantaneous conversion ratios and the operating temperature.

Equation (6) was the basic differential form for the study of thermal kinetic analysis, which represented the relationships between the instantaneous conversion ratios of the reactant with T under certain heating rates (β). The Friedman method [34] and Coats–Redfern method [35] were obtained by rearranging Equation (6) and applied to the thermal kinetic analysis in the pyrolysis process of oily sludge, effectively [12]. However, limitations for differential methods still existed and were mainly attributed to $d\alpha/dT$, which was dramatically affected by the background noise of the TGA test [36–40]. Therefore, the activation energy (E_a) obtained in differential methods was imprecise.

Based on Equation (6), the integral method for the study of thermal kinetic analysis could be deduced as follows:

$$d\alpha/dT = \frac{A}{\beta}exp(-E_a/RT)f(\alpha) \Rightarrow \frac{1}{f(\alpha)}d\alpha = \frac{A}{\beta}e^{-E_a/RT}dT \tag{7}$$

$$G(\alpha) = \int_0^1 \frac{1}{f(\alpha)}d\alpha = \frac{A}{\beta}\int_{T_o}^{T} e^{-E_a/RT}dT \tag{8}$$

where $G(\alpha)$ is the integral Equation of $f(\alpha)$; and T_o and T are the initial and final operating temperature, respectively.

The activation energy (E_a) and pre-exponential factor (A) could be obtained by rearranging Equation (8), and the negative effect of background noise could be avoided. However, the solution of the nth-order reaction rate equation $f(\alpha)$ and the reaction order (n) were hard to obtain. Therefore, both the differential method and integral method have their advantages and limitations during the acquisition of the activation energy, pre-exponential factor, and the nth-order reaction rate equation.

The pyrolysis kinetics analysis of oily sludge or plastic was reported in previous studies and the methods that were used are listed in Table 1, including the utilized thermal test method, the modeling method, and the basic three elements (E_a, A, and $f(\alpha)$ or n).

Table 1. Thermal kinetics analysis and modeling methods reported in the previous studies.

Materials	Atmosphere	Thermal Test Method	Modeling Method	Basic Three Elements	Reference
Oil sludge, Phenolic plastic	Nitrogen	TGA/DTG	Integral	E_a	[11,34]
Oil sludge	Nitrogen	TGA	Differential	E_a, A and n	[12]
Polyurethane Foams	Nitrogen	TGA	Differential	E_a	[30,31]
Chalcogenide $Ge_2Sb_2Te_5$	Nitrogen	DSC	Integral	E_a	[36]
Polyurethane	Nitrogen	TGA	Differential	E_a, A and n	[37]
Rice husk	Nitrogen	TGA	Integral	E_a	[39]

TGA: thermogravimetric analysis; DTG: derivative thermogravimetric analysis; DSC: differential scanning calorimetry.

The reaction schemes of upstream oily sludge incineration were extremely complex due to the complicated composition; meanwhile, the reaction mechanism and the corresponding kinetic parameters for the incineration of various intermediate products and by-products may differ with the change of heating rate and operating temperature regions. It is difficult to identify or distinguish whether or not the kinetics model is suitable for the different reaction stages during the incineration process only on the basis of TGA curves. However, few studies have been available concerning the incineration reaction kinetics or in both differential and integral modeling methods for upstream oily sludge.

The aims of the present work were as follows. (1) Both the multiple non-isothermal thermogravimetric analysis (TGA) and differential scanning calorimetry (DSC) were performed to study the incineration thermal kinetic of oily sludge, simultaneously. (2) The work aimed to provide a new viewpoint to sectionalize the reaction stages in TGA/DSC curves. (3) The work aimed to present and utilize a comprehensive differential integral method to obtain the incineration kinetics model in different reaction stages.

2. Materials and Methods

2.1. Materials and Reagents

Purified air for DSC/TGA analysis was purchased from Qingdao Fengtai Co., Ltd. (Qingdao, China). Oxygen gas for the heating value test, with 99.99% purity, was purchased from Qingdao

Chunfeng Co., Ltd. (Qingdao, China). The upstream oily sludge that was utilized in this study was obtained from a temporary storage bin in Shengli Oil Field, Dongying, Shandong province, China, and the samples appear to be viscous and black block.

2.2. Apparatus and Methods

The upstream oil sludge was dried in a recycle ventilation drier for 24 h at 105 °C; then, the heating value and ash content of the dried sample was analyzed by an automatic oxygen bomb calorimeter (SDAC6000, Sunday, Changsha, China) and automatic ash Fusion Tester (SDAF105b, Sunday, Changsha, China), respectively. The test of bulk density and moisture content for the undried upstream oily sludge was the same as that in previous studies [27,29].

First, seven to 15 mg of undried upstream oily sludge was employed for the TGA/DSC test and carried out in a SDT Q600 thermal analyzer (TA instrument, New Castle, PA, USA) under air atmosphere. In order to simulate the oxygen-rich conditions applied in practical incineration processes, ratios of the purified air to the initial weight of oily sludge were maintained over 10.00 mL/min to 1.00 mg. The operating temperatures for the TGA/DSC test were performed from 30 ± 2.5 °C to 900 °C. The heating rates (β_n) were five K/min, 10 K/min, 15 K/min, 20 K/min, and 25 K/min, and were labeled as β_a, β_b, β_c, β_d, and β_e, respectively. The sample of each experiment was weighted within a thousandth of an error and loaded into the quartz disk, which settled in the center of the equipment. A K-type thermocouple was inserted beside the quartz disk for measuring the operating temperature. Before the formal thermal test, the purified air controlled by a rotameter was injected into the equipment and lasted at least two hours for the purpose of purging. When the run was finished, the air was kept flowing until the temperature of the system returned to room temperature. The TGA and DSC data were simultaneously recorded, and the results of each experiment were repeated twice and averaged.

3. Results and Discussion

3.1. Sectionalized Rules and Peak-Thermal Kinetic Analysis

3.1.1. Characteristics of the Upstream Oily Sludge

Features of the upstream oily sludge are represented in Table 2. Compared with the characteristics of the oily sludge, which was studied by Jing [8] and Xu [9], the upstream oily sludge had higher ash content and heating values, but lower moisture, which was attributed to the quality of the crude oil and the additives that were utilized in the recovery and dehydration processes.

Table 2. Features of upstream oily sludge.

Parameters	Heating Values (MJ/Kg, dry basis)	Ash Content (wt.%, dry basis)	Moisture (wt.%)	Bulk Density (kg/m^3)
Values	35.45 ± 3.64	57.56 ± 1.23	3.24 ± 1.08	1366.25 ± 46.73

3.1.2. Sectionalized Rules for the Incineration Process of Upstream Oily Sludge

The results of TGA/DSC tests under oxygen-rich conditions for upstream oily sludge are shown in Figure 1 for when the heating rates of βn were β_a = 5 K/min (Figure 1A), β_b = 10 K/min (Figure 1B), β_c = 15 K/min (Figure 1C), β_c = 20 K/min (Figure 1D), and β_d = 25 K/min (Figure 1E).

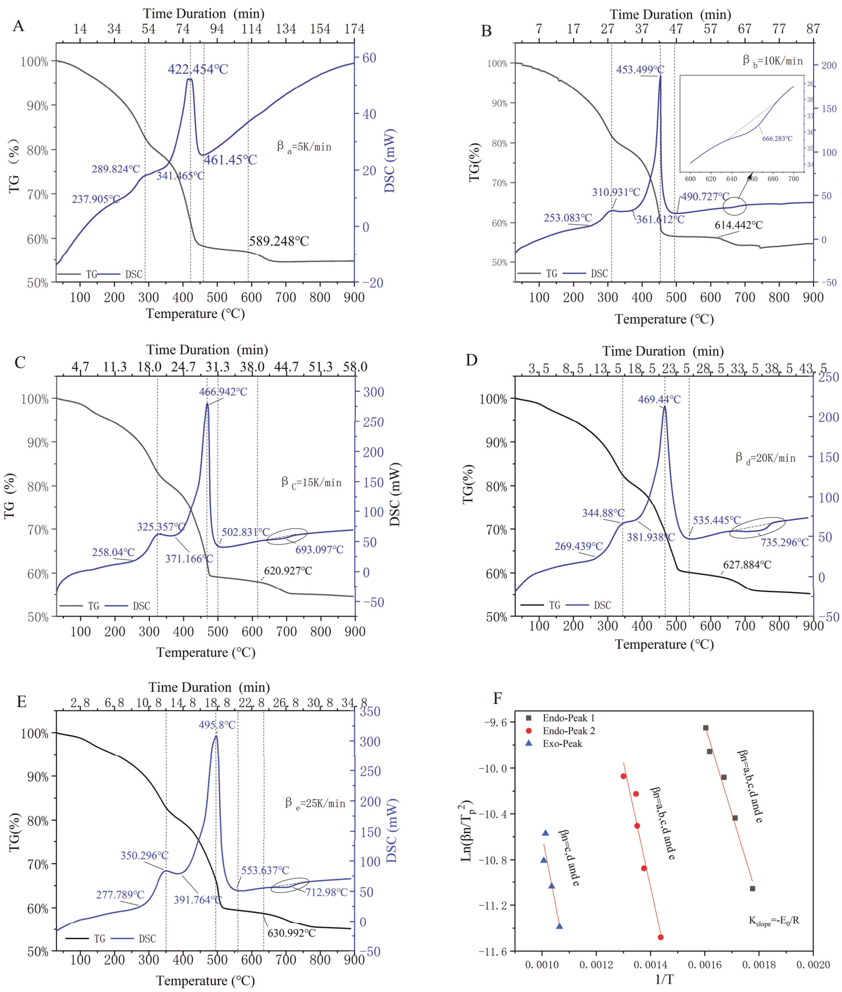


Figure 1. The DSC/TGA analysis of upstream oily sludge at β_n (n = a, b, c, d, and e); $\beta_a = 5$ K/min (**A**), $\beta_b = 10$ K/min (**B**), $\beta_c = 15$ K/min (**C**), $\beta_c = 20$ K/min (**D**) and $\beta_d = 25$ K/min (**E**), endothermic/exothermic peak analysis (**F**).

As shown in the five TGA curves (Figure 1A–E), the weight loss of upstream oily sludge in the incineration process was 45–47%, and three declining regions were simultaneous obtained. The first weight loss region was obtained from 30 °C to 280–350 °C, and the weight loss was 15–20%. The second weight loss region of 25–30% was obtained with the operating temperature ranging from 280–350 °C to 461.45–553.637 °C. The third weight loss region started at 589.248–630.992 °C with a slight weight loss of 1.5–5%.

As it was shown in the five DSC curves (Figure 1A–E), two significant exothermic reactions were detected at 237.905–391.764 °C and 341.465–553.637 °C, sequentially. It was the same as reported in the TGA/DTG test for the pre-dried oily sludge in nitrogen atmosphere [11,12]. Je-Lueng [12] considered that the former exothermic reactions in the pyrolysis process were attributed to the volatilization of volatile contents such as combined water, small hydrocarbons, and small molecular acids, while the latter exothermic reactions were caused by the decomposition of macromolecular compounds, such as for instance, tar, aromatic hydrocarbons, and cycloalkanes. Based on the above-mentioned studies, we inferred that the continuous exothermic reactions in the incineration process were possibly attributed to the volatilization and/or combustion of the volatile contents (in lower temperature region) and macromolecular compounds (in the higher temperature region), respectively. Although the components of volatile contents and macromolecular compounds were not the main objective of this work, the kinetics of each exothermic reaction seemed to be quite different. When the operating temperature exceeded 600 °C, with the increase of the heating rate, an endothermic phenomenon was gradually detected, which can probably be attributed to the decomposition of inorganic carbonate [27,28].

When the operating temperature was lower than 237.905 °C (Figure 1A–E), no exothermic or endothermic phenomenon occurred in the DSC curves, but a significant weight loss was observed in the TGA curves. When the exothermic or endothermic phenomenon appeared, the weight loss simultaneously accelerated. Both the endothermic and exothermic peak simultaneously shifted to the right with the increase of heating rate from βa to βe, and it was more sensitive and conspicuous in the exothermic regions.

Different types of volatilization decomposition and/or combustion reactions probably occurred in different weight loss regions, which were attributed to the complex components that existed in the upstream oily sludge. Therefore, both the mechanism and kinetic model changed with the increase of operating temperature. Based on whether the endothermic and/or exothermic reactions occurred (or not) in DSC tests, the TGA curves could be sectionalized as four weight loss stages; these are shown in Figure 2A.

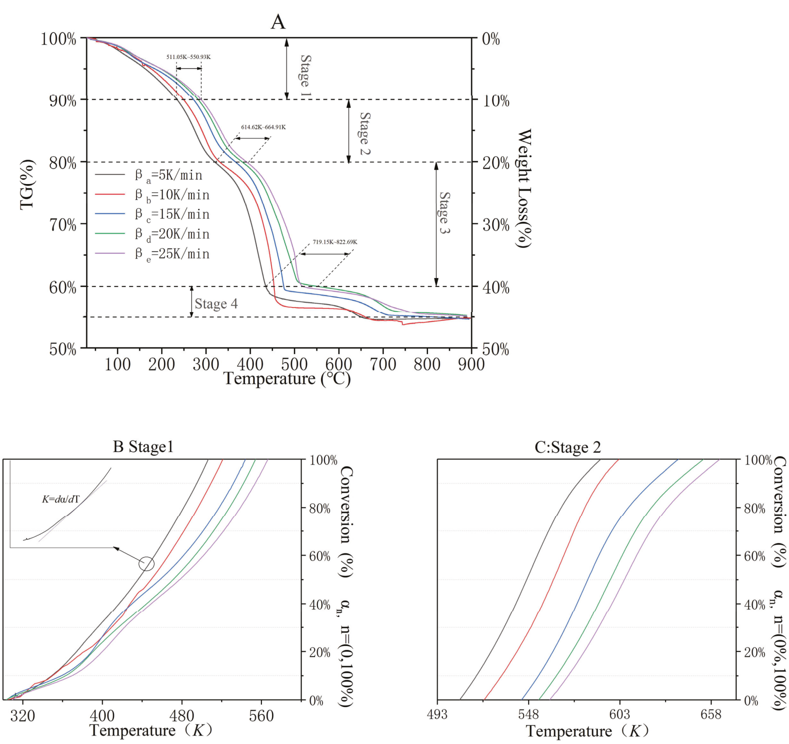


Figure 2. *Cont.*

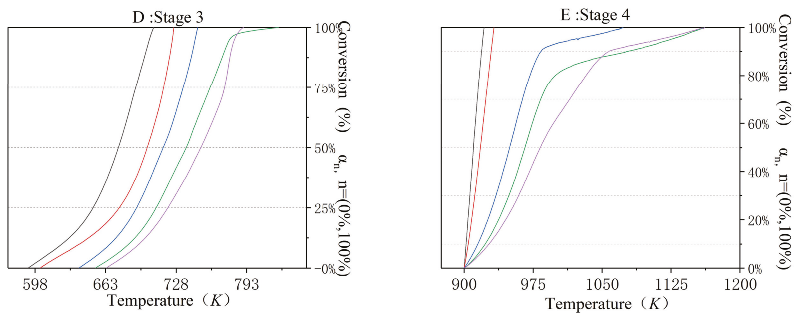


Figure 2. Four sectionalized stages in TGA. TGA test carried out in different heating rates (**A**); TGA test of stage 1 (**B**); TGA test of stage 2 (**C**); TGA test of stage 3 (**D**); TGA test of stage 4 (**E**).

As it was shown in Figure 1A, the ending temperature was 511.05–550.93 K in Stage 1, and no endothermic or exothermic reaction was detected, but a 10% weight loss ratio was obtained. The first and second exothermic reaction occurred in Stage 2 (weight loss = 10%) and Stage 3 (weight loss = 20%), and the ending temperatures were 614.62–664.91 K and 719.15–822.69 K, respectively. As shown in Figure 1B–E, the weight loss in the TGA curve caused by the endothermic phenomenon started at 614.422 °C, 620.927 °C, 627.884 °C, and 630.992 °C, respectively. Meanwhile, the DSC curve of each β_n (10 K/min, 15 K/min, 20 K/min, and 10 K/min) showed that the endothermic temperature regions were 635.968–684.208 °C, 655.017–726.35 °C, 659.994–776.673 °C, and 664.713–759.575 °C, respectively. Therefore, the temperature regions for the incineration kinetic modeling of Stage 4 should be set as 635.968–900.00 °C. The thermal parameters such as the weight loss ratios and peak temperature at βn (n = a, b, c, d, and e) were simultaneously obtained in Table 3.

Table 3. Thermal characteristics of upstream oily sludge in stages one through four.

Parameters		Stage 1	Stage 2	Stage 3	Stage 4
DSC	Endo/Exo [1]	ND [2]	Exo	Exo	Endo
TGA	Weight Loss	10%	10%	20%	5%
Peak temperature at β_n (K)	$\beta_a = 5$ K/min	ND	562.974	695.604	ND
	$\beta_b = 10$ K/min	ND	584.081	726.649	939.433
	$\beta_c = 15$ K/min	ND	598.507	740.092	965.624
	$\beta_d = 20$ K/min	ND	618.030	742.590	993.691
	$\beta_e = 25$ K/min	ND	623.446	768.950	987.380
E_p (KJ/mol)	$ln(\beta_n/T_p{}^2)–1/T_p$	ND	64.43 ± 5.34	90.71 ± 13.35	102.11 ± 28.93
	R^2	ND	0.9798	0.9389	0.8616

Note: 1. Endothermic (Endo)/Exothermic (Exo); 2. Not Detected.

It was not the weight loss, but rather the conversion ratios of the reactant that were employed in the equations of the incineration kinetics model. Therefore, the instantaneous weight of the samples detected in the TGA curves (Figure 1A–E) could be rearranged as instantaneous conversion ratios at certain operating temperature. The results of the conversion ratios versus operating temperature of stages one through four are shown in Figure 2B–E, respectively. In addition, the relationship between the instantaneous conversion ratio and the operating temperature ($d\alpha/dT$) was the slope or the first-order derivative of each curve, as shown in Figure 2B–E. For instance, when the conversion rate was 55% and the heating rate was five K/min (in Stage 1), $d\alpha_{55\%}/dT$ equaled the slope (K), which is shown in the enlarged area of Figure 2B.

3.1.3. Peak-Thermal Kinetic Analysis of Endothermic/Exothermic Reactions

Based on Equations (1), (3), and (4), the basic differential form for the study of thermal kinetic analysis could also be represented as Equation (9).

$$d\alpha/dt = Ae^{-E_a/RT} \times (1-\alpha)^n \tag{9}$$

taking the quadratic differential on both sides of Equation (9), the following equations were obtained:

$$\begin{aligned} \frac{d}{dt}\left[\frac{d\alpha}{dt}\right] &= A(1-\alpha)^n \frac{de^{-E_a/RT}}{dt} + Ae^{-E_a/RT}\frac{d(1-\alpha)^n}{dt} \\ &= \frac{d\alpha}{dt} \times \frac{E_a}{RT^2} \times \frac{dT}{dt} - Ae^{-E_a/RT} \times n(1-\alpha)^{n-1}\frac{d\alpha}{dt} \\ &= \frac{d\alpha}{dt}\left[\beta E_a/RT^2 - Ae^{-E_a/RT} \times n(1-\alpha)^{n-1}\right] \end{aligned} \tag{10}$$

when the quadratic differential $\frac{d}{dt}\left[\frac{d\alpha}{dt}\right]$ equals zero, which means the maximum or minimum value could be obtained. In DSC curves, the exothermal peak was the maximum value, and the endothermic peak was in response to the minimum value. At the limit value, Equation (10) equals zero, and the peak thermal kinetic equation was expressed as Equation (11):

$$\beta E_p/RT_p{}^2 = Ae^{-E_p/RT_p} \times n\left(1-\alpha_p\right)^{n-1} \tag{11}$$

where E_p and T_p are the exothermal/endothermic peak activation energy and exothermal/endothermic peak operating temperature; and α_p is the exothermal/endothermic peak conversion ratio of the reactant;

Kissinger [36] considered that the formula $n\left(1-\alpha_p\right)^{n-1}$ equaled one, taking the natural logarithm of Equation (11) and rewriting it as Equation (12):

$$ln\frac{\beta}{T_p^2} \cong ln\frac{AR}{E_p} - \frac{E_p}{RT_p};\ ln\frac{\beta_n}{T_{n-p}^2} \sim \frac{1}{T_{n-p}} \tag{12}$$

where T_{n-p} was the peak temperature obtained in the DSC curve, and changed with β_n.

A straight line with slope $-E_p/R$ could be obtained by plotting $ln\frac{\beta_n}{T_{n-p}^2}$ versus $\frac{1}{T_{n-p}}$ at every endothermic or exothermic peak parameter. This method was called as "Kissinger approach" in this study. The results of the linear fittings for the endothermic peak and the two exothermic peaks were shown in Figure 1F and Table 2. The peak activation value (E_p) and the coefficient of determination (R^2) of stages two, three, and four were 64.43 ± 5.34 KJ/mol (0.9798), 90.71 ± 13.35 KJ/mol (0.9389), and 102.11 ± 28.93 KJ/mol (0.8616), respectively.

3.2. The Reasoning Process of the Modeling Method Applied for Oily Sludge Incineration

3.2.1. The Reasoning Process of Differential Methods

By rearranging and taking the natural logarithm in Equation (6), the activation value obtained in the differential method was shown in Equation (13). Friedman [37] considered that the activation value could be solved in spite of both the nth-order reaction rate equation and the pre-exponential factor. At each heating rate, a straight line with slope $-E_a/R$ could be obtained by plotting $ln\beta_n(d\alpha/dT)$ versus $\frac{1}{T}$:

$$\beta \times d\alpha/dT = Aexp(-E_a/RT)f(\alpha) \Rightarrow ln\beta(d\alpha/dT) = lnA + lnf(\alpha) - E_a/RT \tag{13}$$

If the activation value had been solved and the nth-order reaction rate equation was fitted to Equation (3), simultaneously, the intercept $(lnA + lnf(\alpha))$ of the straight lines could be applied to solve the reaction order (n) and pre-exponential factor (A) from the equation in two unknowns established under different heating rates. Based on Friedman methods (Equation 13), another pathway to gain the reaction order (n) and the pre-exponential factor (A) was shown in Equation (14):

$$\beta \times d\alpha/dT = Aexp(-E_a/RT)f(\alpha) \Rightarrow ln\frac{\beta(d\alpha/dT)}{exp(-E_a/RT)} = lnA + nln(1-\alpha) \tag{14}$$

Repeating the method of plotting the straight line by $ln\frac{\beta(d\alpha/dT)}{exp(-E_a/RT)}$ versus $ln(1-\alpha)$, the slope of the line was the reaction order, and the intercept was lnA. In addition, for this study, if all of the activation values obtained by the Friedman methods were shown with a high coefficient of determinations at a variety of heating rates, Equation (14) was fit for the solution of the reaction order and the pre-exponential factor. Otherwise, two intercepts $(lnA + lnf(\alpha))$ in Equation (13) obtained with a higher coefficient of determination at a certain conversation ratio would be applied, and two linear equations in two unknowns were simultaneously established for the solution of the reaction order and the pre-exponential factor.

3.2.2. The Reasoning Process of Integral Methods

Flynn [38] confirmed that the $G(\alpha)$ in the integral method (Equation 8) could be rearranged by the temperature integral ($P(\mu)$). The solution of $P(\mu)$ was shown in Figure S1 (Supplementary Materials).

$$\begin{cases} G(\alpha) = \frac{A}{\beta}\int_{T_o}^{T} e^{-E_a/RT}dT \cong \frac{AE_a}{\beta R}\int_{\infty}^{\mu} -e^{-\mu}\mu^{-2}d\mu \\ \quad\quad = \frac{AE_a}{\beta R}P(\mu); \mu = \frac{E_a}{RT} \\ P(\mu) = \int_{\infty}^{\mu} -e^{-\mu}\mu^{-2}d\mu \Rightarrow \frac{e^{-\mu}}{\mu^2}\left(1 - \frac{2!}{\mu} + \frac{3!}{\mu^2} - \frac{4!}{\mu^3}\cdots\right) \\ \quad\quad = \frac{e^{-\mu}}{\mu^2} \times \sum\limits_{N=1}^{\infty}(-1)^{N-1}\frac{N!}{\mu^{N-1}};\ N \geq 1 \end{cases} \tag{15}$$

where $G(\alpha)$ is the integral equation of $f(\alpha)$, $P(\mu)$ is the temperature integral; and N is positive integer, which is greater or equal to one.

In the integral method, $d\alpha/dt$ disappeared, and the noise effect was avoided. However, the activation energy at certain heating rate equations (Equation (15)) was extremely intractable to acquire, which was attributed to $P(\mu)$. Thus, some methods were provided to simplify the solution of Equation (15). Akahira-Sunose [39] deduced that if the N in Equation (15) was equal to one, the $G(\alpha)$ could be simplified and expressed as:

$$\begin{cases} P(\mu) \approx \frac{e^{-\mu}}{\mu^2}; N = 1 \\ G(\alpha) \approx \frac{AE_a}{\beta R} \times \frac{e^{-\mu}}{\mu^2} = \frac{T^2}{\beta}\frac{AR}{E}e^{-E_a/RT} \\ ln\frac{\beta}{T^2} = ln\frac{AR}{G(\alpha)E} - E_a/RT \end{cases} \tag{16}$$

Equation (16) was similar to Kissinger approach (Equation (12)), and the activation energy could be solved by plotting $ln\frac{\beta}{T^2}$ VS. $\frac{1}{T}$. This type of integral method was utilized in the following sections and was labeled as the Kissinger-Akahira-Sunose method (abbreviated as the KAS (Kissinger-Akahira-Sunose) method) in the following studies. Similarly, if the N in Equation (15) was equal to two, combined with the Doyle approach [33], another type of integral method (Equation (17)) was acquired and labeled as the Flynn–Wall–Ozawa method (abbreviated as the FWO (Flynn–Wall–Ozawa) method) [40]. The slope ($-1.0516\ E_a/R$) of the straight line that was plotted by $ln\beta$ VS. $\frac{1}{T}$ was more convenient to solve the acquired energy:

$$\begin{cases} P(\mu) \approx \frac{e^{-\mu}}{\mu^2}\left(1 - \frac{2!}{\mu}\right) \approx 0.00484e^{-1.0516\mu}; N = 2 \\ \beta \approx \frac{AE}{RG(\alpha)} \times 0.00484e^{-1.0516\mu} \Rightarrow ln\beta \approx ln\frac{AE}{RG(\alpha)} - 5.311 - 1.0516E/RT \end{cases} \tag{17}$$

3.3. The Incineration Kinetic Analysis and Model of Stage One

No endothermic or exothermic phenomenon was obtained in Stage 1. Therefore, the TGA curves were utilized for the incineration kinetic analysis. Both $d\alpha/dT$ and T under various conversion rates α_n (n = 10%, 30%, 50%, 70%, and 90%) and heating rates β_n (n = five K/min, 10 K/min, 15 K/min, 20 K/min, and 25 K/min) were shown in Table 4. Based on equations (13), (16), and (17), the results of linear fitting under the KAS method (Figure 3A), the Friedman method (Figure 3B), and the FWO method (Figure 3C) are shown in Figure 3. The activation energies under various conversion rates are shown in Table 5.

Table 4. Parameters for incineration kinetic modeling in Stage 1.

α_n	β_a = 5 K/min		β_b = 10 K/min		β_c = 15 K/min		β_d = 20 K/min		β_e = 25 K/min	
	$d\alpha/dT$ ($\times 10^{-3}$)	T (K)	$d\alpha/dT$ ($\times 10^{-3}$)	T (K)	$d\alpha/dT$ ($\times 10^{-3}$)	T (K)	$d\alpha/dT$ ($\times 10^{-3}$)	T (K)	$d\alpha/dT$ ($\times 10^{-3}$)	T (K)
$\alpha_{10\%}$	3.26	349.781	3.33	350.605	2.16	358.763	2.65	365.125	2.29	369.640
$\alpha_{30\%}$	4.51	394.986	4.35	409.822	5.24	406.490	3.64	416.323	4.22	421.908
$\alpha_{50\%}$	5.59	435.879	5.19	448.973	3.63	458.567	4.09	470.557	3.66	477.807
$\alpha_{70\%}$	6.90	467.820	6.75	481.817	5.78	501.161	5.77	511.133	5.51	521.753
$\alpha_{90\%}$	7.97	495.421	7.99	509.078	7.42	531.442	7.73	541.839	7.19	553.763

Table 5. The activation energies of Stage 1 obtained under the KAS, FWO, and Friedman methods.

α_n (β_a–β_e)	KAS Method $ln(\beta_n/T^2)$–$1/T$		FWO Method $ln(\beta_n)$–$1/T$		Firedman Method $ln(\beta_n \times d\alpha/dT)$–$1/T$	
	$E_{0-\alpha n}$(KJ/mol)	R^2	$E_{0-\alpha n}$(KJ/mol)	R^2	$E_{0-\alpha n}$(KJ/mol)	R^2
$\alpha_{10\%}$	59.26 ± 15.62	0.8453	65.84 ± 14.79	0.8676	53.56 ± 17.26	0.7572
$\alpha_{30\%}$	56.83 ± 17.05	0.8622	63.51 ± 16.13	0.8818	45.41 ± 23.09	0.7633
$\alpha_{50\%}$	54.53 ± 5.44	0.9738	61.93 ± 5.11	0.9798	47.32 ± 5.83	0.9571
$\alpha_{70\%}$	54.04 ± 5.15	0.9691	60.45 ± 4.83	0.9774	49.06 ± 9.38	0.9558
$\alpha_{90\%}$	53.78 ± 6.24	0.9565	60.24 ± 5.87	0.9685	55.80 ± 7.08	0.9555

In Figure 3 and Table 4, when the conversion ratios are over 50% (Table 4), the corresponding coefficient of determinations under three methods were dramatically higher than 0.9555. Only 5% weight loss was detected in Stage 1 before $\alpha_{50\%}$, which was attributed to the dehydration process of oily sludge. When the conversion ratios exceed 50%, the volatilization (not combustion) process of volatiles was performed and confirmed in the thermal kinetic analysis of Stage 2 (Section 3.3). Compared with the coefficient of determinations obtained in the three methods, the Friedman method (Figure 3B) was more fitting for the kinetic modeling of volatilization. Furthermore, the activation energy changed with the conversion ratios in Stage 1, which means that neither the dehydration process nor the volatilization of volatiles was an elementary reaction [39,40]. For the volatilization process, the average activation energy $\overline{E}_{0\alpha-(50\%-100\%)}$ that was acquired under the Friedman method was 60.87 ± 5.27 KJ/mol. Then, the reaction orders and pre-exponential factors under various heating rates could be solved via Equation (14), as shown in Figure S2 and Table 6. In addition, the average reaction order (or the average pre-exponential factor) was not the arithmetic mean value obtained at a variety of heating rates, but rather the slope (or intercept) of the straight line plotting by the $\overline{ln\frac{\beta n(d\alpha/dT)}{exp(-E_a/RT)}}$ versus $ln(1-\alpha)$.

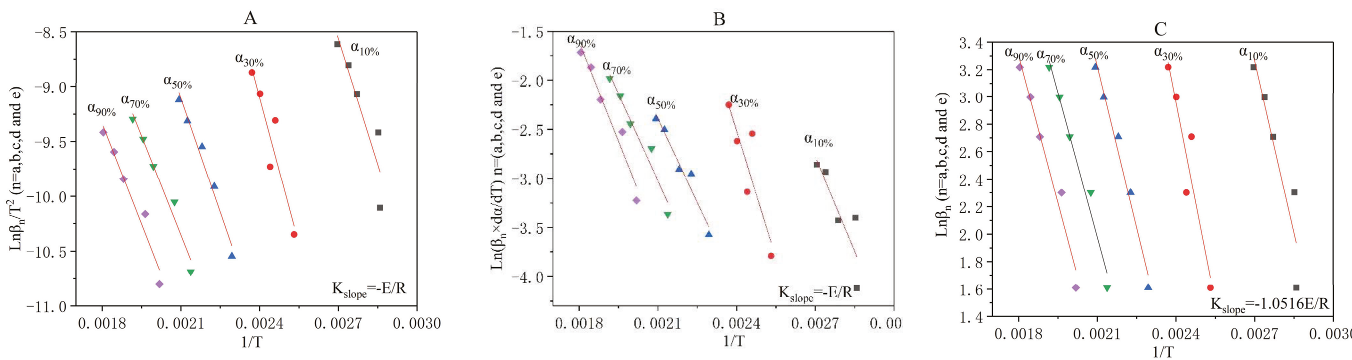


Figure 3. The linear fitting results for Stage 1 under the KAS, Friedman, and FWO methods ((**A**)—KAS method, (**B**)—Friedman method, and (**C**)—FWO method).

Table 6. The reaction order and pre-exponential factor of Stage 1 at a variety of heating rates.

Parameters	Reaction Order (n)	Pre-Exponential Factor ($ln\ A$)	Coefficient of Determination (R^2)
$\beta = 5$ K/min	0.97 ± 0.28	13.72 ± 0.44	0.9236
$\beta = 10$ K/min	0.87 ± 0.25	13.79 ± 0.40	0.9202
$\beta = 15$ K/min	0.85 ± 0.29	13.46 ± 0.45	0.9012
$\beta = 20$ K/min	0.81 ± 0.30	13.42 ± 0.47	0.8766
$\beta = 25$ K/min	0.82 ± 0.30	13.32 ± 0.45	0.8868
Average	0.82 ± 0.30	13.32 ± 0.45	0.9023

The average reaction order was n = 0.82 ± 0.30, and the average pre-exponential factor was lnA = 13.32 ± 0.45. The volatilization kinetic model expressed in differential form and integral form were shown in Equation (18), respectively. Due to the simplification of the temperature integral ($P(\mu)$), the differential form was better to state the volatilization kinetic model in Stage 1.

$$d\alpha/dt = exp(13.32 - 60870/RT)(1-\alpha)^{0.82};\ \alpha \in [0.5,\ 1], T \in (435K,\ 511K) \tag{18}$$

3.4. The Incineration Kinetic Analysis and Model of Stages Two, Three, and Four

Based on Figure 2C–E, the $d\alpha/dT$ and T values for stages two, three, and four under various conversion rates α_n (n = 10%, 30%, 50%, 70%, and 90%) and heating rates β_n (n = five K/min, 10 K/min, 15 K/min, 20 K/min, and 25 K/min) are shown in Table 7.

Table 7. Parameters for incineration kinetic modeling in stages two, three, and four.

α_n	Stage	β_a=5 K/min		β_b=10 K/min		β_c=15 K/min		β_d=20 K/min		β_e=25 K/min	
		$d\alpha/dT$ ($\times 10^{-3}$)	T (K)	$d\alpha/dT$ ($\times 10^{-3}$)	T (K)	$d\alpha/dT$ ($\times 10^{-3}$)	T (K)	$d\alpha/dT$ ($\times 10^{-3}$)	T (K)	$d\alpha/dT$ ($\times 10^{-3}$)	T (K)
	Stage 2	10.57	516.785	10.08	531.892	10.49	554.185	8.89	565.030	9.81	572.626
$\alpha_{10\%}$	Stage 3	3.51	622.887	2.89	639.315	4.39	667.270	4.36	684.588	4.43	694.448
	Stage 4	50.82	902.590	21.16	904.648	8.05	914.979	5.77	923.805	4.72	927.736
	Stage 2	13.60	533.479	13.03	549.453	14.35	570.510	12.99	582.810	12.09	590.517
$\alpha_{30\%}$	Stage 3	9.25	657.685	7.96	682.952	9.50	697.241	8.41	715.626	7.60	727.265
	Stage 4	56.23	907.129	21.91	911.520	12.60	934.999	10.61	948.819	7.82	959.173
	Stage 2	15.53	546.946	15.83	563.042	15.21	583.725	13.97	597.489	13.73	606.171
$\alpha_{50\%}$	Stage 3	13.73	674.935	13.56	701.422	10.61	716.113	10.48	737.392	9.32	750.831
	Stage 4	45.91	911.038	30.19	917.644	14.53	949.547	11.59	965.994	7.69	983.627
	Stage 2	14.17	560.069	15.98	575.404	11.44	598.262	11.76	612.476	10.76	622.225
$\alpha_{70\%}$	Stage 3	16.51	687.783	19.34	714.310	15.14	731.354	11.44	755.153	14.27	769.902
	Stage 4	41.20	914.145	31.44	923.775	12.63	963.750	9.04	983.212	5.73	1015.759
	Stage 2	8.51	577.680	10.3	590.512	6.53	621.731	6.09	635.915	5.94	646.828
$\alpha_{90\%}$	Stage 3	15.38	700.390	27.47	722.863	19.72	743.134	11.03	772.596	22.36	778.701
	Stage 4	25.62	919.106	32.64	929.565	6.11	982.866	0.97	1076.961	2.45	1058.503

The linear fitting results of stage two, three, and four under the KAS method (following Equation (13)), Friedman method (following Equation (16)), and FWO method (following Equation (17)) were shown in Figures 4–6, respectively. The activation energies of each stage obtained under the three methods were shown in Table 8.

Both Stage 2 and Stage 3 were exothermic stages, and the $E_{0-\alpha n}$ (Table 8) values that were separately obtained by the KAS method, Friedman method, and FWO method at α_n (n = 10%, 30%, 50%, 70%, and 90%) were quite different. Meanwhile, $E_{0-\alpha n}$ apparently changed with α_n in each method. In Table 8, the average $\overline{E_0}$ of Stage 2 (Stage 3) obtained under the KAS method, Friedman method, and FWO method were 78.11 KJ/mol (98.82 KJ/mol), 58.07 KJ/mol (77.68 KJ/mol), and 65.63 KJ/mol (81.10 KJ/mol), respectively. In stages two and three (Table 8), the R^2 at a variety of heating rates in the KAS method were all higher than that in the Friedman method and FWO method. Therefore, $\overline{E_0}$ = 78.11 KJ/mol and 98.82 KJ/mol were appropriate for the incineration kinetic modeling of Stage 2 and Stage 3, respectively.

In Stage 2, R^2 was apparently changed with the heating rates, and relatively high R^2 values by Friedman method (Table 8 and Figure 4B) were obtained at $\alpha_{10\%}$ (R^2 = 0.9693) and $\alpha_{30\%}$ (R^2 = 0.9774). The intercepts of $\alpha_{10\%}$ and $\alpha_{30\%}$ linear fitting curves (Figure 4B) plotted by $ln(\beta_n \times d\alpha/dT)$ versus $\frac{1}{T}$

were 19.33 ± 1.24 and 18.45 ± 1.39, respectively. Thus, two linear equations in two unknowns were simultaneously established and expressed in Equation (19):

$$\begin{cases} lnA + nln(1-\alpha_{10\%}) = 19.33 \pm 1.24 \\ lnA + nln(1-\alpha_{30\%}) = 18.45 \pm 1.39 \end{cases} \tag{19}$$

The reaction order and the pre-exponential factor for Stage 2 were $lnA = 19.69$ and $n = 3.50$, respectively. The reaction rate equation of Stage 2 was $f(\alpha) = (1-\alpha)^{3.5}$. The kinetic model for the combustion of volatile components was expressed as:

$$d\alpha/dt = exp(19.69 - 78110/RT)(1-\alpha)^{3.5}; \alpha \in [0,\ 1],\ T \in (511K,\ 658K) \tag{20}$$

Similarly for Stage 3, relatively high R^2 values (Table 8 and Figure 5B) were obtained at $\alpha_{10\%}$ and $\alpha_{50\%}$ by the Friedman method. The $ln(\beta_n \times d\alpha/dT)$ intercepts and R^2 values of the $\alpha_{10\%}$ and $\alpha_{50\%}$ linear fitting curves (Figure 5B) were 20.74 ± 0.74, 0.9942 and 19.24 ± 1.32, 0.9602, respectively. The reaction order and the pre-exponential factor were solved from the followed equations:

$$\begin{cases} lnA + nln(1-\alpha_{10\%}) = 20.74 \pm 0.74 \\ lnA + nln(1-\alpha_{50\%}) = 19.24 \pm 1.32 \end{cases} \tag{21}$$

The reaction order, the pre-exponential factor, and the reaction rate equation for Stage 3 were $lnA = 21.00$, $n = 2.50$, and $f(\alpha) = (1-\alpha)^{2.5}$, respectively. The kinetic model was expressed as:

$$d\alpha/dt = exp(21.00 - 98820/RT)(1-\alpha)^{2.5};\ \alpha \in [0,\ 1],\ T \in (658K,\ 793K) \tag{22}$$

Table 8. The activation energies of stages two, three, and four obtained from the KAS, FWO, and Friedman methods.

Stage	α_n (β_a–β_e)	KAS Method $ln(\beta_n/T^2)$–$1/T$		Friedman Method $ln(\beta_n \times d\alpha/dT)$–$1/T$		FWO Method $ln(\beta_n)$–$1/T$	
		$E_{0-\alpha n}$ (KJ/mol)	R^2	$E_{0-\alpha n}$ (KJ/mol)	R^2	$E_{0-\alpha n}$ (KJ/mol)	R^2
Stage 2	$\alpha_{10\%}$	75.40 ± 5.81	0.9825	61.80 ± 6.34	0.9693	62.84 ± 5.53	0.9772
	$\alpha_{30\%}$	79.10 ± 5.56	0.9854	66.58 ± 5.84	0.9774	69.77 ± 5.31	0.9810
	$\alpha_{50\%}$	80.59 ± 5.88	0.9843	64.64 ± 7.59	0.9602	67.25 ± 5.62	0.9795
	$\alpha_{70\%}$	80.29 ± 6.81	0.9789	54.68 ± 11.97	0.8743	66.74 ± 6.50	0.9723
	$\alpha_{90\%}$	75.16 ± 8.69	0.9614	42.63 ± 15.31	0.7210	61.57 ± 8.26	0.9487
Stage 3	$\alpha_{10\%}$	94.42 ± 8.09	0.9738	92.41 ± 4.07	0.9942	70.83 ± 7.71	0.9654
	$\alpha_{30\%}$	103.31 ± 5.03	0.9930	84.01 ± 10.33	0.9565	87.32 ± 4.85	0.9908
	$\alpha_{50\%}$	100.83 ± 6.18	0.9888	66.66 ± 7.83	0.9602	84.63 ± 5.94	0.9853
	$\alpha_{70\%}$	97.26 ± 6.77	0.9857	67.74 ± 15.01	0.8703	80.98 ± 6.52	0.9808
	$\alpha_{90\%}$	98.28 ± 9.40	0.9732	77.59 ± 35.79	0.6103	81.76 ± 8.99	0.9647
Stage 4	$\alpha_{10\%}$	15.68 ± 0.074	0.9999	−34.35 ± 6.88	0.8925	89.05 ± 56.65	0.4516
	$\alpha_{30\%}$	15.71 ± 0.073	0.9999	179.03 ± 45.85	0.8355	78.65 ± 66.51	0.3179
	$\alpha_{50\%}$	15.40 ± 0.068	0.9999	92.93 ± 60.19	0.4687	103.69 ± 127.91	0.1797
	$\alpha_{70\%}$	16.54 ± 0.197	0.9996	95.77 ± 170.19	0.0955	50.20 ± 33.35	0.4303
	$\alpha_{90\%}$	16.46 ± 0.195	0.9996	−43.57±31.13	0.3949	94.35 ± 14.70	0.9321

The endothermic phenomenon that existed in Stage 4 and the R^2 of the linear fitting curves in the KAS method were dramatically higher than those in the Friedman method and FWO method. Therefore, $\overline{E_0} = 15.96$ KJ/mol was the optimum parameter for the incineration kinetic modeling of Stage 4. As it was shown in Figure 6B, significant errors appeared in the linear fitting curves of Stage 4 under the Friedman method, and a relatively low R^2 value was obtained in each heating rate. Thus, the reaction order and the pre-exponential factor could not be obtained in Equation (13) or Equation (14). We inferred that the probe that was utilized to detect the weight of the reactant was significantly affected by the operating temperatures and caused the apparent errors of $d\alpha/dt$ or $d\alpha/dT$.

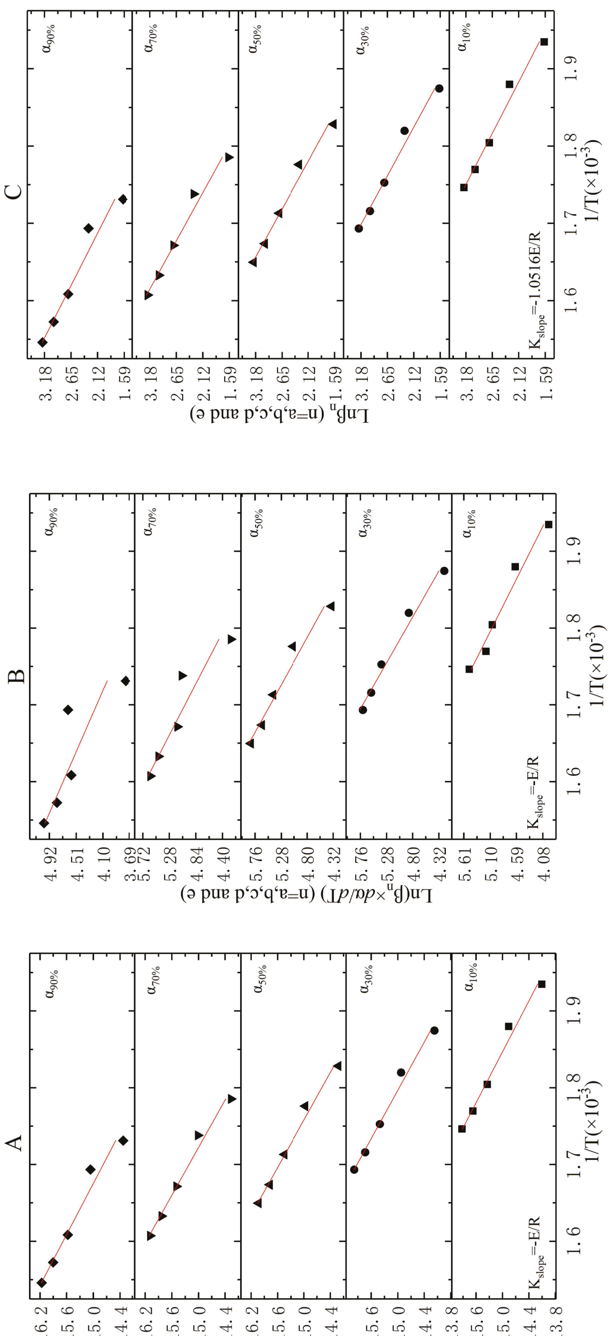


Figure 4. The linear fitting results for Stage 2 under the KAS, Friedman, and FWO methods ((**A**)—KAS method, (**B**)—Friedman method, and (**C**)—FWO method).

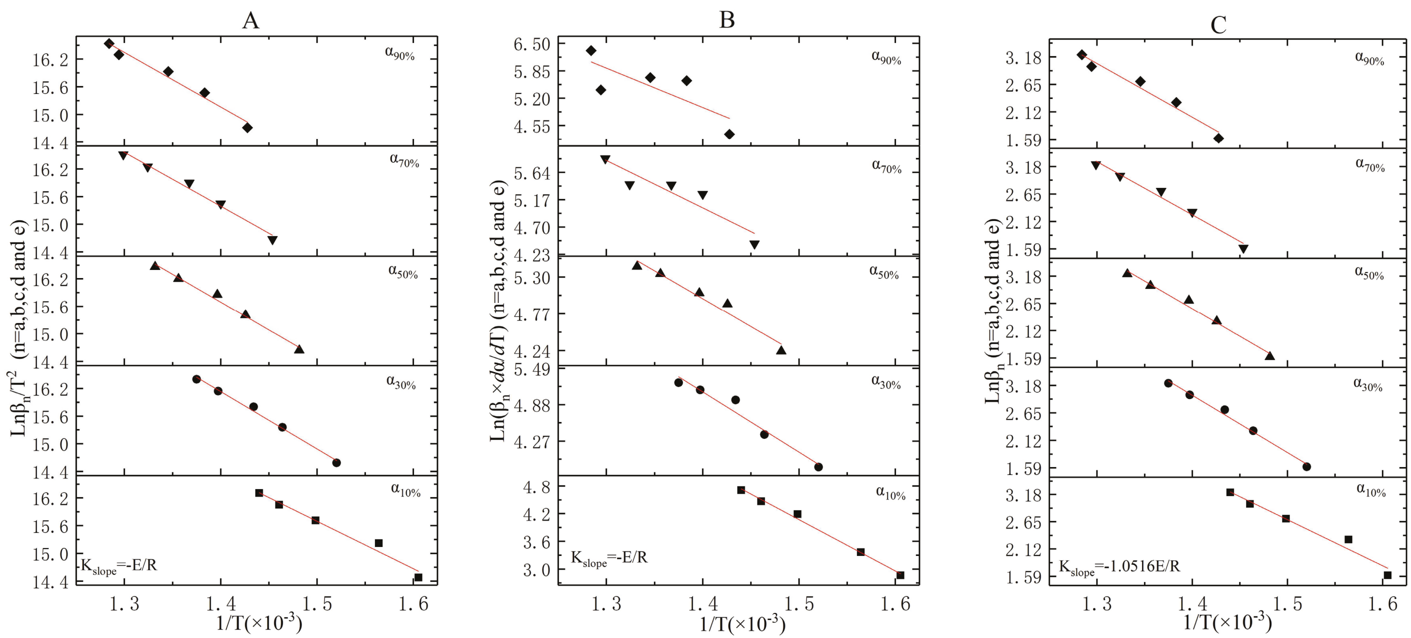


Figure 5. The linear fitting results for Stage 3 under the KAS, Friedman, and FWO methods ((**A**)—KAS method, (**B**)—Friedman method, and (**C**)—FWO method).

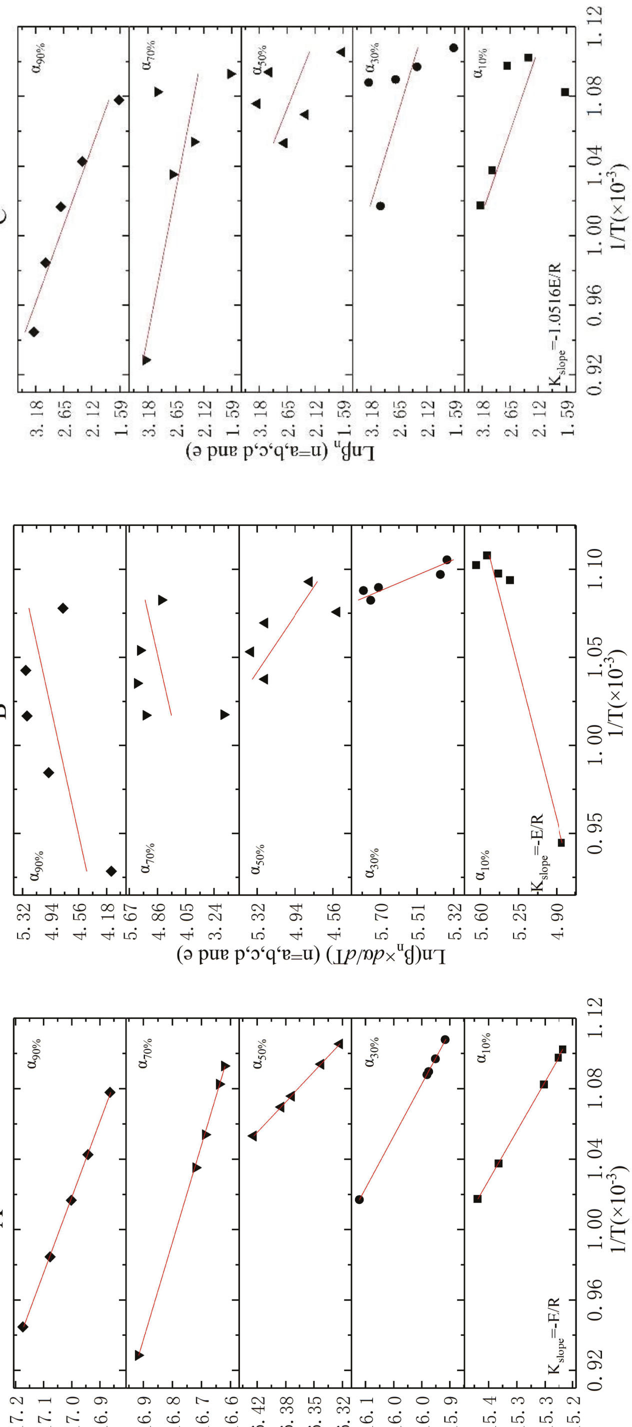


Figure 6. The linear fitting results for Stage 4 under the KAS, Friedman, and FWO methods ((**A**)—KAS method, (**B**)—Friedman method, and (**C**)—FWO method).

3.5. The Judgement of Sectionalized Modeling in Differential/Integral Method

For the incineration kinetics modeling of upstream oily sludge, the E_p values of Stage 2 (or Stage 3, Table 1) obtained under the Kissinger approach were lower than the $\overline{E_0}$ values from the KAS method, which means that the peak operating temperature was more appropriate as the incineration temperature in engineering use. Attributed to the background noise or sensitivity of the probe, the activation energy (E_a) that was obtained in differential methods (the Friedman method) was imprecise, and the value of R^2 also demonstrated that the integral method was more suitable than the differential method.

In comparison with the previous reports in Table 1, the differential method (Friedman method) was more convenient to obtain the pre-exponential factor (A) and the representation of the nth-order reaction rate equation $f(\alpha)$ or the reaction order (n). It seems that the comprehensive differential integral method was more reasonable to solve the basic three elements for the incineration kinetics analysis of upstream oily sludge. However, both the approximate solution of temperature integral $P(\mu)$ that existed in the integral method and the model that was utilized in engineering use should be evaluated and adjusted.

4. Conclusions

Based on whether the endothermic and/or exothermic reactions occurred (or not) in DSC tests, the TGA curves of upstream oily sludge could be sectionalized as four weight loss stages. No endothermic or exothermic reaction was detected, but a 10% weight loss ratio was obtained in Stage 1. The first and second exothermic reaction occurred in stages two (weight loss = 10%) and three (weight loss = 20%), and the ending temperatures were 614.62–664.91 K and 719.15–822.69 K, respectively. The temperature region of Stage 4 was between 635.968–900.00 °C, and the weight loss was 5%.

Five types of thermal reactions existed in the four incineration stages, i.e., dehydration and volatilization (Stage 1), the combustion of light components (Stage 2), the combustion of heavy components (Stage 3), and the decomposition of inorganic carbonate (Stage 4). The two combustion reactions caused the exothermic phenomenon, while the endothermic phenomenon was attributed to the decomposition reaction.

The integral methods (the FWO method and the KAS method) were efficient to obtain the activation energy, while the differential method (the Friedman method) was more suitable to solve the reaction order and pre-exponential factors. The average activation energies of stages one, two, three, and four were 60.87 KJ/mol, 78.11 KJ/mol, 98.82 KJ/mol, and 15.96 KJ/mol, respectively. The reaction order and pre-exponential factors of stages one, two, and three were 0.82, 3.50, and 2.50, and $e^{13.32}min^{-1}$, $e^{19.69}min^{-1}$, and $e^{21.00}min^{-1}$, respectively. Due to the significant errors and relatively low R^2 values that appeared in the linear fitting curves of Stage 4, the reaction order and the pre-exponential factor could not be obtained under the Friedman method.

The E_p values of Stage 2 (or Stage 3, as shown in Table 1) obtained under the Kissinger approach were lower than the $\overline{E_0}$ values from the KAS method, which means that the peak operating temperature was more suitable as the incineration temperature in engineering use.

Supplementary Materials: The following are available online at http://www.mdpi.com/1660-4601/16/3/384/s1, Figure S1: The solution of $P(\mu)$. Figure S2: The linear fitting results for reaction orders and pre-exponential factors of stage 1.

Author Contributions: Conceptualization, Y.Z. and Y.Q.; Data curation, X.W.; Formal analysis, Y.Z. and Y.Q.; Investigation, X.W.; Methodology, Y.Q.; Resources, X.W.; Visualization, F.X.; Writing—original draft, Y.Q.; Writing—review & editing, F.X.

Funding: This research received no external funding.

Acknowledgments: This study was supported by the Zhejiang Heze environmental Tech Shares Co.,LTD. (support number:B220170264). Furthermore, all the authors would like to thank Ding Zhinong, the CEO of Zhejiang Heze environmental Tech Shares Co., LTD, and thankful to Gu Lifeng for the constructive suggestions to this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hu, G.; Li, J.; Zeng, G. Recent development in the treatment of oily sludge from petroleum industry: A review. *J. Hazard. Mater.* **2013**, *261*, 470–490. [CrossRef] [PubMed]
2. Mrayyan, B.; Battikhi, M.N. Biodegradation of total organic carbon (TOC) in Jordanian petroleum sludge. *J. Hazard. Mater.* **2005**, *120*, 127–134. [CrossRef] [PubMed]
3. Al-Futaisi, A.; Jamrah, A.; Yaghi, B.; Taha, R. Assessment of alternative management techniques of tank bottom petroleum sludge in Oman. *J. Hazard. Mater.* **2007**, *141*, 557–564. [CrossRef]
4. Greg, M.H.; Robert, A.H.; Zdenek, D. Paraffinic sludge reduction in crude oil storage tanks through the use of shearing and resuspension. *Acta Montanistica Slovaca* **2004**, *9*, 184–188.
5. Lima, T.M.S.; Fonseca, A.F.; Leão, B.A.; Mounteer, A.H.; Tótola, M.R.; Borges, A.C. Oil recovery from fuel oil storage tank sludge using biosurfactants. *J. Bioremediat. Biodegrad.* **2011**, *2*, 125.
6. Tang, J.; Lu, X.; Sun, Q.; Zhu, W. Aging effect of petroleum hydrocarbons in soil under different attenuation conditions. *Agric. Ecosyst. Environ.* **2012**, *149*, 109–117. [CrossRef]
7. Cuypers, C.; Pancras, T.; Grotenhuis, T.; Rulkens, W. The estimation of PAH bioavailability in contaminated sediments using hydroxypropyl-beta-cyclodextrin and Triton X-100 extraction techniques. *Chemosphere* **2002**, *46*, 1235–1245. [CrossRef]
8. Wang, J.; Yin, J.; Ge, L.; Shao, J.; Zheng, J. Characterization of Oil Sludges from Two Oil Fields in China. *Energy Fuels* **2010**, *24*, 973–978. [CrossRef]
9. Long, X.; Zhang, G.; Shen, C.; Sun, G.; Wang, R.; Yin, L.; Meng, Q. Application of rhamnolipid as a novel biodemulsifier for destabilizing waste crude oil. *Bioresour. Technol.* **2013**, *131*, 1–5. [CrossRef]
10. Gong, B.; Deng, Y.; Yang, Y.; Tan, S.N.; Liu, Q.; Yang, W. Solidification and Biotoxicity Assessment of Thermally Treated Municipal Solid Waste Incineration (MSWI) Fly Ash. *Int. J. Environ. Res. Public Health* **2017**, *14*, 626. [CrossRef]
11. Liu, J.; Jiang, X.; Zhou, L.; Han, X.; Cui, Z. Pyrolysis treatment of oil sludge and model-free kinetics analysis. *J. Hazard. Mater.* **2009**, *161*, 1208–1215. [CrossRef] [PubMed]
12. Shie, J.; Chang, C.; Lin, J.; Wu, C.; Lee, D. Resources recovery of oil sludge by pyrolysis: Kinetics study. *J. Chem. Technol. Biotechnol.* **2000**, *75*, 443–450. [CrossRef]
13. Schmidt, H.; Kaminsky, W. Pyrolysis of oil sludge in a fluidised bed reactor. *Chemosphere* **2001**, *45*, 285–290. [CrossRef]
14. Zhang, Z.Y.; Li, L.H.; Zhang, J.S.; Ma, C.; Wu, X. Solidification of oily sludge. *Pertol. Sci. Technol.* **2018**, *36*, 273–279. [CrossRef]
15. Stasiuk, E.N.; Schramm, L.L. The influence of solvent and demulsifier additions on nascent froth formation during flotation recovery of Bitumen from Athabasca oil sands. *Fuel Process. Technol.* **2001**, *73*, 95–110. [CrossRef]
16. Zubaidy, E.; Abouelnasr, D. Fuel recovery from waste oily sludge using solvent extraction. *Process Saf. Environ. Prot.* **2010**, *88*, 318–326. [CrossRef]
17. Rocha, O.; Dantas, R.; Duarte, M.; Duarte, M.; Silva, V. Oil sludge treatment by photocatalysis applying black and white light. *Chem. Eng. J.* **2010**, *157*, 80–85. [CrossRef]
18. Nii, S.; Kikumoto, S.; Tokuyama, H. Quantitative approach to ultrasonic emulsion separation. *Ultrason. Sonochem.* **2009**, *16*, 145–149. [CrossRef] [PubMed]
19. Christofi, N.; Ivshina, I.B. A review: Microbial surfactants and their use in field studies of soil remediation. *J. Appl. Microbiol.* **2002**, *93*, 915–929. [CrossRef]
20. Lima, T.; Procópio, L.; Brandão, F.; Leão, B.; Tótola, M.; Borges, A. Evaluation of bacterial surfactant toxicity towards petroleum degrading microorganisms. *Bioresour. Technol.* **2011**, *102*, 2957–2964. [CrossRef]
21. Mater, L.; Sperb, R.M.; Madureira, L.; Rosin, A.; Correa, A.; Radetski, C.M. Proposal of a sequential treatment methodology for the safe reuse of oil sludge-contaminated soil. *J. Hazard. Mater. B* **2006**, *136*, 967–971. [CrossRef] [PubMed]
22. Liu, J.; Jiang, X.; Zhou, L.; Wang, H.; Han, X. Co-firing of oil sludge with coal–water slurry in an industrial internal circulating fluidized bed boiler. *J. Hazard. Mater.* **2009**, *167*, 817–823. [CrossRef] [PubMed]
23. Sankaran, S.; Pandey, S.; Sumathy, K. Experimental investigation on waste heat recovery by refinery oil sludge incineration using fluidised-bed technique. *J. Environ. Sci. Health* **1998**, *33*, 829–845. [CrossRef]

24. Scala, F.; Chirone, R. Fluidized bed combustion of alternative solid fuels. *Exp. Therm. Fluid Sci.* **2004**, *28*, 691–699. [CrossRef]
25. Zhou, L.; Jiang, X.; Liu, J. Characteristics of oily sludge combustion in circulating fluidized beds. *J. Hazard. Mater.* **2009**, *170*, 175–179. [CrossRef] [PubMed]
26. Wang, R.; Liu, J.; Gao, F.; Zhou, J.; Cen, K. The slurrying properties of slurry fuels made of petroleum coke and petrochemical sludge. *Fuel Process. Technol.* **2012**, *104*, 57–66. [CrossRef]
27. Yue, Q.; Zhao, Y.; Li, Q.; Li, W.; Gao, B.; Han, S.; Qi, Y.; Yu, H. Research on the characteristics of red mud granular adsorbents (RMGA) for phosphate removal. *J. Hazard. Mater.* **2010**, *176*, 741–748. [CrossRef] [PubMed]
28. Qi, Y.; Dai, B.; He, S.; Wu, S.; Huang, J.; Xi, F. Effect of chemical constituents of oxytetracycline mycelia residue and dredged sediments on characteristics of ultra-lightweight ceramsite. *J. Taiwan Inst. Chem. Eng.* **2016**, *65*, 225–232. [CrossRef]
29. Qi, Y.; Yue, Q.; Han, S.; Yue, M.; Gao, B. Preparation and mechanism of ultra-lightweight ceramics produced from sewage sludge. *J. Hazard. Mater.* **2010**, *176*, 76–84. [CrossRef] [PubMed]
30. Garcia, A.N.; Font, R.; Esperanza, M.M. Thermogravimetric Kinetic Model of the Combustion of a Varnish Waste Based on Polyurethane. *Energy Fuels* **2001**, *15*, 848–855. [CrossRef]
31. Hui, W.; Qing, W.; Jia, H.; Zhan, M.; Jin, S. Study on the Pyrolytic Behaviors and Kinetics of Rigid Polyurethane Foams. *Procedia Eng.* **2013**, *52*, 377–385.
32. Lefebvre, J.; Duquesne, S.; Mamleev, V.; le Bras, M.; Delobel, R. Study of the kinetics of pyrolysis of a rigid polyurethane foam: Use of the invariant kinetics parameters method. *Polym. Adv. Technol.* **2003**, *14*, 796–801. [CrossRef]
33. Jiang, L.; Zhang, D.; Li, M.; He, J.; Gao, Z.; Zhou, Y.; Sun, J. Pyrolytic behavior of waste extruded polystyrene and rigid polyurethane by multi kinetics methods and Py-GC/MS. *Fuel* **2018**, *222*, 11–20. [CrossRef]
34. Friedman, H.L. Kinetics of thermal degradation of char-forming plastics from thermogravimetry. Application to a phenolic plastic. *J. Polym. Sci.* **1965**, *6*, 183–195. [CrossRef]
35. Coat, A.W.; Redfern, J.P. Kinetic Parameters from Thermogravimetric Data. *Nature* **1964**, *201*, 68–69. [CrossRef]
36. Orava, J.; Greer, A.L. Kissinger method applied to the crystallization of glass-forming liquids: Regimes revealed by ultra-fast-heating calorimetry. *Thermochim. Acta* **2015**, *603*, 63–68. [CrossRef]
37. Font, R.; Garrido, M.A. Friedman and n-reaction order methods applied to pine needles and polyurethane thermal decompositions. *Thermochim. Acta* **2018**, *660*, 124–133. [CrossRef]
38. Flynn, J.H. The 'Temperature Integral'—Its use and abuse. *Thermochim. Acta* **1997**, *300*, 83–92. [CrossRef]
39. Lima, A.C.R.; China, B.L.F.; Jawada, Z.A. Kinetic analysis of rice husk pyrolysis using Kissinger-Akahira-Sunose (KAS) method. *Procedia Eng.* **2016**, *148*, 1247–1251. [CrossRef]
40. Opfermann, J.; Kaisersberger, E. An advantageous variant of the Ozawa-Flynn-Wall analysis. *Thermochim. Acta* **1992**, *203*, 167–175. [CrossRef]

International Journal of
Environmental Research and Public Health

MDPI

Review

"Back to the Future": Time for a Renaissance of Public Health Engineering

Richard J. Gelting [1], Steven C. Chapra [2], Paul E. Nevin [3], David E. Harvey [4] and David M. Gute [2,*]

1 Division of Global Health Protection, Center for Global Health, Centers for Disease Control and Prevention, Atlanta, GA 30329 USA; rug7@cdc.gov
2 Department of Civil and Environmental Engineering, Tufts University, Medford, MA 02155, USA; steven.chapra@tufts.edu
3 Department of Global Health, University of Washington, Seattle, WA 98195, USA; penevin@uw.edu
4 Division of Sanitation Facility Construction, Indian Health Service, Rockville, MD 20857, USA; david.harvey@ihs.gov
* Correspondence: david.gute@tufts.edu

Received: 21 December 2018; Accepted: 26 January 2019; Published: 29 January 2019

Abstract: Public health has always been, and remains, an interdisciplinary field, and engineering was closely aligned with public health for many years. Indeed, the branch of engineering that has been known at various times as sanitary engineering, public health engineering, or environmental engineering was integral to the emergence of public health as a distinct discipline. However, in the United States (U.S.) during the 20th century, the academic preparation and practice of this branch of engineering became largely separated from public health. Various factors contributed to this separation, including an evolution in leadership roles within public health; increasing specialization within public health; and the emerging environmental movement, which led to the creation of the U.S. Environmental Protection Agency (EPA), with its emphasis on the natural environment. In this paper, we consider these factors in turn. We also present a case study example of public health engineering in current practice in the U.S. that has had large-scale positive health impacts through improving water and sanitation services in Native American and Alaska Native communities. We also consider briefly how to educate engineers to work in public health in the modern world, and the benefits and challenges associated with that process. We close by discussing the global implications of public health engineering and the need to re-integrate engineering into public health practice and strengthen the connection between the two fields.

Keywords: engineering; public health; curriculum proposals

1. Background

Public health and engineering were closely aligned as professional fields for many years; indeed, engineering was integral in the emergence of public health as a distinct discipline from clinical medicine [1]. However, the practice and profession of environmental engineering have become partially separated from public health in the latter half of the 20th century in the United States (U.S.) [2]. Several factors contributed to this separation, including an evolution in leadership roles within public health; increasing specialization within public health, including the formation of the sanitarian/environmental health field, to which some tasks migrated that were formerly within public health engineering; and the emerging environmental movement, which led to the creation of the U.S. Environmental Protection Agency, with its emphasis on human health and the natural environment. In this paper, we consider these factors in turn, and also highlight an example of public health engineering in current practice in the U.S. We also consider how to educate public health engineers in the modern world, and the global implications of public health engineering.

Engineering is obviously found in other content areas and occupations within public health beyond environmental applications. For example, biomedical engineering is prominent within the Food and Drug Administration (FDA), as is safety engineering and industrial hygiene in the National Institute of Occupational Safety and Health (NIOSH) [3]. In this paper, however, we specifically focus on the aspects of engineering that were most closely aligned with the emergence of public health, and that have been known at various times as sanitary engineering, public health engineering, and environmental engineering. We believe that there is a current need to strengthen the connection between this type of engineering and the practice of public health, and an appropriate moniker for this profession is "public health engineering".

2. The Divergence of Public Health and Engineering in the Latter Half of the 20th Century

Public health has always been, and remains, an interdisciplinary field, but leadership within the field has evolved over time. Starting in the late 19th century, with the emergence of bacteriology and laboratory analysis, a medical perspective became increasingly prominent in public health [4]. This shift in leadership is evident in the U.S., and the evolution of roles is apparent in the history of the U.S. Centers for Disease Control and Prevention (CDC), although this transition within public health clearly started earlier. The first director of the CDC when it initially became a permanent U.S. government agency (the Communicable Disease Center) in 1946, Mark Hollis, ScD (Figure 1), was an engineer [3], a fact little known today, even amongst CDC staff. The agency, which grew out of the Malaria Control in War Areas program created during World War II, was largely staffed by engineers and entomologists during its initial years, reflecting the professions working in the control of mosquito-borne diseases [5]. A review of the history of CDC notes that there were only seven physicians on the staff out of 369 employees when the agency was formed in 1946 [6]. At the time, public health was not always considered the most preferable career option for physicians in the U.S. [7].

Figure 1. Mark Hollis (U.S. Centers for Disease Control, CDC, Public Health Image Library ID# 1304).

As many mosquito-borne diseases were largely controlled in the U.S.—the nation was considered malaria free by 1951 [8]—CDC leaders recognized that the agency's focus needed to expand to include all communicable diseases, and that "To survive, it had to become a center for epidemiology" [5]. This expansion led to the development and refinement of public health surveillance programs [9]. The Epidemic Intelligence Service (EIS) program was started within CDC in 1951, with the stated intent to train epidemiologists to counter the possibility of biological warfare during the Korean War [5]. The program also fulfilled a vision within CDC to create a cadre of public health epidemiologists to investigate disease outbreaks of any kind. The first EIS class of 23 included 22 physicians and a single sanitary engineer [10]. Physicians continue to be prominent within the program, and made up 78% of the nearly 2500 graduates during the first 50 years of the EIS from 1951–2000, while only a handful of engineers participated during this period [10]. The EIS program helped to change negative perceptions among physicians about public health careers, and EIS graduates went on to become leaders and administrators in public health programs, in both the CDC and at the state level. Indeed, since 1953, all CDC directors have been physicians.

At about the same time that physicians were becomingly increasingly prominent in public health in the U.S., further differentiation of other public health professions was also taking place. The term "Sanitarian" was used in the early part of the 20th century as a general descriptor for all public health professionals, including sanitary engineers and physicians working in public health [11–14]. Later, it was sometimes used to describe those working in the field of public health engineering, but without specific engineering education or experience [15]. By the middle of the 20th century, use of the designation had evolved to describe those engaged in more specific environmental health tasks, especially related to ensuring safe food, water, and sanitation [16]. In 1961, the American Public Health Association (APHA) proposed a model act for registering Sanitarians, who were defined as "(persons) who by education and experience in the physical, biological, and sanitary sciences, (are) qualified to carry out educational, investigational and technical duties in the field of sanitation." These tasks were previously considered to be part of the realm of the public health engineer, but were now migrating to the emerging profession of environmental health. More recently, the terms Environmental Health Professional, Environmental Health Scientist or Environmental Health Officer have largely replaced the designation of Sanitarian. Nonetheless, the practice of environmental health has always been closely linked to public health (indeed, the proceedings of some early APHA meetings were published in a journal entitled "The Sanitarian"), whereas the practice of public health engineering has largely become divorced from public health.

While these developments were taking place within the field of public health, another outside force that would also shape the relationship between engineering and public health soon emerged. The publication of the book "Silent Spring" by Rachel Carson in 1962 catalyzed the environmental movement, and raised awareness among both the public and policy makers of the impacts of human action on the environment. The environmental movement incorporated not only the concept of preservation of the natural world, but also enforcement to punish those who pollute it [17]. These attitudes eventually led to the creation of the U.S. Environmental Protection Agency (EPA) in 1970, which consolidated numerous federal programs dealing with environmental pollution issues under one organization. Consequently, most of the engineers working within the U.S. Public Health Service (USPHS) at the time moved to the EPA. Many states mirrored this federal action, moving environmental activities out of health departments [18]. As a result, engineers and engineering as a profession increasingly became aligned with environmental rather than public health issues. This transition also reinforced the emerging use of the term environmental engineer to describe practitioners in the field, replacing the earlier descriptors sanitary engineer and public health engineer.

Although public health engineering may have largely lost recognition as a profession, the practice of it is alive and well in some programs in the U.S. We offer the following case which explores the contributions of the intellectual drivers of public health engineering as it shaped the activities of the USPHS and the Indian Health Service.

3. A Case Study of Public Health Engineering: The Indian Health Service Experience in Native American and Alaska Native Communities

One excellent, but not widely known, example of public health engineering in practice in the U.S. is the Sanitation Facilities Construction (SFC) Program within the Indian Health Service (IHS), and related tribal programs. IHS is an Operating Division within the U.S. Department of Health and Human Services (DHHS), and is responsible for providing federal health services, including public health services, for federally recognized American Indian and Alaska Native (AI/AN) communities.

The Indian Sanitation Facilities Act (Public Law (PL) 86–121, commonly referred to as PL 86–121) was signed into law in 1959 in recognition of the lack of water and sanitation services for tribal communities. Through this law, the U.S. Congress authorized the Surgeon General to construct essential sanitation facilities for Indian homes, communities and lands. This authorization was to help ensure Indian homes and communities would have access to safe drinking water supply and sewage disposal systems which in turn would positively impact health. As a result of this authorization the Division of Sanitation Facilities Construction (DSFC) within the IHS was created. PL 86–121 also directed the Surgeon General of the USPHS to consult with and encourage the participation of tribes impacted in developing sanitation facility projects [19]. This was one of the first federal requirements for tribal participation in Native American programs, and DSFC has worked with tribes since the beginning of the program to identify needs and projects and ensure support for operation and maintenance. Under the 1993 Indian Self-Determination and Education Assistance Act (Public Law 93–638, commonly referred to simply as "638"), 20 tribal governments have taken on some DSFC program functions as of 2015. These functions are still at least partially funded by IHS, and still maintain the principles of a public health based engineering program.

The work of the DSFC has made significant progress since its creation in 1959, increasing the proportion of AI/AN (American Indian/Alaska Native) homes without essential water and sanitation facilities from less than 20% in 1955 to over 90% in 2012 [20]. Nonetheless, significant challenges remain to maintain and advance this progress. Keeping pace with population growth, increasing regulations, and upgrading or replacement of existing facilities when their useful design life is reached are ongoing challenges of the SFC program. As noted by IHS in 2015, safe and adequate water supply and/or waste disposal facilities are lacking in approximately 9 percent of AI/AN homes, which is significantly higher when compared to the overall U.S. population [21]. Uneven distribution among tribal communities is also an issue, with the majority of unserved homes in rural locations on large reservations such as the Navajo Nation in Arizona, New Mexico, and Utah, and in remote Alaskan villages.

The creation and ongoing support of the SFC Program is due in large part to a vision of health within IHS that has integrated prevention principles, including public health engineering, rather than promoting a narrower, more clinically focused view of health. As stated in SFC's annual report, "The IHS considers the provision of sanitation facilities to be a logical extension of its primary health care delivery efforts" [20]. This integration has led to very large scale positive impacts on the health of Native people in the U.S. through prevention programs directed by the DSFC. As stated by Dr. Everett Rhoades, the first Native American Director of IHS (1982–1993), "PL 86–121 had the biggest impact on Indian Health since the smallpox vaccination campaign of the 1830s". The most evident impacts have been on rates of gastrointestinal disease, which have dropped dramatically in the AI/AN population since the creation of the SFC program [20,22,23]. Recent studies in Alaska have also shown that piped water service was associated with fewer respiratory, skin and soft tissue infections and pneumococcal disease among Alaska Natives [23,24]. In addition, there are other ancillary health benefits from the SFC Program that have not been as well documented. For example, anecdotal evidence from clinical practitioners working on the Navajo Nation indicates that elderly tribal members suffer fewer falls when indoor plumbing is installed and residents are not forced to use outdoor privies, especially during harsh and icy winter conditions.

The success of the SFC Program can be attributed to several factors, including the dedication of significant public resources to the program, including nearly U.S. $1 billion in IHS-provided project funding over the 10 years from fiscal years 2007 to 2016, the commitment and hard work of public health engineers and technicians (even if they would not self-identify as such), consistent tribal input to identify needs and appropriate projects to address them, and long-term institutional vision and commitment to improving sanitation conditions in AI/AN communities over decades.

4. Educating Public Health Engineers in the Future

Because engineering and public health have become partially separated as professions and in practice, there is no longer a standardized program of training for public health engineers in the U.S. Few engineers set out initially to work in public health, but that needs to change if we are to effectively address many public health challenges. Engineers working in public health often graduate from "typical" engineering programs, and then gravitate toward public health engineering work either through personal interest or because of unique opportunities. Many engineers come into public health through global work, in which engineering expertise is applied to basic public health problems, such as improving water, sanitation, and hygiene (WASH) conditions in low and middle income countries. Prior work has reviewed the core competencies required for engineers to work in global settings, in addition to making recommendations for educating "globally competent engineers" [2].

In discussing education and curricula for public health engineering, we do not intend to present a comprehensive review of existing programs, but rather some examples of how engineering and public health education can be combined. Neither is our purpose to propose a definitive solution, but rather to pose the question of how best to educate public health engineers. In some fields, such as biomedical and safety engineering, the current form of educating engineers for working in public health fields may be suitable and appropriate. In others, however, a combination of formal education and experience in both engineering and public health are increasingly needed to ensure effective work. We believe this to be the case in the environmental aspects of engineering related to public health, the domain in which we propose to resurrect the term public health engineering.

There are a small number of longstanding graduate programs in the U.S. that combine elements of public health and engineering within schools of public health, including Johns Hopkins University and the University of North Carolina at Chapel Hill. On the other hand, schools of engineering at Tufts University and Columbia University integrate public health into environmental engineering programs. More recently, several programs have more formally integrated elements of public health into schools of engineering, including Stanford University and the Georgia Institute of Technology. This growth has often been at least partially driven by student demand and interest in combining these two fields. That interest is also evident in the growth of Engineers Without Borders (EWB) programs at many U.S. universities. EWB is a voluntary organization of students and professionals that partner with communities to meet basic needs through sustainable engineering projects, a large fraction of which focus on health-related infrastructure.

Although we focus on the U.S. experience in this paper, universities outside the U.S. also have longstanding programs combining engineering and public health. These often focus on water and sanitation issues in lower income countries, and include, for example, programs at the University of Leeds and Loughborough University in the United Kingdom. A more comprehensive review of programs combining public health and engineering, as well as further considerations about potential curricula for public health engineering is beyond the scope of this paper.

In the U.S., we have observed what may be a trend of engineers with an undergraduate engineering degree and a Master of Public Health (MPH) working in the field of public health, although the numbers to date have not been large enough to substantiate any conclusions. According to recent graduates, this combination of academic training is a good one if employers know how to use it to take advantage of and integrate both skill sets.

Whether engineers working in public health obtain their education in both public health and engineering in integrated or separate programs, having such a combined background is optimal to ensuring that they are most effective. A combined background allows them to understand and function across both fields and apply both the engineering and public health skill sets to public health challenges. For example, an engineering education builds critical thinking and problem-solving skills that can be applied to public health issues. Public health education is complementary in that it helps engineers apply those skills in assessing public health impacts, especially in activities involving engineering interventions. Expanding opportunities for students to obtain backgrounds in both fields is also essential for reinvigorating public health engineering as a profession.

5. Conclusions: Re-Integrating Public Health Engineering into Practice

The SFC Program within the Indian Health Service discussed above provides an excellent example of a public health engineering program that has positively impacted public health in the U.S. At the same time, since this successful program is not widely known outside of American Indian and Alaska

Native communities, it also demonstrates the lack of recognition of public health engineering as a profession. Restoring recognition to public health engineering and expanding its practice will help to lay the foundation for large scale public health improvements.

In addition, the preceding case highlights many of the attributes of the public health engineer as seen in practice. These qualities were also championed during the extraordinary career of Abel Wolman (Figure 2) and his advocacy for engineering to be integrated in the practice of public health. Wolman was a prolific writer and thought leader in public health engineering, authoring hundreds of papers over the course of a career spanning eight decades [25]. In his last paper, presented to the 1986 World Health Forum when he was 94, Wolman emphasized the importance of the environment as a determinant of health throughout the world [26]. The scale of global health challenges argue for interdisciplinary approaches to public health practice that are very consistent with Wolman's vision for the application of the engineering skill set to population health challenges around the globe. For example, the case study of the Indian Health Service program discussed in this paper may provide a potential model for improving WASH conditions in other underserved locations, and consequently improving public health. It has been widely recognized that the severity of the cholera epidemic which started in 2010 in Haiti was due to the lack of water and sanitation infrastructure and services in that country [27]. Because the WASH deficiencies in Haiti evolved during decades of limited attention and resources, a long term, sustained public health engineering approach with dedicated resources similar to that of the SFC program will be needed to improve such situations. This would, in turn, require increased development of public health engineers with the necessary skills and experience to manage and sustain such programs. There are some hopeful signs of progress in this area. As mentioned above, university programs incorporating engineering and public health, while still relatively uncommon in the U.S., have expanded in recent years, especially focused on WASH issues in lower income countries. These programs have been met with an overwhelmingly positive response from students, and will help to lay the foundation for a resurgence of engineers working in the public health arena. That will, in turn, underline the need for wider recognition of public health engineering as a profession.

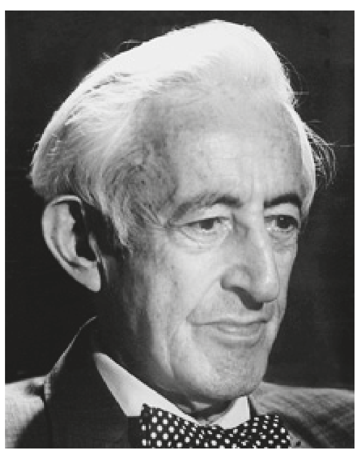


Figure 2. Abel Wolman (Tyler Prize for Environmental Achievement, 2016). (Received consent to publish from Tyler Prize).

In summary, we believe that there is a current need to strengthen the connection between engineering and the practice of public health, and that an appropriate descriptor for this connection is "public health engineering".

Author Contributions: R.J.G. wrote the first draft. R.J.G. and D.M.G. conceived of the organization of the P.E.N. contributed bibliographic references. D.E.H. contributed elements of the Indian Health Service case study. S.C.C. shaped curricular proposals. All authors reviewed and edited the final version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare that they have no competing interests.

Disclaimer: The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

References

1. Gute, D.M. Public Health Engineering: Water as a Focus in Restoring the Connection between Engineering and Public Health. *J. Water Resour. Plan. Manag.* **2004**, *130*, 425–428. [CrossRef]
2. Mihelcic, J.R.; Naughton, C.C.; Verbyla, M.E.; Zhang, Q.; Schweitzer, R.W.; Oakley, S.M.; Wells, E.C.; Whiteford, L.M. The Grandest Challenge of All: The Role of Environmental Engineering to Achieve Sustainability in the World's Developing Regions. *Environ. Eng. Sci.* **2017**, *34*, 16–41. [CrossRef]
3. Earnest, G.S.; Reed, L.D.; Conover, D.; Estill, C.; Gjessing, C.; Gressel, M.; Hall, R.; Hudock, S.; Hudson, H.; Kardous, C.; et al. Engineering and Public Health at CDC. *MMWR Suppl.* **2006**, *55*, 10–13. [PubMed]
4. Fox, D.M. The Professions of Public Health. *Am. J. Public Health* **2001**, *91*, 1362–1364. [CrossRef] [PubMed]
5. Etheridge, E.W. Historical Perspectives: History of the CDC. *Morb. Mortal. Wkly. Rep.* **1996**, *45*, 526–530.
6. CDC. Celebrating 7 Decades of Firsts. Available online: https://www.cdc.gov/museum/history/celebrating-7decades.html (accessed on 10 December 2018).
7. Kelsey, H.J. *"Bright, Aggressive, and Abrasive:" A History of the Chief Epidemic Intelligence Officer of the U.S. Centers for Disease Control and Prevention, 1951–2006*; Georgia State University: Atlanta, GA, USA, 2006.
8. Centers for Disease Control and Prevention (CDC). Malaria—About Malaria-History-CDC's Origins and Malaria. Available online: https://www.cdc.gov/malaria/about/history/history_cdc.html (accessed on 10 December 2018).
9. Thacker, S.B.; Qualters, J.R.; Lee, L.M. Public Health Surveillance in the United States: Evolution and Challenges. *MMWR Suppl.* **2012**, *61*, 3–9. [PubMed]
10. Thacker, S.B.; Dannenberg, A.L.; Hamilton, D.H. Epidemic Intelligence Service of the Centers for Disease Control and Prevention: 50 Years of Training and Service in Applied Epidemiology. *Am. J. Epidemiol.* **2001**, *154*, 985–992. [CrossRef] [PubMed]
11. Bracken, H.M. Executive Sanitarians as Specialists. *Am. J. Public Hyg.* **1906**, *16*, 497–500.
12. Anonymous. The Sanitarian and War. *Am. J. Public Health N. Y. N 1912* **1917**, *7*, 479–480.
13. Anonymous. Sanitarians. *Am. J. Public Health N. Y. N 1912* **1922**, *12*, 422–423.
14. Anonymous. The Recruitment of Sanitarians. *Am. J. Public Health N. Y. N 1912* **1926**, *16*, 508.
15. Kroeze, H.A. The Expanded Role of the Sanitarian. *Am. J. Public Health Nations Health* **1942**, *32*, 611–614. [CrossRef] [PubMed]
16. Light, I.; Pennell, M.; Taylor, D. Sanitarian Manpower. *Public Health Rep.* **1963**, *78*, 793–798. [CrossRef] [PubMed]
17. U.S. Environmental Protection Agency. EPA History: Earth Day. Available online: https://www.epa.gov/history/epa-history-earth-day (accessed on 10 December 2018).
18. Mullan, F. *Plagues and Politics: The Story of the United States Public Service*; Basic Books: New York, NY, USA, 1989.
19. Indian Health Service, Office of Environmental Health and Engineering. *Public Law 86-121 50th Anniversary*; U.S. Department of Health and Human Services, Indian Health Service: Rockville, MD, USA, 2009.
20. Rogers, C.L. *Public Law 86-121 Annual Report for 2012*; U.S. Department of Health and Human Services, Indian Health Service: Rockville, MD, USA; p. 14.
21. Safe Water and Waste Disposal Facilities | Fact Sheets. Available online: https://www.ihs.gov/newsroom/factsheets/safewater/ (accessed on 10 December 2018).
22. Rhoades, E.; Dangelo, A.; Hurlburt, W. The Indian-Health-Service Record of Achievement. *Public Health Rep.* **1987**, *102*, 356–360. [PubMed]

23. Thomas, T.K.; Ritter, T.; Bruden, D.; Bruce, M.; Byrd, K.; Goldberger, R.; Dobson, J.; Hickel, K.; Smith, J.; Hennessy, T. Impact of Providing In-Home Water Service on the Rates of Infectious Diseases: Results from Four Communities in Western Alaska. *J. Water Health* **2016**, *14*, 132–141. [CrossRef] [PubMed]
24. Wenger, J.D.; Zulz, T.; Bruden, D.; Singleton, R.; Bruce, M.G.; Bulkow, L.; Parks, D.; Rudolph, K.; Hurlburt, D.; Ritter, T.; et al. Invasive Pneumococcal Disease in Alaskan Children Impact of the Seven-Valent Pneumococcal Conjugate Vaccine and the Role of Water Supply. *Pediatr. Infect. Dis. J.* **2010**, *29*, 251–256. [CrossRef] [PubMed]
25. Rosen, H. *Who Is Abel Wolman?* APWA (American Public Works Association) Reporter: Kansas City, MO, USA, 1989.
26. Wolman, A. Give Health a Chance—With Healthy Surroundings. *World Health Forum Round* **1986**, *7*, 107–113.
27. Gelting, R.; Bliss, K.; Patrick, M.; Lockhart, G.; Handzel, T. Water, Sanitation and Hygiene in Haiti: Past, Present, and Future. *Am. J. Trop. Med. Hyg.* **2013**, *89*, 665–670. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

Arbuscular Mycorrhizal Fungi Alter Plant and Soil C:N:P Stoichiometries Under Warming and Nitrogen Input in a Semiarid Meadow of China

Linlin Mei, Xue Yang, Hongbing Cao, Tao Zhang * and Jixun Guo *

Institute of Grassland Science, Northeast Normal University, Key Laboratory of Vegetation Ecology, Ministry of Education, Changchun 130024, China; meill641@nenu.edu.cn (L.M.); yangx014@nenu.edu.cn (X.Y.); meilinlinmll@126.com (H.C.)

* Correspondence: zhangt946@nenu.edu.cn (T.Z.); gjixun@nenu.edu.cn (J.G.)

Received: 28 December 2018; Accepted: 29 January 2019; Published: 31 January 2019

Abstract: Ecological stoichiometry has been widely used to determine how plant-soil systems respond to global change and to reveal which factors limit plant growth. Arbuscular mycorrhizal fungi (AMF) can increase plants' uptake of nutrients such as nitrogen (N) and phosphorus (P), thereby altering plant and soil stoichiometries. To understand the regulatory effect of AMF feedback on plants and soil stoichiometry under global change, a microcosm experiment was conducted with warming and N input. The C_4 grass *Setaria viridis*, C_3 grass *Leymus chinensis*, and Chenopodiaceae species *Suaeda corniculata* were studied. The results showed that the mycorrhizal benefits for the C_4 grass *S. viridis* were greater than those for the C_3 grass *L. chinensis*, whereas for the Chenopodiaceae species *S. corniculata*, AMF symbiosis was antagonistic. Under N input and a combination of warming and N input, AMF significantly decreased the N:P ratios of all three species. Under N input, the soil N content and the N:P ratio were decreased significantly in the presence of AMF, whereas the soil C:N ratio was increased. These results showed that AMF can reduce the P limitation caused by N input and improve the efficiency of nutrient utilization, slow the negative influence of global change on plant growth, and promote grassland sustainability.

Keywords: arbuscular mycorrhizal fungi; global change; grassland ecosystem; stoichiometry; phosphorus limitation

1. Introduction

Global climate change, including warming and increased atmospheric nitrogen deposition, has influenced the structure and function of terrestrial ecosystems over the last century [1]. It is predicted that global warming will continue to increase in the next 100 years [2]. Changes in temperature can affect plant productivity and community composition [3,4], e.g., by influencing nutrient demand and uptake [5]. Previous studies have shown that warming may alter carbon (C), nitrogen (N) and phosphorus (P) availability in soil [6] and nutrient cycles [7–9], simultaneously altering the stoichiometry of plants [4,10]. A previous study demonstrated that plant N and P contents declined because of warming, and plants invested fewer nutrients to produce proteins for sustaining biochemical reactions under warming conditions [11]. In addition, studies have shown that warming is likely to accelerate biological processes [12] and enhance decomposition and mineralization rates [13,14], which will enhance the supply rates of N and P to plants [11,15].

Anthropogenic N deposition, i.e., N deposition caused by human activities, is another important threat to ecosystem stability that reduces plant species richness and diversity, increases plant productivity and litter production [16], and affects biogeochemical cycles and net ecosystem C accumulation [17,18]. China is one of the most serious N deposition zones worldwide. Over the past 30 years, the average annual bulk deposition of N increased by approximately 80 kg ha^{-1} y^{-1} N [19]. N input decreases C mineralization, accelerates net N mineralization in soil [20,21], causes N:P imbalances in soil and indirectly affects plant growth [22]. Yue et al. [23] compiled a large number of studies to determine how plant and soil C:N:P stoichiometries respond to individual and combined effects of warming, N input and elevated CO_2 concentrations. The results showed that the effects of N addition on plant C:N:P stoichiometry were stronger than those of warming. Most plants are N limited, but the increasing N input has significantly altered nutrient availability and has resulted in the transformation of N-limited to N and P colimited or P-limited ecosystems [24–26]. Our previous study [9] showed that the addition of N increased the amount of available N and the rate of N mineralization, but reduced soil available P; thus, the soil N:P ratio will decrease under climate change. A high P limitation in saline alkali soil is obvious in the Songnen meadow steppe. Therefore, exploration of the influence of warming and N input on plant and soil stoichiometries of grassland ecosystems, identification of a way to reduce the threat of grassland degradation caused by global change, and determination of a means to increase nutrient use efficiency will be beneficial to maintain sustainable grassland development and reduce environmental impacts.

Arbuscular mycorrhizal fungi (AMF) are among the most important microorganisms in terrestrial ecosystems [27,28]. AMF can form symbiotic relationships with the majority of land plants [29] and can increase plant N and P uptake and obtain C from the host plant, thereby altering the plant C:N:P stoichiometry [30]. Moreover, AMF enhance ecosystem stability and sustainability [31,32].

However, the development and species composition of AMF are often influenced by global change, such as warming and N input [33,34]. In our previous research [35], we found that AMF altered the species composition and productivity under warming and nitrogen (N) addition. However, the mechanism by which AMF affect plant community composition and productivity under warming and N addition is not well understood.

Here, we selected Setaria viridis (*L. Gramineae*) Leymus chinensis (Trin. Tzvel, Poaceae), and Suaeda corniculata (C. A. Mey, Chenopodiaceae), three typical dominant plants in the Songnen meadow steppe, to determine the effects of AMF on the C:N:P stoichiometry of plants and soils under global change. Chenopodiaceae species are highly tolerant to high salinity and drought stress, and an increasing number of studies have found that many chenopods can be colonized by AMF [36–39]. Our previous study showed that AMF increased the relative abundance and aboveground biomass of *S. viridis* and *L. chinensis*, but significantly reduced the relative abundance and aboveground biomass of *S. corniculata* under both warming and N addition [35].

An experiment was conducted using microcosms in a greenhouse. We assessed the total C, N and P contents of plants and soil. We hypothesized that (1) AMF improve C_4 *Setaria viridis* (*L. Gramineae*) and C_3 *Leymus chinensis* (Trin. Tzvel, Poaceae) N and P nutrition and decrease the N and P contents of Suaeda corniculata (C. A. Mey, Chenopodiaceae) under global change; (2) AMF have different contributions to the C:N:P stoichiometry of the three species under global change; and (3) AMF decrease the plant N:P ratio.

2. Materials and Methods

2.1. Plant and Soil Preparation

We established experimental grassland microcosms with typical sodic saline meadow soil from the Songnen Meadow Ecological Research Station (44°45′ N, 123°45′ E), Northeast Normal University, Jilin Province, in northeastern China; this area has a semiarid temperate-zone monsoon climate and typical characteristics of a continental climate. The soil pH was 8.2. The vegetation of the experimental site is dominated primarily by *S. viridis*, *L. chinensis*, and *S. corniculata*. *S. viridis* is an indigenous C_4 grass that associates strongly with AMF. *L. chinensis* is a C_3 grass that gains little benefit from association with AMF [40,41], and *S. corniculata* is a member of the Chenopodiaceae family. The seeds of the three species and soil were collected from the Songnen meadow and were stored in a refrigerator at 4 °C before being used.

Topsoil (0–30 cm) was collected from the same site from which the plant seeds were collected. The soil was sieved (2 mm sieve) to remove large stones and plant roots and was sterilized twice using high-pressure steam at 121 °C for two hours each time to eliminate indigenous AMF.

We collected soil (500 g) from the Songnen meadow where warming and N input had been applied for five years and where Medicago sativa had been cultivated in 200 g of sterilized soil for four months. After the four month period, the aboveground biomass was removed, and the inoculum comprised spores, infected root fragments, hyphae, and soil (approximately 2000 spores per 50 g of soil).

2.2. Experimental Design

This experiment included four global change treatments: Control (CK), warming (W), N input (N), and a combination of warming and N input (WN). Each global change treatment included two AMF treatments: with AMF ('AM') and without AMF ('NM'); each treatment was replicated four times, i.e., a total of 32 microcosms.

Sterilized soil was placed in microcosm pots (2.5 kg of dry soil per pot), and 200 g of AMF inoculum was added to produce the mycorrhizal ('AM') treatment. The same amount of sterilized soil (121 °C for two hours) was placed in the microcosm pots to produce the nonmycorrhizal ('NM') treatment. To ensure the rhizosphere microbial communities in the NM treatment were consistent with those in the AM treatment, a microbial suspension was prepared by sequential filtration of a soil extract obtained by orbital shaking (150 rpm) of a nonsterile soil, sterile dH_2O (1:9 w/v) mixture for 30 min. The AM treatment received 10 ml of deionized water, and the NM treatment received 10 ml of filtrates that were free of mycorrhizal propagules [35].

Before planting, the seeds were surface disinfected with 10% (v/v) hydrogen peroxide for 5 min and washed five times with deionized water. The seeds were allowed to germinate at 20 °C for 48 h. Uniform seedlings were transplanted after 3 days into the grassland microcosms that had a height of 23 cm and a volume of approx. 2.5 kg (based on soil dry weight). The microcosms contained *Setaria viridis* (10 plants per pot), *L. chinensis* (15 plants per pot), and *S. corniculata* (5 plants per pot) [35].

The experiments were performed in phytotrons (LT/ACR-2002, E-Sheng Tech., Beijing, China) from April to August 2014 at Northeast Normal University. The microcosms were placed in phytotrons under a light intensity of 350 $\mu mol^{-2}\ S^{-1}$ (06:00–20:00) and a relative humidity of 40%–60%. The temperature of the microcosms matched the average summer temperature in the Songnen meadow over the last 10 years. The temperature and N input in the phytotrons were based on those described by Zhang et al. [35]. The control and N input treatments were set up as follows: 22 °C from 06:00–10:00, 25 °C from 10:00–15:00, 22 °C from 15:00–20:00, and 22 °C from 20:00–06:00. In the warming and combined warming and N input treatments, the temperature was increased by 3 °C in all time periods relative to the control and N input treatments. The N input and combined warming and N input treatments received N (10 g $m^{-2}\ yr^{-1}$). The soil water content was maintained at 50%–60% of the field capacity by adding water every two days.

After twelve weeks of growth, we harvested the plants. The aboveground samples were cut at the soil surface, removed from the microcosms, rinsed with deionized water, dried at 65 °C for 48 h, and then weighed. The plant roots were collected and washed using deionized water to measure mycorrhizal colonization. To account for soil heterogeneity, soil was collected from three random locations in each microcosm and was sieved using a 2 mm soil sieve. Dried plant leaf and soil samples were milled in a ball mill prior to C, N, and P chemical analysis.

The total C and N contents in the soil and plant leaves were determined using a stable isotope mass spectrometer (Isoprime 100, Isoprime Ltd, Manchester, UK). Soil and plant leaves were digested in sulfuric acid, and then total P contents were determined photometrically using the molybdenum blue ascorbic acid method [42]. The root samples were cut into 1 cm segments and were cleared with

10% (w/v) KOH and then stained with trypan blue at 90 °C for 2 h. AMF colonization was estimated using a previously described method [35].

2.3. Statistical Analysis

All of the data were tested for normality and homogeneity of variance before analysis. All plant and soil data were analyzed using ANOVA with warming, N addition, and benomyl treatment as factors, and their interactions were assessed. In case of significant interactions, means were compared using Tukey's HSD test. These ANOVAs were followed by an individual *t*-tests to detect differences between the NM and AM treatments. All statistical analyses were performed using SPSS software (SPSS 16.0 for Windows, Chicago, IL, USA).

3. Results

3.1. Mycorrhizal Colonization

No colonization was observed in the NM treatments. However, in AM treatments, mycorrhizal colonization of the three plant species varied within the same treatment; specifically, the colonization of *S. viridis* and *L. chinensis* was significantly greater than that of *S. corniculata* ($p < 0.05$) except in the treatment with warming and N input combined (Table 1).

Table 1. The colonization status (%) of arbuscular mycorrhizal fungi (AMF) treatment under warming (W) and N input (N). CK, control; N, N input; W, warming; W × N, combination of warming and N input.

Treatment	*S. viridis*	*L. chinensis*	*S. corniculata*
CK	83.33 ± 10.00^{b}	68.34 ± 1.67^{b}	33.34 ± 3.34^{a}
W	88.33 ± 5.00^{b}	83.33 ± 5.00^{b}	16.67 ± 3.34^{a}
N	65.00 ± 8.33^{b}	77.26 ± 0.59^{b}	38.33 ± 5.00^{a}
W × N	50.00 ± 6.67^{a}	53.50 ± 16.83^{a}	45.61 ± 10.00^{a}

Note: The colonization status with different superscripts (a or b) differ significantly at $p = 0.05$ among different plant species in the same treatment.

3.2. Plant Nutrients and Stoichiometric Ratios

Three-way factorial analyses of variance revealed significant main effects of warming on the C, N, and P contents and stoichiometric ratios of the three plant species except for C and P contents in *S. corniculata* and the C:P ratio in *S. viridis* and *S. corniculata*. Significant main effects of N input on the C, N, and P contents of *S. corniculata*; the N content of *S. viridis*; the C:N and N:P ratios of all three species; and the C:P ratio of *L. chinensis* were detected. Significant interaction effects of W × N on the C, N, and P contents of *S. viridis*; the C:N ratio of *L. chinensis*; the C:P ratios of S. viridis and *L. chinensis*; and the N:P ratios of *L. chinensis* and *S. corniculata* were observed (Table 2).

Table 2. Results of three-way factorial analyses of variance (ANOVAs) of the effects of warming, N input and arbuscular mycorrhizal fungi (AMF) on plant C, N, and P contents and ratios of C:N, C:P, and N:P.

Plant	Variable	W	N	W × N	AMF	AMF × W	AMF × N	AMF × W × N
S. viridis	C	15.38 **	2.583	15.32 **	20.79 ***	2.19	4.65 *	3.41
L. chinensis		12.25 **	2.43	1.29	15.55 **	10.18 **	2.69	2.67
S. corniculata		0.50	11.43 **	4.15	1.92	0.34	2.11	2.33
S. viridis	N	9.76 **	22.99 ***	18.47 ***	0.207	0.01	0.22	0.01
L. chinensis		19.47 ***	0.68	0.49	13.36 **	14.50 **	0.02	0.20
S. corniculata		8.01 **	26.41 ***	0.19	10.91 **	1.54	7.95 **	0.15
S. viridis	P	8.74 **	0.81	15.80 **	16.94 ***	0.99	1.21	2.47
L. chinensis		10.89 **	0.23	0.02	11.77 **	7.85 **	0.96	1.40
S. corniculata		0.80	10.76 **	3.97	2.93	0.02	2.28	3.34 **
S. viridis	C:N	13.78 **	88.52 ***	3.713	77.74 ***	0.29	0.7	0.01

Table 2. *Cont.*

Plant	Variable	W	N	W × N	AMF	AMF × W	AMF × N	AMF × W × N
L. chinensis		212.46 ***	38.73 ***	183.63 ***	3.97	33.64 ***	77.07 ***	1.41
S. corniculata		33.57 ***	64.74 ***	1.41	9.31**	2.34	2.43	0.01
S. viridis	C:P	2.32	1.755	13.28 **	0.501	0.08	13.69 **	0.69
L. chinensis		5.48 *	22.07 ***	26.50 ***	1.01	0.03	4.55 *	2.29
S. corniculata		2.18	0.04	0.34	3.04	6.45*	0.35	6.59 *
S. viridis	N:P	6.77 *	91.98 ***	1.49	88.19 ***	0.89	12.03 **	0.05
L. chinensis		51.63 ***	29.90 ***	25.06 ***	1.85	21.72 ***	38.42 ***	6.92 *
S. corniculata		70.84 ***	134.85 ***	19.23 ***	40.67 ***	16.85 ***	28.08 ***	20.59 ***

* represents a significant difference at $p < 0.05$; ** represents a significant difference at $p < 0.01$; *** represents a significant difference at $p < 0.001$.

AMF increased the C content of *S. viridis* by 119% and 138% ($p < 0.05$; Figure 1A) in the N and WN treatments, respectively, and increased the C content of *L. chinensis* by 282% ($P < 0.01$; Figure 1B) in the N treatment. There were significant effects of AMF × N and AMF × W on C content in *S. viridis* and *L. chinensis*, respectively (Table 2), whereas C contents in *S. corniculata* and N content in *S. viridis* were not affected by AMF (Figure 1C,D; Table 2).

AMF increased the N content of *L. chinensis* by 231% ($p < 0.01$; Figure 1E) in the N treatments and decreased the N content of *S. corniculata* by 52% ($p < 0.001$, Figure 1F) in the WN treatments. There were significant effects of AMF ×W and AMF × N on N content in *L. chinensis* and *S. corniculata*, respectively (Table 2).

AMF increased the P content of *S. viridis* by 54%, 118% ($p < 0.05$), and 115% ($p < 0.05$, Figure 1G) in the W, N, and WN treatments, respectively, and by 299% ($p < 0.01$; Figure 1H) in the *L. chinensis* in the N treatment but decreased the P content in *S. corniculata* by 42% ($p < 0.05$, Figure 1I) in the WN. AMF × W and AMF × W × N had significant effects on the P content in *L. chinensis* and *S. corniculata*, respectively (Table 2).

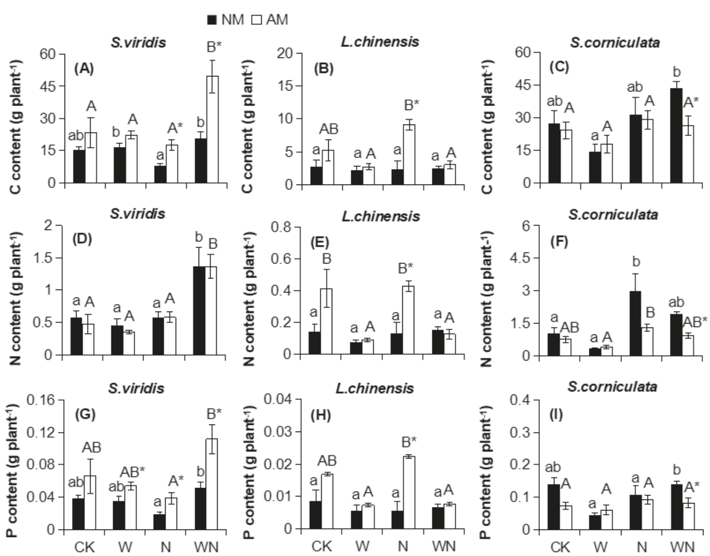


Figure 1. Effects of AMF on the C, N, and P contents of *S. viridis* (**A**,**D**,**G**), *L. chinensis* (**B**,**E**,**H**), and *S. corniculata* (**C**,**F**,**I**) under warming and N input. Different lowercase and capital letters above bars indicate significant differences ($p < 0.05$) among different warming and N input treatments within the same AMF treatment. Asterisks indicate significant differences ($p < 0.05$) between AMF treatments within the same warming and/or N input condition. CK, control; N, N input; W, warming; and WN, combination of warming and N input.

AMF increased the C:N ratio of *S. viridis* by 74% ($p < 0.01$), 57% ($p < 0.05$), 122% ($p < 0.001$), and 121% ($p < 0.01$; Figure 2A) in the control, warming, N input and combination of warming and N input, respectively. AMF decreased the C:N ratio of *L. chinensis* by 36% ($p < 0.001$) in the CK treatment, but in the N and WN treatments, AMF increased the C:N ratio of *L. chinensis* by 17% ($p < 0.01$) and 48% ($p < 0.001$; Figure 2B), respectively. There were significant effects of AMF × W and AMF × N on the C:N ratio of *L. chinensis* (Table 2). AMF significantly increased the C:N ratio of *S. corniculata* by 114% ($p < 0.001$) and 25% ($p < 0.01$; Figure 2C) in the N and WN treatments, respectively.

AMF increased the C:P ratio of *S. viridis* by 11% ($p < 0.05$; Figure 2D) in the WN treatment and that of *S. corniculata* by 18% ($p < 0.05$; Figure 2F) in the CK treatment; it had no impact in the other treatments. There were significant effects of AMF × N on the C:P ratios of *S. viridis* and *L. chinensis* and of AMF × W and AMF × W × N on the C:P ratio of *S. corniculata* (Table 2).

AMF decreased the N:P ratio of *S. viridis* by 51%, 48% ($p < 0.01$), 51% ($p < 0.001$), and 52% ($p < 0.01$; Figure 2G) in the in the control, warming, N input and combination of warming and N input, respectively. AMF increased the N:P ratio of *L. chinensis* by 46% ($p < 0.01$) in the CK treatment but decreased that of *L. chinensis* by 16% ($p < 0.05$) and 27% ($p < 0.01$; Figure 2H) in the N and WN treatments, respectively. AMF decreased the N:P ratio of *S. corniculata* by 49% ($p < 0.001$) and 15% ($p < 0.05$; Figure 2I) in the N and WN treatments, respectively. There were significant effects of AMF × N on the N:P ratio of all three species and of AMF × W and AMF × W × N on the N:P ratio of *L. chinensis* and *S. corniculata* (Table 2).

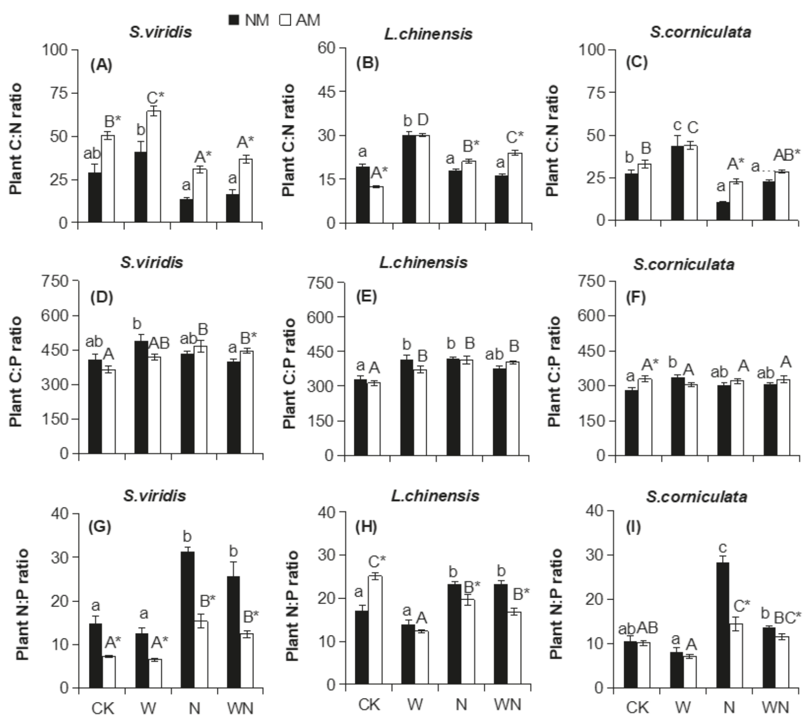


Figure 2. Effects of AMF on the C:N, C:P, and N:P ratios of *S. viridis* (**A**,**D**,**G**, *L. chinensis* (**B**,**E**,**H**) and *S. corniculata* (**C**,**F**,**I**) under warming and N input. Different lowercase and capital letters above bars indicate significant differences ($p < 0.05$) among different warming and N input treatments within the same AMF treatment. Asterisks indicate significant differences ($p < 0.05$) between AMF treatments within the same warming and/or N input condition. CK, control; N, N input; W, warming; and WN, combination of warming and N input.

3.3. Soil Nutrients and Stoichiometric Ratios

Three-way factorial analyses of variance showed significant differences on the effects of warming on soil N and P contents and the C:N ratio; of N input on soil N and P contents and the C:N, C:P and N:P ratios; and of W × N on the C:P ratio (Table 3).

Table 3. Results of three-way factorial analyses of variance (ANOVAs) of the effects of warming, N input and arbuscular mycorrhizal fungi (AMF) on soil total C, N, and P contents and ratios of C:N, C:P, and N:P.

Variable	W	N	W × N	AMF	AMF × W	AMF × N	AMF × W × N
Soil C	0.01	0.20	2.13	2.37	0.07	1.55	0.15
Soil N	33.17 ***	8.50 **	0.22	1.99	10.80 **	0.32	4.04
Soil P	7.88 **	11.05 **	4.04	1.21	11.20 **	8.37 **	2.29
Soil C:N	23.62 ***	6.99 *	2.39	0.66	7.20 *	0.02	2.72
Soil C:P	3.57	6.16 *	5.88*	2.60	10.81 **	8.32 **	1.17
Soil N:P	3.05	17.92 ***	1.58	2.05	20.15 ***	4.95 *	7.35 *

* represents a significant difference at $p < 0.05$; ** represents a significant difference at $p < 0.01$; *** represents a significant difference at $p < 0.001$.

AMF had no effects on soil nutrients and stoichiometric ratios (Table 3); however, under N input, the soil N content and the N:P ratio were reduced by 18% ($p < 0.05$) and 29% ($p < 0.05$; Figure 3B,F), respectively, and the soil C:N ratio was increased by 22% ($p < 0.05$; Figure 3D) in the presence of AMF. In the WN treatment, AMF decreased the soil P content by 22% ($p < 0.05$; Figure 3C) and increased the soil N:P ratio by 44% ($p < 0.05$; Figure 3F). Significant effects of W × AMF on soil N and P contents and the C:N, C:P and N:P ratios; of N × AMF on soil P content and the C:P and N:P ratios; and of W × N × AMF on the soil N:P ratio were observed (Table 3).

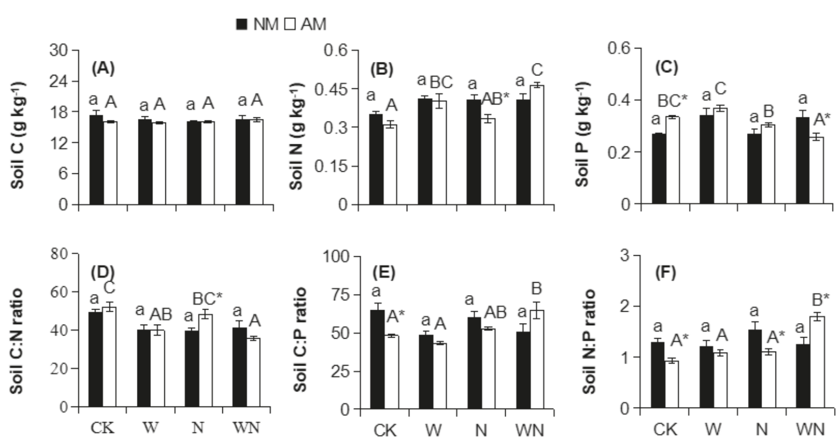


Figure 3. Effects of AMF on soil C (**A**), N (**B**), and P (**C**) contents and C:N (**D**), C:P (**E**), and N:P (**F**) ratios under warming and N input. Different lowercase and capital letters above bars indicate significant differences ($p < 0.05$) among different warming and N input treatments within the same AMF treatment. Asterisks indicate significant differences ($p < 0.05$) between AMF treatments within the same warming and/or N input condition. CK, control; N, N input; and W, warming; WN, combination of warming and N input.

4. Discussion

4.1. Effect of AMF on Plant C:N:P Stoichiometry

Previous studies demonstrated that terrestrial carbon and nitrogen pools can be significantly stimulated by experimental N input [43,44], which could be partly explained by warming-induced increases in net soil N mineralization and nitrification rates [45]. The significant effect of N input on plant C:N and N:P ratios may be attributed to higher soil N availability, which stimulates plant growth [23]. However, our results indicated that AMF alter different plant C:N:P stoichiometries under warming and nitrogen input in a semiarid meadow.

Our previous study has shown that under warming treatment, the mycorrhizal benefits increased by 374.4% for the aboveground biomass of *S. viridis* [35]. In the present study, under warming treatment, AMF significantly increased the P contents and the C:N ratio of *S. viridis*. AMF may increase plant biomass by promoting nutrient cycling [31].

In the N input treatment of our previous study [35], the mycorrhizal benefits conferred to the aboveground biomass of *S. viridis* and *L. chinensis* were increased by 51.1% and 47.4%, respectively, whereas the aboveground biomass of *S. corniculata* decreased significantly under both treatments. These results are generally in agreement with those of the present study. In the present study, under N input and a combination of warming and N input, AMF significantly enhanced the C and P contents and the C:N ratio of the C_4 grass *S. viridis*. Under N input, AMF significantly increased the C, N, and P contents and the C:N ratio of the C_3 grass *L. chinensis.* However, AMF significantly decreased the C, N, and P contents of *S. corniculata* under the combination of warming and N input. The mycorrhizal benefits conferred to C_4 grasses are greater than those conferred to C_3 grasses and should therefore result in higher carbon production and increased AMF development [41,46], an expectation consistent with the findings of the current study. However, in *S. corniculata*, AMF symbiosis was antagonistic. The results suggest that the contribution of AMF to the three dominant species in the Songnen meadow steppe varies under global change. These results support our hypothesis and further suggest that plant stoichiometric responses to global change and ecosystem stability can be adjusted by AMF.

N and P are the most common limiting elements for plant growth and have profound impacts on plant functions [30]. According to some studies, AMF can transfer a considerable amount of N from the soil to host plants [47,48], but other studies have found no evidence that AMF symbioses increase N uptake [49,50]. Phosphorus is the most readily immobilized element in the soil, and its availability is very low [51]; therefore, mycorrhizal P uptake is the dominant pathway [52]. AMF form extensive hyphal networks in the soil and forage effectively for nutrients, especially P, which is supplied to their host plants [10,53]. Fungal nutrient allocation is adjusted through the carbon source strength of individual host plants [54], and plant species affect the AMF response to resource stoichiometry [55].

Furthermore, under all treatments, AMF significantly decreased the N:P ratio of *S. viridis* under N input and a combination of warming and N input, and AMF significantly decreased the N:P ratios of *L. chinensis* and *S. corniculata*. Several studies have suggested that a plant N:P ratio < 14 indicates N limitation and that a plant N:P ratio > 16 indicates P limitation [56,57]. Numerous studies have shown that N input induces an imbalance in the N:P ratio and an increase in P limitation in grasslands [58,59]. In the present study, N input and the combination of warming and N input significantly enhanced the plant N:P ratios of the three species, which may result in an altered balance between N and P. Therefore, the Songnen meadow steppe ecosystem changed from being N limited to being P limited, which is in accordance with an experiment in a temperate steppe ecosystem [60]. However, AMF significantly reduced the N:P ratios of the three species under N input and the combination of warming and N input, which agrees with a previous result [61]. The trade balance model predicts that N enrichment of a P-limited soil will exacerbate the P limitation and increase the amount of P obtained through symbiosis [30]; thus, the plant N:P ratio will decrease in the presence of AMF. The results suggest that AMF might slow the increase in P limitation caused by global change in Songnen meadows. Rational management of soil nutrients in these meadows is critically important to increase plant productivity and to improve the sustainable utilization of grassland ecosystems.

4.2. Effects of AMF on Soil C:N:P Stoichiometry

Stoichiometry is a vital indicator of biogeochemical cycles in terrestrial ecosystems [62]. Soil C:N:P stoichiometry provides a crucial potential diagnostic value for nutrient mineralization and organic matter decomposition [63]. Studies have revealed that C:N:P stoichiometry in soil and plants is tightly linked [64,65]. Soil C:N:P stoichiometry not only regulates microbial activity but also plant N and P uptake [64], while plant C:N:P stoichiometry directly reflects the availability of soil N and P. Many studies have suggested that AMF may transport large numbers of limiting nutrients (N, P) to their host plants from the soil [48,66] and may promote nutrient use efficiency by accelerating the decomposition of organic matter [48,67].

Yue et al. [23] found that high stoichiometric homeostasis, measured as the soil C:N ratio, decreased significantly under N addition. In the present study, under N input, the soil N content and the N:P ratio were decreased significantly in the presence of AMF, but the soil C:N ratio under N input and AMF increased, suggesting that AMF play vital roles in soil nutrient cycling [68], including C and N cycling, in grassland ecosystems [45]. The results indicate that AMF might improve the soil stability and sustainability of plant-soil systems.

5. Conclusions

While N is proposed as the primary limiting nutrient for plant growth, warming and N input resulted in a change from N limitation to P limitation in a semiarid meadow steppe; therefore, P was arguably more limiting in the grassland ecosystem in the study area. Our results showed that AMF play a vital role in maintaining plant nutrient balance and affect plant growth by altering plant nutrient uptake. AMF might slow the rate of P limitation in the Songnen meadow steppe under global change, which would alter plant community composition under future global change. In addition, AMF enhance the soil stability and sustainability of plant-soil systems.

Author Contributions: L.M., T.Z. and J.G. designed the research; L.M., X.Y. and H.C. performed the research; L.M., T.Z. and J.G. wrote and revised the paper.

Funding: This work was supported by the National Natural Science Foundation of China (31470405, 31770359), and the Fundamental Research Funds for the Central Universities (2412018ZD011).

Acknowledgments: We thank the anonymous reviewers for their helpful comments on this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Achat, D.L.; Augusto, L.; Gallet-Budynek, A.D.L. Future challenges in coupled C–N–P cycle models for terrestrial ecosystems under global change: A review. *Biogeochemistry* **2016**, *131*, 173–202. [CrossRef]
2. Peters, G.P.; Andrew, R.M.; Boden, T.; Canadell, J.G.; Ciais, P.; Quéré, C.L.; Marland, G.; Raupach, M.R.; Wilson, C. The challenge to keep global warming below 2 °C. *Nat. Clim. Chang.* **2013**, *3*, 4–6. [CrossRef]
3. Boeck, H.J.D.; Lemmens, C.M.H.M.; Gielen, B.; Bossuyt, H.; Malchair, S.; Carnol, M.; Merckx, R.; Ceulemans, R.; Nijs, I. Combined effects of climate warming and plant d Combined effects of climate warming and plant diversity loss on above- and below-ground grassland productivity. *Environ. Exp. Bot.* **2007**, *60*, 95–104. [CrossRef]
4. Fonty, E.; Sarthou, C.; Larpin, D.; Ponge, J.A. 10-year decrease in plant species richness on a neotropical inselberg: Detrimental effects of global warming? *Glob. Chang Biol.* **2009**, *15*, 2360–2374. [CrossRef]
5. Velthuis, M.; Van Deelen, E.; Van Donk, E.; Zhang, P.; Bakker, ES. Impact of temperature and nutrients on carbon: nutrient tissue stoichiometry of submerged aquatic plants: an experiment and meta-analysis. *Front. Plant Sci.* **2017**, *8*, 655. [CrossRef] [PubMed]

6. Daebeler, A.; Bodelier, P.L.E.; Hefting, M.M.; Rütting, T.; Jia, Z.; Laanbroek, H.J. Soil warming and fertilization altered rates of nitrogen transformation processes and selected for adapted ammonia-oxidizing archaea in sub-arctic grassland soil. *Soil Biol. Biochem.* **2017**, *107*, 114–124. [CrossRef]
7. D'Orangeville, L.; Houle, D.; Côté, B.; Duchesne, L. Soil response to a 3-year increase in temperature and nitrogen deposition measured in a mature boreal forest using ion-exchange membranes. *Environ. Monit. Assess.* **2014**, *186*, 8191–8202. [CrossRef]
8. Huang, J.; Xiong, D.; Liu, X.; Yang, Z.; Xie, J.; Yang, Y. Effects of warming on soil organic carbon mineralization: A review. *Acta Ecol. Sin.* **2017**, *37*, 12–24.
9. Zhang, N.Y.; Guo, R.; Song, P.; Guo, J.X.; Gao, Y.Z. Effects of warming and nitrogen deposition on the coupling mechanism between soil nitrogen and phosphorus in songnen meadow steppe, northeastern China. *Soil Biol. Biochem.* **2013**, *65*, 96–104. [CrossRef]
10. Bucher, K.; Belli, S.; Wunderli-Allenspach, H.; Krämer, S.D. P-Glycoprotein in Proteoliposomes with Low Residual Detergent: The Effects of Cholesterol. *Pharm. Res.* **2007**, *24*, 1993–2004. [CrossRef]
11. Xia, C.X.; Yu, D.; Wang, Z.; Xie, D. Stoichiometry patterns of leaf carbon, nitrogen and phosphorous in aquatic macrophytes in eastern China. *Ecol. Eng.* **2014**, *70*, 406–413. [CrossRef]
12. Schaeffer, S.M.; Sharp, E.; Schimel, J.P.; Welker, J.M. Soil-plant N processes in a high Arctic ecosystem, NW Greenland are altered by long-term experimental warming and higher rainfall. *Glob. Chang. Biol.* **2013**, *19*, 3529–3539. [CrossRef] [PubMed]
13. Jonasson, S.; Castro, J.; Michelsen, A. Litter, warming and plants affect respiration and allocation of soil microbial and plant C, N and P in arctic mesocosms. *Soil Biol. Biochem.* **2004**, *36*, 1129–1139. [CrossRef]
14. Fröberg, M.; Grip, H.; Tipping, E.; Svensson, M.; Strömgren, M.; Dan, B.K. Long-term effects of experimental fertilization and soil warming on dissolved organic matter leaching from a spruce forest in Northern Sweden. *Geoderma* **2013**, *200–201*, 172–179.
15. Butler, S.M.; Melillo, J.M.; Johnson, J.E.; Mohan, J.; Steudler, P.A.; Lux, H.; Burrows, E.; Smith, R.M.; Vario, C.L.; Scott, L.; et al. Soil warming alters nitrogen cycling in a New England forest: Implications for ecosystem function and structure. *Oecologia* **2012**, *168*, 819–828. [CrossRef] [PubMed]
16. Sirulnik, A.G.; Allen, E.B.; Meixner, T.; Fenn, M.E; Allen, M.F. Changes in N cycling and microbial N with elevated N in exotic annual grasslands of southern California. *Appl. Soil Ecol.* **2007**, *36*, 1–9. [CrossRef]
17. Van der Heijden, M.G.A. Mycorrhizal fungi reduce nutrient loss from model grassland ecosystems. *Ecology* **2010**, *91*, 1163–1173. [CrossRef]
18. Johansson, O.; Palmqvist, K.; Olofsson, J. Nitrogen deposition drives lichen community changes through differential species responses. *Glob. Chang. Biol.* **2012**, *18*, 2626–2635. [CrossRef]
19. Liu, X.J.; Zhang, Y.; Han, W.X.; Tang, A.H; Shen, J.L.; Cui, Z.L.; Vitousek, P.; Erisman, J.W.; Goulding, K.; Christie, P.; et al. Enhanced nitrogen deposition over China. *Nature* **2013**, *494*, 459–462. [CrossRef]
20. Lu, M.; Yang, Y.H.; Luo, Y.Q.; Fang, C.M.; Zhou, X.H.; Chen, J.K.; Yang, X.; Li, B. Responses of ecosystem nitrogen cycle to nitrogen addition: A meta-analysis. *New Phytol.* **2010**, *189*, 1040–1050. [CrossRef]
21. Gong, S.; Zhang, T.; Guo, R.; Cao, H.; Shi, L.; Guo, J.; Sun, W. Response of soil enzyme activity to warming and nitrogen addition in a meadow steppe. *Soil Res.* **2015**, *53*, 242–252. [CrossRef]
22. Menge, D.N.L.; Field, C.B. Simulated global changes alter phosphorus demand in annual grassland. *Glob. Chang. Biol.* **2007**, *13*, 2582–2591. [CrossRef]
23. Yue, K.; Fornara, D.A.; Yang, W.; Peng, Y.; Li, Z.; Wu, F.; Peng, C. Effects of three global change drivers on terrestrial C:N:P stoichiometry: a global synthesis. *Glob. Chang. Biol.* **2016**, *23*, 2450. [CrossRef] [PubMed]
24. Sardans, J.; Rivas-Ubach, A.; Peñuelas, J. The C:N:P stoichiometry of organisms and ecosystems in a changing world: A review and perspectives. *Perspect. Plant Ecol. Evolut. Syst.* **2012**, *14*, 33–47. [CrossRef]
25. Camenzind, T.; Hempel, S.; Homeier, J.; Horn, S.; Velescu, A.; Wilcke, W.; Rillig, M.C. Nitrogen and phosphorus additions impact arbuscular mycorrhizal abundance and molecular diversity in a tropical montane forest. *Glob. Chang. Biol.* **2014**, *20*, 3646–3659. [CrossRef] [PubMed]
26. Peñuelas, J.; Sardans, J.; Rivas-Ubach, A.; Janssens, I.A. The human-induced imbalance between C, N and P in Earth's life system. *Glob. Chang. Biol.* **2015**, *18*, 3–6.
27. Verbruggen, E.; Van der Heijden, M.G.A.; Weedon, J.T.; Kowalchuk, G.A.; Röling, W.F.M. Community assembly, species richness and nestedness of arbuscular mycorrhizal fungi in agricultural soils. *Mol. Ecol.* **2012**, *21*, 2341–2353. [CrossRef] [PubMed]

28. Sun, X.F.; Su, Y.Y.; Zhang, Y.; Wu, M.Y.; Zhang, Z.; Pei, K.Q.; Sun, L.; Wan, S.; Liang, Y. Diversity of arbuscular mycorrhizal fungal spore communities and its relations to plants under increased temperature and precipitation in a natural grassland. *Chin. Sci. Bull.* **2013**, *58*, 4109–4119. [CrossRef]
29. Sikes, B.A.; Cottenie, K.; Klironomos, J.N. Plant and fungal identity determines pathogen protection of plant roots by arbuscular mycorrhizas. *J. Ecol.* **2009**, *97*, 1274–1280. [CrossRef]
30. Johnson, N.C. Resource stoichiometry elucidates the structure and function of arbuscular mycorrhizas across scales. *New Phytol.* **2010**, *185*, 631–647. [CrossRef]
31. Bender, S.F.; Conen, F.; Van der Heijden, M.G.A. Mycorrhizal effects on nutrient cycling, nutrient leaching and N_2O production in experimental grassland. *Soil Biol. Biochem.* **2015**, *80*, 283–292. [CrossRef]
32. Veresoglou, S.D.; Shaw, L.J.; Sen, R. *Glomus intraradices* and *Gigaspora margarita* arbuscular mycorrhizal associations differentially affect nitrogen and potassium nutrition of *Plantago lanceolata* in a low fertility dune soil. *Plant Soil* **2011**, *340*, 481–490. [CrossRef]
33. Heinemeyer, A.; Fitter, A.H. Impact of temperature on the arbuscular mycorrhizal (AM) symbiosis: Growth responses of the host plant and its AM fungal partner. *J. Exp. Bot.* **2004**, *55*, 525–534. [CrossRef] [PubMed]
34. Zavalloni, C.; Vicca, S.; Büscher, M.; Providencia, I.E.D.L.; Boulois, H.D.D.; Declerck, S.; Nijs, I.; Ceulemans, R. Exposure to warming and CO_2 enrichment promotes greater above-ground biomass, nitrogen, phosphorus and arbuscular mycorrhizal colonization in newly established grasslands. *Plant Soil* **2012**, *359*, 121–136. [CrossRef]
35. Zhang, T.; Yang, X.; Guo, R.; Guo, J.X. Response of AM fungi spore population to elevated temperature and nitrogen addition and their influence on the plant community composition and productivity. *Sci. Rep.* **2016**, *6*, 24749. [CrossRef] [PubMed]
36. Wilde, P.; Manal, A.; Stodden, M.; Sieverding, E.; Hildebrandt, U.; Bothe, H. Biodiversity of arbuscular mycorrhizal fungi in roots and soils of two salt marshes. *Environ. Microbiol.* **2009**, *11*, 1548–1561. [CrossRef] [PubMed]
37. Aleman, R.; Tiver, F. Endomycorrhizalinfection Levels among Chenopod Plant Species at Port Wakefield, South Australia. *Trans. R. Soc. S. Aust.* **2010**, *134*, 1–4. [CrossRef]
38. Zhang, T.; Shi, N.; Bai, D.; Chen, Y.; Feng, G. Arbuscular Mycorrhizal Fungi Promote the growth of *Ceratocarpus arenarius* (Chenopodiaceae) with no enhancement of phosphorus nutrition. *PLoS ONE* **2012**, *7*, e41151. [CrossRef]
39. Zhao, Y.; Yu, H.; Zhang, T.; Guo, J. Mycorrhizal colonization of chenopods and its influencing factors in different saline habitats, China. *J. Arid Land* **2017**, *9*, 143–152. [CrossRef]
40. Wilson, G.W.T.; Hartnett, D.C. Interspecific variation in plant responses to mycorrhizal colonization in tallgrass prairie. *Am. J. Bot.* **1998**, *85*, 1732–1738. [CrossRef]
41. Grman, E. Plant species differ in their ability to reduce allocation to non-beneficial arbuscular mycorrhizal fungi. *Ecology* **2012**, *93*, 711–718. [CrossRef] [PubMed]
42. Watanabe, F.S.; Olsen, S.R. Test of an ascorbic acid method for determining phosphorus in water and $NaHCO_3$ extracts from soil. *Soil Sci. Soc. Am. J.* **1965**, *29*, 677–678. [CrossRef]
43. Zhao, C.Z.; Liu, Q. Growth and photosynthetic responses of two coniferous species to experimental warming and nitrogen fertilization. *Can. J. For. Res.* **2008**, *39*, 1–11. [CrossRef]
44. Zhang, T.; Yang, S.B.; Guo, R.; Guo, J.X. Warming and nitrogen addition alter photosynthetic pigments, sugars and nutrients in a temperate meadow ecosystem. *PLoS ONE* **2016**, *11*, e0155375.
45. Zhang, Q.; Xie, J.; Lyu, M.; Xiong, D.; Wang, J.; Chen, Y.; Li, Y.; Wang, M.; Yang, Y. Short-term effects of soil warming and nitrogen add. Short-term effects of soil warming and nitrogen addition on the N:P stoichiometry of *Cunninghamia lanceolata* in subtropical regions. *Plant Soil* **2016**, *411*, 395–407. [CrossRef]
46. Verbruggen, E.; Veresoglou, S.D.; Anderson, I.C.; Caruso, T.; Hammer, E.C.; Kohler, J.; Rillig, M.C. Arbuscular mycorrhizal fungi-short-term liability but long-term benefits for soil carbon storage? *New Phytol.* **2013**, *197*, 366–368. [CrossRef]
47. Govindarajulu, M.; Pfeffer, P.E.; Jin, H.; Abubaker, J.; Douds, D.D; Allen, J.W.; Bücking, H.; Lammers, P.J.; Shachar-Hill, Y. Nitrogen transfer in the arbuscular mycorrhizal symbiosis. *Nature* **2005**, *435*, 819–823. [CrossRef]
48. Leigh, J.; Hodge, A.; Fitter, A.H. Arbuscular mycorrhizal fungi can transfer substantial amounts of nitrogen to their host plant from organic material. *New Phytol.* **2009**, *181*, 199–207. [CrossRef]

49. Koide, R.T. Nutrient supply, nutrient demand and plant response to mycorrhizal infection. *New Phytol.* **1991**, *117*, 365–386. [CrossRef]
50. Reynolds, H.L.; Hartley, A.E.; Vogelsang, K.M.; Bever, J.D.; Schultz, P.A. Arbuscular mycorrhizal fungi do not enhance nitrogen acquisition and growth of old-field perennials under low nitrogen supply in glasshouse culture. *New Phytol.* **2005**, *167*, 869–880. [CrossRef]
51. Querejeta, J.I.; Allen, M.F.; Caravaca, F.; Roldán, A. Differential modulation of host plant $\delta^{13}C$ and $\delta^{18}O$ by native and nonnative arbuscular mycorrhizal fungi in a semiarid environment. *New Phytol.* **2006**, *169*, 379–387. [CrossRef] [PubMed]
52. Camenzind, T.; Rillig, M.C. Extraradical arbuscular mycorrhizal fungal hyphae in an organic tropical montane forest soil. *Soil Biol. Biochem.* **2013**, *64*, 96–102. [CrossRef]
53. Van der Berg, L.J.L.; Vergeer, P.; Rich, T.C.G.; Smart, S.M.; Guest, D.; Ashmore, M.R. Direct and indirect effects of nitrogen deposition on species composition change in calcareous grasslands. *Glob. Chang. Biol.* **2010**, *17*, 1871–1883. [CrossRef]
54. Fellbaum, C.R.; Mensah, J.A.; Cloos, A.J.; Strahan, G.E.; Pfeffer, P.E.; Kiers, E.T.; Bücking, H. Fungal nutrient allocation in common mycorrhizal networks is regulated by the carbon source strength of individual host plants. *New Phytol.* **2014**, *203*, 646–656. [CrossRef]
55. Grman, E.; Robinson, T.M. Resource availability and imbalance affect plant-mycorrhizal interactions: A field test of three hypotheses. *Ecology* **2013**, *94*, 62–71. [CrossRef] [PubMed]
56. Koerselman, W.; Meuleman, A.F.M. The vegetation N:P ratio: a new tool to detect the nature of nutrient limitation. *J. Appl. Ecol.* **1996**, *33*, 1441–1450. [CrossRef]
57. Güsewell, S. N:P ratios in terrestrial plants: variation and functional significance. *New Phytol.* **2004**, *164*, 243–266. [CrossRef]
58. Clark, C.M.; Tilman, D. Loss of plant species after chronic low-level nitrogen deposition to prairie grasslands. *Nature* **2008**, *451*, 712–715. [CrossRef]
59. Huang, J.; Yu, H.; Liu, J.; Luo, C.; Sun, Z.; Ma, K.; Kang, Y.; Du, Y. Phosphorus addition changes belowground biomass and C:N:P stoichiometry of two desert steppe plants under simulated N deposition. *Sci. Rep.* **2018**, *8*, 3400. [CrossRef]
60. Zhan, S.X.; Wang, Y.; Zhu, Z.C.; Li, W.H.; Bai, Y.F. Nitrogen enrichment alters plant N:P stoichiometry and intensifies phosphorus limitation in a steppe ecosystem. *Environ. Exp. Bot.* **2017**, *134*, 21–32. [CrossRef]
61. Chen, M.M.; Yin, H.B.; O'Connor, P.; Wang, Y.S.; Zhu, Y.G. C:N:P stoichiometry and specific growth rate of clover colonized by arbuscular mycorrhizal fungi. *Plant Soil* **2010**, *326*, 21–29. [CrossRef]
62. Zhao, F.Z.; Sun, J.; Ren, C.G.; Kang, D.; Deng, J.; Han, X.H.; Gaihe, Y.; Yongzhong, F.; Guangxin, R. Land use change influences soil C, N, and P stoichiometry under 'Grain-to-Green Program' in China. *Sci. Rep.* **2015**, *5*, 10195.
63. Bui, E.N.; Henderson, B.L. C:N:P stoichiometry in Australian soils with respect to vegetation and environmental factors. *Plant Soil* **2013**, *373*, 553–568. [CrossRef]
64. Bell, C.; Carrillo, Y.; Boot, C.M.; Rocca, J.D.; Pendall, E.; Wallenstein, M.D. Rhizosphere stoichiometry: Are C:N:P ratios of plants, soils, and enzymes conserved at the plant species-level? *New Phytol.* **2014**, *201*, 505–517. [CrossRef] [PubMed]
65. Fan, H.; Wu, J.; Liu, W.; Yuan, Y.; Hu, L.; Cai, Q. Linkages of plant and soil C:N:P stoichiometry and their relationships to forest growth in subtropical plantations. *Plant Soil* **2015**, *392*, 127–138. [CrossRef]
66. Hodge, A.; Storer, K. Arbuscular mycorrhiza and nitrogen: implications for individual plants through to ecosystems. *Plant Soil* **2015**, *386*, 1–19. [CrossRef]
67. Cheng, L.; Booker, F.L.; Tu, C.; Burkey, K.O.; Zhou, L.; Shew, H.D.; Rufty, T.W.; Hu, S. Arbuscular mycorrhizal fungi increase organic carbon decomposition under elevated CO_2. *Science* **2012**, *337*, 1084–1087. [CrossRef] [PubMed]
68. Rillig, M.C.; Mummey, D.L. Mycorrhizas and soil structure. *New Phytol.* **2006**, *171*, 41–53. [CrossRef]

International Journal of
Environmental Research and Public Health

MDPI

Article

Evaluating Targeted Intervention on Coal Miners' Unsafe Behavior

Ruipeng Tong *, Yanwei Zhang, Yunyun Yang, Qingli Jia, Xiaofei Ma and Guohua Shao

School of emergency management and safety engineering, China University of Mining and Technology (Beijing), Beijing 100083, China; 13121957417@163.com (Y.Z.); yangyunyun0225@126.com (Y.Y.); jiaqingli123456@126.com (Q.J.); 18811739823@163.com (X.M.); 18810700823@163.com (G.S.)

* Correspondence: tongrp@cumtb.edu.cn; Tel.: +86-137-1843-1777

Received: 25 December 2018; Accepted: 30 January 2019; Published: 1 February 2019

Abstract: Miners' unsafe behavior is the main cause of roof accidents in coal mines, and behavior intervention plays a significant role in reducing the occurrence of miners' unsafe behavior. However, traditional behavior intervention methods lack pertinence. In order to improve the intervention effect and reduce the occurrence of coal mine roof accidents more effectively, this study proposed a targeted intervention method for unsafe behavior. The process of targeted intervention node locating was constructed, and based on the analysis of 331 coal mine roof accidents in China, three kinds of targeted intervention nodes were located. The effectiveness of targeted intervention nodes was evaluated by using structural equation model (SEM) through randomly distributing questionnaires to miners of Pingdingshan coal. The results show that, in preventing roof accidents of coal mines, the targeted intervention nodes have a significant positive impact on the intervention effect. The method can also be applied to the safety management of other industries by adjusting the node location and evaluation process.

Keywords: unsafe behavior; targeted intervention; coal miners; safety management

---

1. Introduction

At present, roof accidents are still frequent in coal mine production. According to a previous analysis of roof accidents in coal mines, engineering technical means cannot completely control the occurrence of roof accidents. However, research on the influence of human behavior is not clear [1–3]. Currently, the research on roof accidents is mainly divided into two categories: management and technology. The former does not point out concrete operation mistakes or management mistakes, and is a barrier to measuring formulation pertinence. The latter is combined with specific coal mining faces or roadways, so the mechanism of roof deformation and failure instability has been studied, and relevant engineering measures have been worked out [4]. However, despite the continuous improvement of the technical level of roof support in China, roof accidents still occur from time to time. This shows that engineering technical means are not the best measures to solve roof accidents. With continuous improvement and a deepening of people's understanding and the research on behavior safety, researchers have gradually discovered that unsafe behavior is a more significant cause of accidents [5]. Similarly, the vast majority of coal mine casualties are caused by unsafe behavior by miners [6]. Unsafe behavior refers to that that may cause casualties, property damage, and environmental damage in violation of rules of operation and safety regulations. In order to effectively prevent roof accidents, human factors must be taken into account. Combined with BBS (Behavior Based Safety) theory, explicit behavioral intervention measures should be put forward [7]. Reducing the incidence of unsafe behavior among workers through behavior intervention is conducive to reducing the risk of enterprise safety, thus enhancing the effectiveness of the enterprise safety management system [8].

A miner's behavior is based on a complex decision-making process, and behavior safety research is mainly based on objective theory and operational condition theory. The emphasis is to identify key unsafe behavior and correct unsafe behavior through intervention. Currently, the research on unsafe behavior intervention mainly explores the intervention countermeasures of unsafe behavior by combining organizational factors with individual factors [9]. Using planning behavior theory, accident cause theory, the structural equation model, and other methods, research has been done on coal mines, buildings, aviation, and other fields, respectively, with respect to safety training, safety culture, performance feedback, material incentives, and other aspects of appropriate intervention measures to reduce the occurrence of unsafe behavior [10–12]. Namian et al. found that the prolonged use of ineffective safety training methods seriously affected the safety in a building. They collected safety training data at the project level to measure workers' risk identification ability and safety risk perception level, and the results provided references for improving safety training work. It shows that traditional safety training methods, to enhance their effectiveness, need to be combined with modern information technology [13]. Kouabenan's research on two French nuclear power plants proves that the safety climate has a certain substantial impact on the promotion of safety management, but it was also found that the encouragement of the direct supervisor in the enterprise is more influential than the view of senior management on safety. This also indicates that more in-depth research on human behavior characteristics is needed in enterprise security management [14]. Warszawska et al. found that a weak safety culture is the main cause of many catastrophic events and that, in order to avoid this situation, enterprise safety culture must be strengthened [15]. These traditional methods of unsafe behavior intervention are mostly based on the observations and records of workers' behaviors, the assessment of behavioral risks to overall conduct, and extensive behavior intervention of workers. In general, traditional methods are only used for the intervention of unsafe behavior itself and lack the in-depth analysis of its internal characteristics and the in-depth study of the root factors leading to unsafe behavior [16]. Therefore, this intervention method lacks pertinence. People's unsafe behavior cannot be fundamentally changed, and the root factors of unsafe behavior will reappear after the intervention [17].

With the continuous expansion of the data volume of security information, analyzing and mining the hidden value of the data have become important in behavior security research [18]. The purpose of this study is to propose a targeted intervention method for unsafe behavior on the basis of behavior safety management. This method pays more attention to the inherent characteristics of workers' unsafe behaviors on the basis of traditional intervention methods through data mining to fully master the risk level, position, behavior individual, behavior trace, behavioral property, time, and type of unsafe action dimension information for worker's unsafe behavior. Statistical analyses and the association rule mining method are applied to analyze various dimensions of information regarding distribution and the inner link between them. According to the internal information of unsafe behaviors, intervention nodes are located. The ultimate goal is to improve the effect of behavioral intervention by formulating corresponding intervention measures for each targeted intervention node in actual safety management work. Through data analysis of coal mine roof accident cases, three types of targeted intervention nodes are identified: the key types of behavior-targeted intervention nodes, the single-dimensional characteristic-targeted intervention nodes, and the multi-dimensional characteristic-targeted intervention nodes. The structural equation model (SEM) was used to evaluate each targeted intervention node to prove the effectiveness and practicability of the method.

2. The Targeted Intervention Nodes of Miners' Unsafe Behavior

The significance of "targeted" has different meanings in different industries, and in medicine, marketing, and economic development, it mainly refers to targeted drugs, targeted marketing, and targeted poverty alleviation. "Targeted" used in the field of behavior safety management refers to the realization of precise positioning, precision intervention, and the accurate management of unsafe behavior of miners based on existing research. The most critical step to achieve targeted intervention is to locate the targeted intervention node through data mining, statistical analysis, and other methods to identify the key behavior of each specialty and the characteristics of the single dimensional, multi-dimensional, and other deep-seated information [19]. Finally, the corresponding intervention measures are made for each intervention node, so as to correct the unsafe behavior of miners and prevent accidents.

The work of coal mine safety management is difficult, and the unsafe behavior of miners is complex and diverse [20]. Traditional behavioral interventions often use methods such as material reward, goal setting and performance feedback, management intervention, and so on. The accuracy of "netting" and "one-size-fits-all" intervention methods is low, and they not only cannot fully capture the unsafe behavior of the site but will also certainly lead to a miner's aversion and a waste of money. At present, with the construction of digital mines, coal mines have realized the overall perception of underground personnel, equipment and facilities, the environment, and other objects, and the accident reports and various rules and regulations have also gradually improved. This can provide a large amount of data support for the location of target nodes. Based on the above analysis, the process of targeted intervention nodes for unsafe behavior of miners was constructed, as shown in Figure 1.

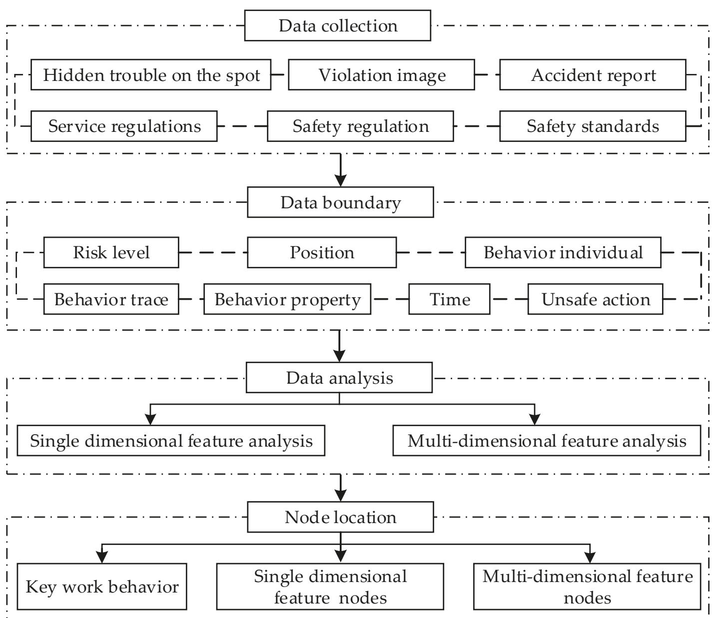


Figure 1. Process of targeted intervention node locating.

2.1. Data Sources

The hidden danger of job sites and violation images can be used as realistic data, recording the occurrence time of the accident, the behavior process of the individual, and the situation of all things, which can clearly and truly reflect the unsafe behavior and the safety risk points of miners [21]. The accident report records the process, the cause, the corrective action, and the result of the accident. Many regulations and standards provide guidance for the standardized description of unsafe behavior, which can be used as abstract data for data analysis.

2.2. Data Boundary

Seven dimensions are used to describe the unsafe behavior comprehensively, and the information of unsafe behavior in the form of text and pictures is transformed into structured data by coding. Risk level (RL) indicates that the severity of unsafe behavior is divided into three levels: high-risk, medium-risk, and low-risk; position (P) describes the location where unsafe behavior occurs; behavior individual (BI) expresses the information of the person who issued the unsafe behavior, which includes age, length of service, educational background, and so on; behavior trace (BT) shows whether the unsafe behavior can be traced after its occurrence and whether the behavior can be divided into traced unsafe behavior and non-traced unsafe behavior; behavior property (BP) is divided into four parts: violation of command, violation of operation, violation of action, and non-violation unsafe action; time (T) records when unsafe behavior occurs; unsafe action (UA) describes the specific unsafe behavior that may lead to accidents, casualties, and environmental disruption.

2.3. Data Analysis

Through the single-dimensional analysis of unsafe behavior data, the distribution characteristics of risk level, position, behavior individual, behavior trace, behavior property, time, unsafe action, and specialty have been explored. Through the interaction analysis of each dimension, the interaction rules of unsafe behavior between different dimensions are explored. Data analysis provides sufficient evidence for locating the targeted intervention nodes.

2.4. Locating Nodes

Firstly, the key behavior of each position should be determined. The targeted intervention nodes oriented to the key behavior of different types of work will be located. Secondly, the distribution differences of miners' unsafe behavior in one particular dimension and different categories will be studied, and the distribution characteristics in the specific dimension are obtained, so as to locate the targeted intervention node oriented to the single-dimensional feature of unsafe behavior. The occurrence of miners' unsafe behavior has its inherent complexity, which is influenced by the natural environment, geological conditions, construction technology, personnel characteristics, management level, and so on [22]. The interaction between variables in theory will be combined with the requirement of practical security management. The interaction between different dimensions of unsafe behavior is exploring, and the potential characteristics and deep regularity of unsafe behavior should be found. Finally, targeted intervention nodes oriented to multi-dimensional association will be located.

3. Research Methods

3.1. Unsafe Behavior Data Analysis Based on Roof Accidents

The number of roof accidents is very large, it is very difficult to integrate the detailed case data of each accident. In order to ensure the scientific nature of the study, the selection of cases follows two principles: accident integrity and case authority. The main sources of accident cases are Internet, accident analysis report, coal mine typical accident compilation, and so on. The application of data analysis methods has the following advantages: the large amount of data, rather than the selection of random samples, greatly reduces the impact of random events on the overall conclusion; the intrinsic characteristics associated with behavior, not just superficial causality, can be studied; the processing speed is fast, and valuable information can be quickly obtained in a short period of time [23].

Unsafe behavior refers to the behavior of the person who has caused the accident or may cause the accident. There are many causes of unsafe behaviors, including individual factors, psychological factors, organizational factors, environmental factors, and so on. Different researchers have different views on the classification of unsafe behaviors. Unsafe behaviors in a coal mine mainly refer to the "three disobeying" of coal mine safety production [24]. "Three disobeying" is the general term of disobeying command, disobeying operation, and disobeying labor discipline. The elimination of "three disobeying" has always been an important issue in the safety management of all industries, especially coal mining enterprises and other high-risk industries. In order to fundamentally explore the objective rule of unsafe behaviors and the complex relationship between its internal factors, many unsafe behaviors need to be objectively analyzed. Previous coal mine accidents can provide enough data for the analysis of miners' unsafe behaviors, but the unsafe behaviors that lead to different types of accidents are different in nature. To eliminate the heterogeneity of a large number of unsafe behavior data, we selected unsafe behavior leading to roof accidents in coal mines as the research object [25].

A total of 331 roof accidents were collected from 1983 to 2014 in China, including 8 major accidents, 159 major accidents, and 164 general accidents. All unsafe behaviors of miners causing roof accidents are collected from the accident report. For example, a roof accident occurred in a mine in Pingdingshan city, Henan province, which directly resulted in a death, several serious injuries, and a total economic loss of about RMB 585,000. According to the time of this accident, four unsafe behaviors could be analyzed successively, expressed as follows: "the top plate was out of the slag and was not withdrawn in time," related to the coal miner; "there are no timely measures to prevent potential safety hazards," related to the field commanders; "there are no perfect operation procedures or safety technical measures established," related to the middle management staff; and "the arrangement of roadway is in a steep-inclined coal seam," related to the senior management staff.

According to the data boundary divided by seven dimensions, 1215 data points were obtained. The data of miners' unsafe behavior was transformed from unstructured text records into structured data. A preliminary single-dimensional statistical analysis of the data points was performed, as shown in Table 1.

Table 1. Statistics of a single dimension.

Dimensional	Result (Frequency, Scale)
RL	High-risk (456, 37.5%); medium-risk (539, 44.4%); low-risk (220, 18.1%)
P	Coal face (504, 41.5%); tunneling working site (333, 27.4%); main roadway (198, 16.3%); others (180, 14.8%)
UA	Supporting (190, 18.36%); safety inspection (119, 11.50%); general type (114, 11.01%)
BI	Coal mining worker (628, 51.7%); field commanders (213, 17.5%); middle management staff (255, 21.0%); senior management staff (119, 9.8%)
BT	Traced behavior (574, 47.2%); non-traced behavior (641, 52.8%)
BP	Violation of action (706, 62.5%); violation of operation (316, 26.0%); violation of command (70, 1.4%); non-violation unsafe action (123, 10.1%)
T	More unsafe actions occur in January, March, and August.

Notes: RL: Risk Level; P: Position; UA: Unsafe action; BI: Behavior individual; BT: Behavior trace; BP: Behavior property; T: time.

According to the description of seven dimensions of miners' unsafe behaviors, interactive analyses can be performed on any two or more of them. The interaction analysis between different variables can map out different practical meanings, the purpose of which is to explore the deep regularity of the unsafe behavior data of miners and to provide the basis for determining the targeted intervention nodes. Through the Apriori algorithm [26], unsafe action is the consequent of the study, and other dimensions are antecedents. The relationship between information and unsafe action in each dimension is discussed (the association between specific unsafe action and time dimension is not obvious enough and is not considered for the time being), as shown in Table 2.

Table 2. Multidimensional interaction relationship analysis.

Dimension	Antecedent	Consequent	Association Rule
RL	High-risk	Empty roof operation	Different risk levels correspond to the most frequent miners' unsafe action
	Medium-risk	No safety measures	
	Low-risk	Failure to clean up float coal in time	
P	Coal face	Inadequate supporting	Different workplaces correspond to the most frequent miners' unsafe action
	Tunneling working site	Empty roof operation	
	Main roadway	Failure to strengthen roadway support	
BI	Coal mining worker	Empty roof operation	Different behavior individuals correspond to the most frequent miners' unsafe action
	Field commanders	Inadequate supporting	
	Middle management staff	No site supervision of the workers' work	
	Senior management staff	Illegal organization of production	
BT	Traced behavior	Inadequate supporting	The most frequent occurrence of the traced unsafe behavior is failure to support in time
	Non-traced behavior	The surrounding environment was not checked before the operation	The most frequent occurrence of the traced unsafe behavior is failure to check the surrounding environment before the operation
BP	Violation of action	Empty roof operation	Different behavior properties correspond to the most frequent miners' unsafe action
	Violation of operation	Inadequate supporting	
	Violation of command	Illegal organization of production	
	Non-violation unsafe action	Nocking and drumming before the work are not careful	

Support is the ratio of the number of consequents and antecedents to the total data set in the database, and the probability of their occurrence is determined. Confidence is the ratio of the support for the occurrence of a consequent and an antecedent to the support for the consequent. It is used to denote the probability of an antecedent derived from an association rule under the condition of a consequent. Lift is the ratio of the confidence of "antecedent $\rightarrow$ consequent" to the support of the consequent. It reflects the size of the consequent influenced by the antecedent [27]. Minimum support and minimum confidence were 8% and 30%, respectively, and this information was used to obtain effective strong association rules. The results of the analysis are shown in Figure 2.

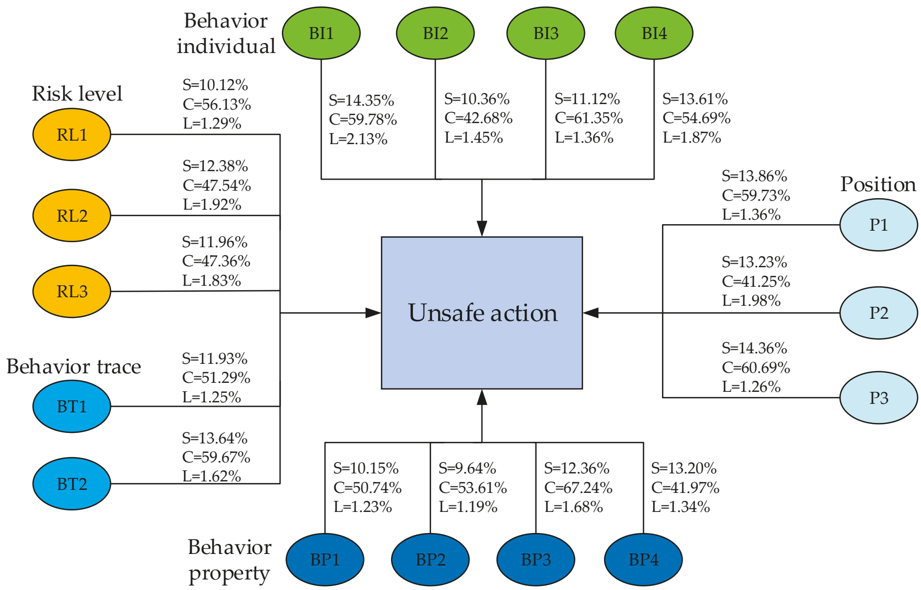


Figure 2. Association rule results of the miners' unsafe behavior.

3.2. The Miners' Unsafe Behavior Targeted Intervention Node Establishment

The key work behaviors are determined by the frequency, work position, work time, risk level, etc. caused by the "three disobeying" in different types of work. Following the principles of reliability, measurability, controllability, observability, etc., according to the comprehensive distribution characteristics of each dimension of "three disobeying" for various types of work, several "three disobeying" behaviors are selected as the key behaviors of various types of work. These key behaviors are targeted at the first type of intervention nodes for unsafe behavior intervention. In the process of practical safety management, it is expected to improve the effect of safety management by focusing on this kind of key work behavior. It is expected that behavior correction can be achieved by interfering with these key work behaviors, and the incidence of unsafe behavior can be greatly reduced [28].

A single-dimensional feature-targeted intervention node is based on the distribution characteristics and proportion of each dimension to guide the intensity of intervention and the allocation of management resources in the process of security management. In terms of time, for example, more unsafe behavior happened in January, March, and August. Safety management can target these months to strengthen the observation and intervention of unsafe behavior and can directly and effectively reduce the number of unsafe behaviors. Based on the association analysis of each dimension of unsafe behavior, the multi-dimensional characteristic-targeted intervention node was located according to their deep connection and rule, which can provide guidance and direction for practical safety management. A list of three types of intervention nodes based on miner's unsafe behavior-targeted intervention was finally formed through data and theoretical analysis for use in this study, as shown in Table 3.

Table 3. List of miners' unsafe behavior-targeted intervention nodes.

Types	Lable	Contents	
Key Work Behavior K	K1	Coal mining worker	Work under the empty roof
	K2	Safety inspection worker	No inspection of the working surface roof safety condition before the operation
	K3	Survey worker	The geological data around the coal mine were not updated in time
	K4	Field commander	No inspection of the work on the working field
	K5	Middle management staff	Safety technical measures were not made according to the situation of the working field before operation
	K6	Senior management staff	Illegal organization of production
Single-Dimensional Feature S	S1	Medium-risk > High-risk > Low-risk	Rational allocation of management resources according to the frequency of different risk behaviors
	S2	Coal face > Tunneling working site > main roadway > Others	Focus on observing coal face, tunneling working site, and roadway in daily behavior safety management
	S3	Supporting > Safety inspection > General type	Increasing the intervention of supporting and safety inspection work in daily behavior safety management
	S4	Coal mining worker > Middle management staff > Field commanders > Senior management staff	Rational allocation of safety training resources according to the frequency of different workers
	S5	Non-traced behavior > Traced behavior	More attention paid to the observation of non-traced behavior in daily safety management
	S6	Violation of action > Violation of operation > Non-violation unsafe action > Violation of command	Rational allocation of safety training and management resources according to the frequency of different workers
	S7	January, March, August	More attention paid in January, March, and August to safety management work
Multi-Dimensional Feature M	M1	High-risk → Empty roof operation	Emphasis placed on "controlling the empty roof operation" when intervening in high-risk behavior
	M2	Coal face → Inadequate supporting	Emphasis placed on "inadequate supporting" during the safety inspection at coal face
	M3	Main roadway→Failure to strengthen roadway support	Emphasis placed on "failure to strengthen roadway support" during the safety inspection at main roadway
	M4	Traced behavior → Inadequate supporting	Emphasis placed on "inadequate supporting" when intervening in traced behavior
	M5	Non-traced behavior →The Surrounding environment was not checked before the operation	Emphasis placed on "surrounding environment" "not checked before the operation" when intervening in non-traced behavior
	M6	Violation of command → Illegal organization of production	Emphasis placed on "illegal organization of production" when intervening the violation of command

3.3. The Targeted Intervention Node Evaluation

The purpose of the evaluation of targeted intervention nodes is to verify the effectiveness of each intervention node and the application value of targeted intervention nodes in the intervention of unsafe behavior of miners. In this study, the structural equation model (SEM) was constructed to evaluate the targeted intervention nodes [29]. As a general framework of statistical analysis, structural equation model (SEM) is widely used in data analysis [30,31].

3.3.1. Variable Division and Formulate Hypothesis

In order to evaluate the effectiveness of three unsafe behavior-targeted intervention nodes, another latent variable "intervention effect" was introduced. Unsafe behavior incidence (I1), safety management efficiency (I2), and safety climate improvement degree (I3) were selected as the observed variables of latent variable "intervention effect" [32–34]. K1–K5, S1–S6, and M1–M7 in the list of the targeted intervention nodes were the observed variables of the "key work behavior node," the "single-dimensional feature node," and the "multi-dimensional feature node," respectively.

In the process of coal mine safety management, in order to effectively reduce the incidence of roof accidents and improve the targeting of workers' behavior intervention, this paper analyzes and summarizes three kinds of targeted intervention nodes: key work behavior, single-dimensional features, and multi-dimensional features. The practical significance of the three factors and their impact on the effect of intervention were analyzed, and the following hypotheses were formulated:

H1. Key work behavior positively influences the intervention effect.
H2. Single-dimensional features have a direct influence on the intervention effect.
H3. Multi-dimensional features positively influence the intervention effect.

These three types of nodes also affect each other, and the targeted intervention on a single class of nodes will also have an impact on other types of intervention nodes. For example, a single-dimensional feature intervention node may result in a change in the incidence of key work behavior, so the following assumptions are formulated:

H4. Single-dimensional features positively influence key work behavior.
H5. Multi-dimensional features have a direct influence on key work behavior.
H6. Multi-dimensional features have a direct influence on single-dimensional features.

3.3.2. Construct Model and Questionnaire Design

Based on the following hypotheses, an initial model of miners' targeted unsafe behavior intervention node evaluation was established, as shown in Figure 3.

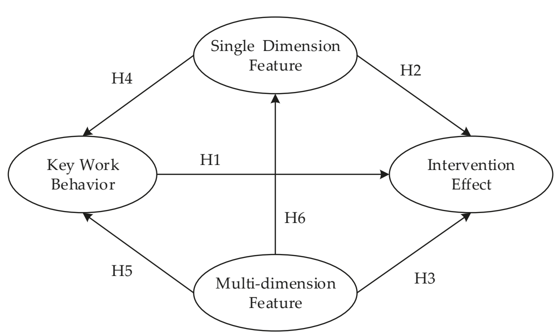


Figure 3. Initial model of miners' targeted unsafe behavior intervention node evaluation.

According to the latent variable and observation variable, the questionnaire was designed, and the experience of other scholars was consulted and verified by experts. The questionnaire's results were compiled with the five-point Likert scale (1–5 represents total disagreement, basic disagreement, partial agreement, basic agreement, and complete agreement, respectively) [35]. This assessment method produces better data distribution. The questionnaire consists of three parts: (1) general information and background information about the interviewees; (2) questions about the role of various targeted intervention nodes in the implementation of behavior interventions; (3) questions about the intervention effect on unsafe behavior through targeted intervention nodes.

A random sampling method was used to survey the employees in a coal mine of the Pingdingshan Coal Industry Group [36]. This coal mine is located in the northwest of Pingdingshan District, Henan province. It was founded in 1956 and put into production on 31 December 1958. It included 4468 registered workers, 376 management personnel, 281 professional and technical personnel, 12 senior titles, and 116 intermediate titles. The questionnaire survey was conducted in this coal mine where several roof accidents resulting in huge casualties and property losses had occurred. A total of 260 questionnaires were sent out, and 248 responses were received. The recovery rate was 95.38%. Among the collected questionnaires, 237 were valid, and the effective rate was 95.56%. Among the interviewees, there were 42 people with a bachelor's degree or above, accounting for 16.15%, and 185 coal miners, accounting for 71.15%; the number of safety administrators at all levels was 38, accounting for 14.62%, and 94, accounting for 36.15%, had worked for more than 15 years.

In order to ensure the applicability and validity of the data in the questionnaire, reliability analysis and validity analysis of the questionnaire data were carried out [37]. In this study, Cronbach's α was used to measure the reliability of setting latent variables [38]. Generally, the higher this coefficient, the higher the reliability. In the exploratory study, results of the questionnaire with a reliability up to 0.7 were acceptable. The data from the questionnaire were input into SPSS 17.0 software (International Business Machines Corporation, Armonk, NY, USA), and the calculated Cronbach's α coefficient showed that the questionnaire data had good reliability, as shown in Table 4.

Table 4. Reliability test results for latent variables.

Latent Variable	Cronbach's α	Cronbach's α Based on Standardization Term	Number of Terms
Key work behavior node	0.813	0.837	6
Single-dimensional feature node	0.824	0.841	7
Multi-dimensional feature node	0.796	0.805	6
Intervention effect	0.811	0.828	3

Validity analysis tests the degree to which the questionnaire reflects the objective reality. Analysis of questionnaire validity was performed via KMO and the Bartlett sphericity test [39]. Under the standard condition, when the KMO is greater than 0.7 and the concomitant probability of the Bartlett sphericity test is less than significance level 0.001, the questionnaire has high validity. The calculated result was KMO = 0.837 > 0.7, and the significance probability of the statistical value of Bartlett sphericity test was less than 0.001, indicating that the index data were normally distributed. This shows that the questionnaire has high validity. According to the above analysis, the questionnaire has good reliability and validity, and the data obtained are suitable for factor analysis.

4. Result and Discussion

Based on the setting latent variables and observed variables (K1–K6, S1–S7, and M1–M6), r_1~r_4 was set as the residual variable of the corresponding latent variable, and e_1~e_{23} was the residual variable of the corresponding observed variable. The model was tested by IBM SPSS AMOS21.0 software (International Business Machines Corporation, Armonk, NY, USA). First, the evaluation data of the structural equation model were calculated, and the model was modified by adjusting the model to meet all kinds of fitting indexes. This paper mainly used the absolute fit index, the comparative fit index, and the parsimonious fit index to judge the fitness of the model [40]. After several revisions of the model and calculation of the fitting data, the final fitting results were obtained. The model fitting reference criteria and the final model fitting data are shown in Table 5.

Table 5. Model fitting criteria and results.

Criteria	X^2/df	RMSEA	NFI	RFI	IFI	TLI	CFI	PGFI
Fit Index	<3	<0.08	>0.9	>0.9	>0.9	>0.9	>0.9	>0.5
Model Index	3.062	0.068	0.923	0.906	0.912	0.956	0.914	0.587

Notes: RMSEA: Root-Mean-Square Error of Approximation; NFI: Comparative Fit Index; RFI: Relative Fit Index; IFI: Increasing Fit Index; TLI: Tucker-Lewis Index; CFI: Comparative Fit Index; PGFI: Goodness of Fit Index.

The result of the model modification shows that the fitting data of the model meets the requirements of the fitting criteria, which shows that the model has good fit [41]. Based on the above analysis, the result of the final model is shown in Figure 4.

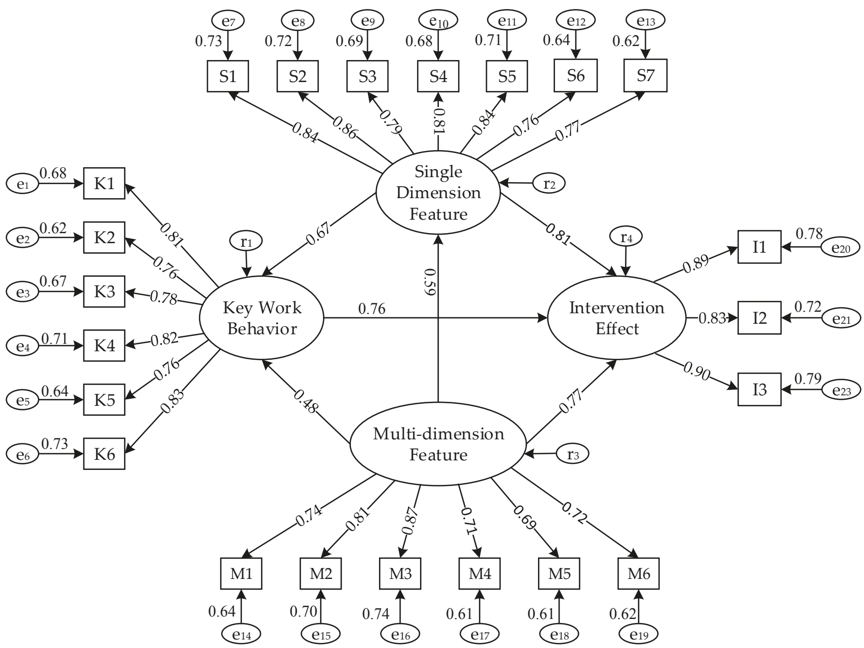


Figure 4. Final evaluation model of targeted intervention nodes for unsafe behavior.

According to the final output results of the evaluation model, the path coefficients of H1–H6 were 0.766, 0.81, 0.77, 0.67, 0.48, and 0.59, respectively, which proved that H1–H6 were significant influences. Based on the path coefficients of H1–H3, the order of the effects of the three target intervention nodes was as follows: single-dimensional feature nodes > multi-dimensional feature nodes > key work behavior nodes.

The results of H1 showed that the positive influence of key work behavior on the prediction of intervention effect was supported, and the accuracy and practicability of targeted intervention nodes K1–K5 were verified. The results of Judi et al. show that the precise positioning of specific unsafe behaviors and the positive strengthening of specific safety practices can effectively reduce occupational accidents [42]. This conclusion is in line with the research results of this paper. There is a very complicated nonlinear relationship between the overall safety state of coal mining enterprises and various unsafe behaviors of miners [43]. Intervening unsafe behavior that can easily lead to coal mine accidents has a positive effect on the overall safety state of coal mines. The key work behaviors K1–K5 involved in H1 come from different groups of workers, including coal mine workers, field commanders, and management staff, all of which have a positive support effect on the prediction results. The results support Luria's view that enhancing the visibility of workers' behavior helps to increase the impact of intervention programs [44]. It is obvious that locating the key types of behavior intervention nodes is an effective way to enhance the visibility of workers' behavior. Because of this remarkable influence relationship, in the daily coal mine safety management, in the prevention of roof accidents, we should pay attention to such unsafe behavior in the process of safety education and training, safety supervision, inspection, etc. [24], as this can effectively reduce the incidence of unsafe behavior.

The significant results of H2 support the positive effect of single-dimensional features on the intervention effect and verify the effectiveness of targeted intervention nodes S1–S7. Bona's view shows that different interventions based on the distribution of risk levels of unsafe behavior among workers can achieve the best results at the lowest management cost [45]. Cheng studied the characteristics of occupational accidents in construction enterprises and found that more accidents occurred on the first day of workers' presence in the workplace. According to the results of the time dimension analysis, enterprises can avoid accidents to a great extent [46]. In the study of human factors in coal mines in China, Chen found that environmental characteristics affected the occurrence of unsafe behavior to some extent. The environmental characteristics mentioned in this study are the spatial distribution of accidents determined by location, working conditions, geological characteristics, etc. It was also shown that mastering the position dimension distribution characteristics of unsafe behavior has a certain role in promoting the safety management efficiency of enterprises [47]. In the analysis of the unsafe behavior of the gas explosion accident in China, Yin et al. found that there are great differences in the frequency of the unsafe behavior of different types of work, which can make safety training easier and more effective [22]. This conclusion is consistent with the results of this paper. Sanmiquel et al. used a database of occupational accident and death reports in Spain's mining industry to analyze the main causes of accidents. Some data mining techniques, such as Bayesian classifiers, decision trees, and contingency tables, were used to discover behavior patterns based on certain rules. The results were helpful in formulating appropriate preventive measures to reduce human injury and death [48]. In the study of road traffic accidents by Kumar, k-mode clustering technology and a correlation analysis algorithm were used to obtain combined characteristics of road traffic accidents. Through the trend analysis of road traffic accidents, it was found that the results of the study have a positive effect on reducing road traffic accidents [25]. The above conclusions are consistent with the results of this paper. The data distribution of dimension information such as risk level, time, position, etc. directly shows the single-dimensional characteristic of insecurity. According to the data distribution of each dimension of unsafe behavior, such dimensions can significantly change the overall safety climate of the enterprise by providing corresponding guidance to the intervention intensity and management resource allocation [49]. In the process of practical safety management of coal mine enterprises, the single-dimensional feature node should be taken as the basis for providing the guidance and basis for the intervention and resource allocation from the aspects of safety responsibility, safety culture, safety education, safety investment, and so on [50].

The positive results of H3 support the influence prediction between the multi-dimensional features and the intervention effect, which proves that the targeted intervention nodes M1–M6 can effectively improve the effect of behavioral intervention as the multi-dimensional feature nodes in the process of targeted intervention. The multi-dimensional intervention node determined by interaction analysis makes the targeted identification of the intervention objects more specific and the intervention process more precise [51]. The combinatorial localization of multi-dimensional information is a deep mining of unsafe behavior characteristics, which is different from general surface information and reveals the deep-seated characteristics of unsafe behavior [52].

The results of H4–H6 show that the three kinds of targeted intervention nodes also have interactive relationships, and the support of the analysis results for this prediction shows that the practical targeted intervention will be a dynamic process of intervention. When using single-dimensional features and multi-dimensional features to intervene, the key work behavior will change, and the single-dimensional feature information data distribution will also change when the intervention is aimed at multi-dimensional features. Therefore, in the process of targeted intervention in enterprises, intervention nodes are constantly changing. Because targeted intervention nodes are based on data analysis and positioning, this fully reflects the advantages of the data analysis method in the process of security management [53]. When implementing targeted intervention in enterprises, it is necessary to collect and update the unsafe behavior data of workers as the intervention continues, and to analyze the data regularly, so as to reposition the unsafe behavior targeted intervention nodes. Finally, in the process of enterprise safety management, the pertinence of worker unsafe behavior intervention is improved, and real-time intervention is realized.

5. Conclusions

The research of targeted interventions in other industries and different fields were combined with behavioral safety management and applied to unsafe behavior interventions. By dividing the data boundary, using safety management data in coal mine production and the method of single-dimensional statistical analysis and multi-dimensional interactive analysis, the targeted intervention process of miners' unsafe behavior was constructed and deemed suitable for coal mines. In the study of miners' unsafe behavior, this process realizes innovation from accident cases to data analysis and to behavior intervention. The method and process of the data analysis of behavior promote further development in the field of behavior intervention, and lay the theoretical foundation and method reference for the realization of accurate identification, accurate intervention, and accurate management of unsafe behavior.

The coal miners' unsafe behavior was taken as the target intervention object. Through the analysis of the data of the coal mine roof accident, the characteristics of data distribution and the interaction rules in the accident case were excavated, and the target intervention nodes of the unsafe behavior were then located. The targeted intervention nodes can help to find unsafe behaviors that can easily lead to roof accidents, and provide a direction for enterprises to reduce the incidence of unsafe behaviors more quickly. According to the different proportion of intervention provided by intervention nodes, enterprises can optimize the allocation of management resources and improve management efficiency. Multiple intervention strategies provided by multiple intervention nodes have an effectively practical significance in controlling roof accidents in coal mines.

The evaluation model of targeted intervention nodes was established by using the structural equation method (SEM), based on which the evaluation process of unsafe behavior-targeted intervention nodes was formed. According to the evaluation results, the effectiveness of each target intervention node was determined, and the scientific and practical nature of the proposed method was further verified.

Based on the roof accident data, we studied the target intervention of miners' unsafe behavior, and the research results have practical significance. The method and process can also be used to intervene on the unsafe behavior of workers in other industries. It is necessary to redefine the data boundary according to the behavior characteristics of workers in different industries, and make adaptive adjustments to the node location process by combining the sources of unsafe behavior data and the behavior characteristics of workers in various industries. This is a possible direction of future research and development.

Author Contributions: R.T. conceived the key ideas and the paper structure; Y.Y. and X.M. conducted the data research; Y.Z. and Q.J. discussed the results of data analysis; R.T. and G.S. wrote the paper; R.T. and Y.Z. reviewed the process.

Funding: The study was financially supported by the National Natural Science Foundation of China (No. 51674268).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jiang, W.; Qu, F.; Zhang, L. Quantitative identification and analysis on hazard sources of roof fall accident in coal mine. *Procedia Eng.* **2012**, *45*, 83–88. [CrossRef]
2. Düzgün, H.S. Analysis of roof fall hazards and risk assessment for zonguldak coal basin underground mines. *Int. J. Coal Geol.* **2005**, *64*, 104–115. [CrossRef]
3. Yang, D.F.; Zhang, L.F.; Chaim Li, B.; Bai, Y.F. Study of roof breaking law of fully mechanized top coal caving mining in ultra-thick coal seam based on fracture mechanics. *Rock Soil Mech.* **2016**, *37*, 2033–2039.
4. Wang, W.; Luo, L.; Yu, W. Study of dynamic pressure roadway supporting scheme under condition of thick composite roof. *J. Coal Sci. Eng.* **2013**, *19*, 119–125. [CrossRef]
5. Cooper, M. Exploratory Analyses of the Effects of Managerial Support and Feedback Consequences on Behavioral Safety Maintenance. *J. Organ. Behav. Manag.* **2006**, *26*, 1–41. [CrossRef]
6. Paul, P.; Maiti, J. The role of behavioral factors on safety management in underground mines. *Saf. Sci.* **2007**, *45*, 449–471. [CrossRef]
7. Chen, D.; Tian, H. Behavior based safety for accident prevention and positive study in China construction projects. *Procedia Eng.* **2012**, *43*, 528–534. [CrossRef]
8. Di Bona, G.; Silvestri, A.; De Felice, F.; Forcina, A.; Petrillo, A. An Analytical Model to Measure the Effectiveness of Safety Management Systems: Global Safety Improve Risk Assessment (G-SIRA) Method. *J. Fail. Anal. Prev.* **2016**, *16*, 1–14. [CrossRef]
9. Priebe, R.; Thomasolson, L. An exploration and analysis on the timeliness of critical incident stress management interventions in healthcare. *Int. J. Emerg. Ment. Health* **2013**, *15*, 39–49.
10. Fogarty, G.; Shaw, A. Safety Climate and the Theory of Planned Behavior: Towards the Prediction of Unsafe Behavior. *Accid. Anal. Prev.* **2010**, *42*, 1455–1459. [CrossRef]
11. Clarke, S. Contrasting perceptual, attitudinal and dispositional approaches to accident involvement in the workplace. *Saf. Sci.* **2006**, *44*, 537–550. [CrossRef]
12. Aryee, S.; Hsiung, H. Regulatory focus and safety outcomes: An examination of the mediating influence of safety behavior. *Saf. Sci.* **2016**, *86*, 27–35. [CrossRef]
13. Namian, M.; Albert, A.; Zuluaga, C.M.; Behm, M. Role of Safety Training: Impact on Hazard Recognition and Safety Risk Perception. *J. Constr. Eng. Manag.* **2016**, *142*, 04016073. [CrossRef]
14. Kouabenan, D.; Ngueutsa, R.; Mbaye, S. Safety climate, perceived risk, and involvement in safety management. *Saf. Sci.* **2015**, *77*, 72–79. [CrossRef]
15. Warszawska, K.; Kraslawski, A. Method for quantitative assessment of safety culture. *J. Loss Prev. Process Ind.* **2016**, *42*, 27–34. [CrossRef]
16. Johnson, S.; Hall, A. The prediction of safe lifting behavior: An application of the theory of planned behavior. *J. Saf. Res.* **2005**, *36*, 63–73. [CrossRef] [PubMed]
17. Lee, Y.S.; Kim, Y.; Kim, S.H.; Kim, C.; Chung, C.H.; Jung, W.D. Analysis of human error and organizational deficiency in events considering risk significance. *Nucl. Eng. Des.* **2004**, *230*, 61–67. [CrossRef]
18. Geller, E. Behavior-based safety and occupational risk management. *Behav. Modif.* **2005**, *29*, 539. [CrossRef]
19. Tong, R.; Chen, C.; Liu, S.; Lu, H.; Ma, J. Research on theory and application of pan-scene data for behavioral safety. *China Saf. Sci. J.* **2017**, *27*, 1–6. [CrossRef]
20. Mahdevari, S.; Shahriar, K.; Esfahanipour, A. Human health and safety risks management in underground coal mines using fuzzy TOPSIS. *Sci. Total Environ.* **2014**, *488*, 85–99. [CrossRef]
21. Tong, R.; Zhang, Y.; Cui, P.; Zhai, C.; Shi, M.; Xu, S. Characteristic analysis of unsafe behavior by coal miners: Multi-dimensional description of the pan-Scene data. *Int. J. Environ. Res. Public Health* **2018**, *15*. [CrossRef] [PubMed]

22. Yin, W.; Fu, G.; Yang, C.; Jiang, Z.; Zhu, K.; Gao, Y. Fatal gas explosion accidents on Chinese coal mines and the characteristics of unsafe behaviors: 2000–2014. *Saf. Sci.* **2017**, *92*, 173–179. [CrossRef]
23. Rocha, R.; Mollo, V.; Daniellou, F. Work debate spaces: A tool for developing a participatory safety management. *Appl. Ergon.* **2015**, *46*, 107–114. [CrossRef] [PubMed]
24. Qing-Gui, C.; Kai, L.; Ye-Jiao, L.; Qi-Hua, S.; Jian, Z. Risk Management and Workers' Safety Behavior Control in Coal Mines. *Int. Symp. Mine Saf. Sci. Eng.* **2012**, *50*, 909–913. [CrossRef]
25. Kumar, S.; Toshniwal, D. A data mining framework to analyze road accident data. *J. Big Data* **2015**, *2*, 26. [CrossRef]
26. Verma, A.; Khan, S.D.; Maiti, J.; Krishna, O.B. Identifying patterns of safety related incidents in a steel plant using association rule mining of incident investigation reports. *Saf. Sci.* **2014**, *70*, 89–98. [CrossRef]
27. Montella, A. Identifying crash contributory factors at urban roundabouts and using association rules to explore their relationships to different crash types. *Accid. Anal. Prev.* **2011**, *43*, 1451–1463. [CrossRef] [PubMed]
28. Guo, S.Y.; Ding, L.Y.; Luo, H.B.; Jiang, X.Y. A Big-Data-based platform of workers' behavior: Observations from the field. *Accid. Anal. Prev.* **2016**, *93*, 299–309. [CrossRef]
29. Eriksson, P.; Larsson, J.; Ossi, P. Managing complex projects in the infrastructure sector-A structural equation model for flexibility-focused project management. *Int. J. Proj. Manag.* **2017**, *35*, 1512–1523. [CrossRef]
30. Iacobucci, D. Structural equations modeling: Fit Indices, sample size, and advanced topics. *J. Consum. Psychol.* **2010**, *20*, 90–98. [CrossRef]
31. Arocena, P.; Nunez, I. An empirical analysis of the effectiveness of occupational health and safety management systems in SMEs. *Int. Small Bus. J.* **2010**, *28*, 398–419. [CrossRef]
32. Morrow, S.L.; McGonagle, A.K.; Dove-Steinkamp, M.L.; Walker, C.T.; Marmet, M.; Barnes-Farrell, J.L. Relationships between psychological safety climate facets and safety behavior in the rail industry: A dominance analysis. *Accid. Anal. Prev.* **2010**, *42*, 1460–1467. [CrossRef]
33. Li, H.; Lu, M.; Hsu, S.C.; Gray, M.; Huang, T. Proactive behavior-based safety management for construction safety improvement. *Saf. Sci.* **2015**, *75*, 107–117. [CrossRef]
34. Zohar, D. Thirty years of safety climate research: Reflections and future directions. *Accid. Anal. Prev.* **2010**, *42*, 1517–1522. [CrossRef]
35. Leung, S. A Comparison of Psychometric Properties and Normality in 4-, 5-, 6-, and 11-Point Likert Scales. *J. Soc. Serv. Res.* **2011**, *37*, 412–421. [CrossRef]
36. Guo, B.; Yiu, T.; Vicente, A. Predicting safety behavior in the construction industry: Development and test of an integrative model. *Saf. Sci.* **2016**, *84*, 1–11. [CrossRef]
37. De, W.; Dodou, D. The driver behavior questionnaire as a predictor of accidents: A meta-analysis. *J. Saf. Res.* **2010**, *41*, 463–470. [CrossRef]
38. Ho, C.; Dzeng, R. Construction safety training via e-Learning: Learning effectiveness and user satisfaction. *Comput. Educ.* **2010**, *55*, 858–867. [CrossRef]
39. Rosenbloom, T.; Wultz, B. Thirty-day self-reported risky driving behaviors of ADHD and non-ADHD drivers. *Accid. Anal. Prev.* **2011**, *43*, 128–133. [CrossRef]
40. Zaira, M.; Hadikusumo, B. Structural equation model of integrated safety intervention practices affecting the safety behavior of workers in the construction industry. *Saf. Sci.* **2017**, *98*, 124–135. [CrossRef]
41. Lenné, M.G.; Salmon, P.M.; Liu, C.C.; Trotter, M. A systems approach to accident causation in mining: An application of the HFACS method. *Accid. Anal. Prev.* **2012**, *48*, 111–117. [CrossRef] [PubMed]
42. Komaki, J.; Barwick, K.; Scott, L. A behavioral approach to occupational safety: Pinpointing and reinforcing safe performance in a food manufacturing plant. *J. Appl. Psychol.* **1978**, *63*, 434–445. [CrossRef] [PubMed]
43. Li, S.; Liu, H.; Yang, Y. Model of Coal Safety Prediction and Evaluation Based on Miners' Unsafe Behavior. *Saf. Coal Mines* **2017**, *48*, 242–245.
44. Luria, G.; Zohar, D.; Erev, I. The effect of workers' visibility on effectiveness of intervention programs: Supervisory-based safety interventions. *J. Saf. Res.* **2008**, *39*, 273–280. [CrossRef] [PubMed]
45. Di Bona, G.; Silvestri, A.; Forcina, A.; Petrillo, A. Total efficient risk priority number (TERPN): A new method for risk assessment. *J. Risk Res.* **2017**, 1–25. [CrossRef]
46. Cheng, C.W.; Leu, S.S.; Lin, C.C.; Fan, C. Characteristic analysis of occupational accidents at small construction enterprises. *Saf. Sci.* **2010**, *48*, 698–707. [CrossRef]

47. Chen, H.; Qi, H.; Long, R.; Zhang, M. Research on 10-year tendency of China coal mine accidents and the characteristics of human factors. *Saf. Sci.* **2012**, *50*, 745–750. [CrossRef]
48. Sanmiquel, L.; Rossell, J.; Vintró, C. Study of Spanish mining accidents using data mining techniques. *Saf. Sci.* **2015**, *75*, 49–55. [CrossRef]
49. Parker, A.; Tones, M.; Ritchie, G. Development of a multilevel health and safety climate survey tool within a mining setting. *J. Saf. Res.* **2017**, *62*, 173–180. [CrossRef]
50. Mohammadfam, I.; Kamalinia, M.; Momeni, M.; Golmohammadi, R.; Hamidi, Y.; Soltanian, A. Evaluation of the Quality of Occupational Health and Safety Management Systems Based on Key Performance Indicators in Certified Organizations. *Saf. Health Work* **2017**, *8*, 156–161. [CrossRef]
51. Haghighi, M.; Taghdisi, M.H.; Nadrian, H.; Moghaddam, H.R.; Mahmoodi, H.; Alimohammadi, I. Safety culture promotion intervention program (SCPIP) in an oil refinery factory: An integrated application of geller and health belief models. *Saf. Sci.* **2017**, *93*, 76–85. [CrossRef]
52. Amiri, M.; Ardeshir, A.; Fazel Zarandi, M.H.; Soltanaghaei, E. Pattern extraction for high-risk accidents in the construction industry: A data-mining approach. *Int. J. Injury Control Saf. Promot.* **2016**, *23*, 264–276. [CrossRef] [PubMed]
53. Cheng, C.; Yao, H.; Wu, T. Applying data mining techniques to analyze the causes of major occupational accidents in the petrochemical industry. *J. Loss Prev. Process Ind.* **2013**, *26*, 1269–1278. [CrossRef]

International Journal of
Environmental Research and Public Health

MDPI

Article

Design and Validation with Influenza A Virus of an Aerosol Transmission Chamber for Ferrets

Nathalie Turgeon [1,2], Marie-Ève Hamelin [3], Daniel Verreault [4], Ariane Lévesque [4], Chantal Rhéaume [3], Julie Carbonneau [3], Liva Checkmahomed [3], Matthieu Girard [4], Guy Boivin [3] and Caroline Duchaine [1,2,*]

1. Centre de Recherche de l'Institut Universitaire de Cardiologie et de Pneumologie de Québec, 2725 Chemin Ste-Foy, Québec City, QC G1V 4G5, Canada; nathalie.turgeon@criucpq.ulaval.ca
2. Département de Biochimie, de Microbiologie et de Bio-Informatique, Faculté des Sciences et de Génie, Université Laval, Québec City, QC G1V 0A6, Canada
3. Centre de Recherche du Centre Hospitalier Universitaire de Québec, CHUL, 2705 boulevard Laurier, Québec City, QC G1V 4G2, Canada; marie-eve.hamelin@crchudequebec.ulaval.ca (M.-È.H.); rheaumec@hotmail.com (C.R.); julie.carbonneau@crchudequebec.ulaval.ca (J.C.); liva.checkmahomed@crchudequebec.ulaval.ca (L.C.); guy.boivin@crchudequebec.ulaval.ca (G.B.)
4. Research and Development Institute for the Agri-Environment, 2700 rue Einstein, Québec City, QC G1P 3W8, Canada; daniel.verreault@environnement.gouv.qc.ca (D.V.); ariane.levesque@irda.qc.ca (A.L.); matthieu.girard@irda.qc.ca (M.G.)

* Correspondence: Caroline.Duchaine@bcm.ulaval.ca; Tel.: +1-418-656-8711 (ext. 5837)

Received: 18 January 2019; Accepted: 12 February 2019; Published: 19 February 2019

Abstract: *Background*: The importance of aerosols in the spread of viruses like influenza is still a subject of debate. Indeed, most viruses can also be transmitted through direct contact and droplets. Therefore, the importance of the airborne route in a clinical context is difficult to determine. The aim of this study was to design a chamber system to study the airborne transmission of viruses between ferrets. *Methods*: A system composed of three chambers connected in series, each one housing one ferret and preventing direct contact, was designed. The chambers were designed to house the ferrets for several days and to study the transmission of viruses from an infected (index) ferret to two naïve ferrets via aerosols and droplets or aerosols only. A particle separator was designed that can be used to modulate the size of the particles traveling between the chambers. The chamber system was validated using standard dust as well as with ferrets infected with influenza A virus. *Conclusions*: The 50% efficiency cut-off of the separator could be modulated between a 5-µm and an 8-µm aerodynamic diameter. In the described setup, influenza A virus was transmitted through the aerosol route in two out of three experiments, and through aerosols and droplets in all three experiments.

Keywords: bioaerosols; influenza virus; ferret animal model; aerosol chamber

1. Introduction

Several infectious diseases are known to be transmissible through the airborne route, such as tuberculosis and measles. The only known disease transmitted only through the airborne route is tuberculosis, as reported by Roy and Milton [1]. Aerosol transmission of other diseases could be preferential or opportunistic [1]. Therefore, it is difficult to assess the importance (or not) of the airborne route in disease transmission. The mode of transmission of some diseases is ambiguous. The evidence of severe acute respiratory syndrome (SARS) airborne transmission was first assessed by indirect evidences such as modeling and epidemiological studies [2]. Moreover, the possible airborne transmission of non-respiratory diseases, like Norovirus, is a subject of investigation [3–5].

For many diseases, dissociating transmission routes such as indirect contact, exposure to large droplets and aerosol transmission through aerosols can be complicated, even in controlled laboratory environments. The World Health Organization considers disease transmission with particles >5 µm as droplets transmission and with particles <5 µm as aerosols transmission [6].

The size of the particles involved in the natural transmission of diseases through the airborne route is hard to establish, especially for viral diseases. In fact, only a few studies have looked at the particle size of airborne viruses that can be found in the environment. Anderson 6 stage cascade impactors, National Institute for Occupational Safety and Health (NIOSH) two-stage bioaerosols cyclone samplers [7–9], and Sioutas personal cascade impactors [10,11] have been used in agricultural and hospital settings. In all these studies, viruses were found in all air sample stages, meaning that large particles as well as small particles can carry viruses. More recently, in a laboratory setting, experiments using ferrets and particle impactors of various cut-off sizes demonstrated that influenza virus can be transmitted via droplets (15.3–5 µm) as well as airborne particles (5–1.5 µm) [12].

Information on the infectious state of airborne viruses is sparse [13]. Culture on appropriate cell lines is still the gold standard to assess virus infectivity. However, the culture of airborne viruses faces several challenges: (1) low concentrations of viruses in the air require large air volume sampling to allow detection (meaning extensive air sampling periods or the use of high-flow air samplers); (2) viruses can be damaged during air sampling; (3) environmental contaminants can interfere with virus or host cell growth (bacteria, mold, dust, etc.).

The use of animals in laboratory settings can overcome most of these challenges. The virus source can be a sick human, an infected animal, or an artificially generated aerosol. By exposing animals to airborne viruses, air sampling can be avoided (preventing virus damage) as well as laboratory virus culture bias in detection. As an example, using animals instead of air samplers can lead to the demonstration that airborne viruses can [14] or cannot [15] infect healthy animals and also that airborne viruses can remain (or not) infectious long enough to travel to a new host. Using a sick animal or human as an aerosol source has also demonstrated that a sick subject can emit aerosols that can potentially infect other susceptible hosts [12].

Unfortunately, the exposure of healthy animals to aerosols emitted by another animal over several days cannot be performed in commercially available apparatus settings. Indeed, cages designed for animal aerosol exposure are meant for a few minutes per day exposure and cannot be used for housing animals for several days. In contrast, animal cages designed to house subjects over extended periods of time are not airtight, and therefore provide limited information about airborne transmission. These cages can be used to prevent direct contact between index and healthy animals and can be placed at various distances but cannot control the size of particles traveling between cages.

In this study, we designed, constructed and tested a system composed of three airtight cages to study the transmission of infectious agents between animals through large droplets and through airborne particles. The system can house three ferrets for up to 10–12 days and is designed to prevent direct contact between animals. We designed a particle separator to prevent large droplets transmission between cages. The cage system is under negative pressure, with high-efficiency particulate air (HEPA) filters on the air inlet and outlet for the users' and environment's protection. This communication describes the main components of this cage system, the particle separator validation using standard aerosol generators as well as a test trial with ferrets and the influenza virus.

2. Materials and Methods

2.1. Design

An aerosol transmission chamber was developed to study the transmission of infectious agents between infected (index) and naïve ferrets. It was designed to expose naïve animals to naturally produced infectious aerosols containing either both droplets and aerosols or aerosols only. A total of three ferrets could be housed in three individual cages for 10–12 consecutive days (Figure 1).

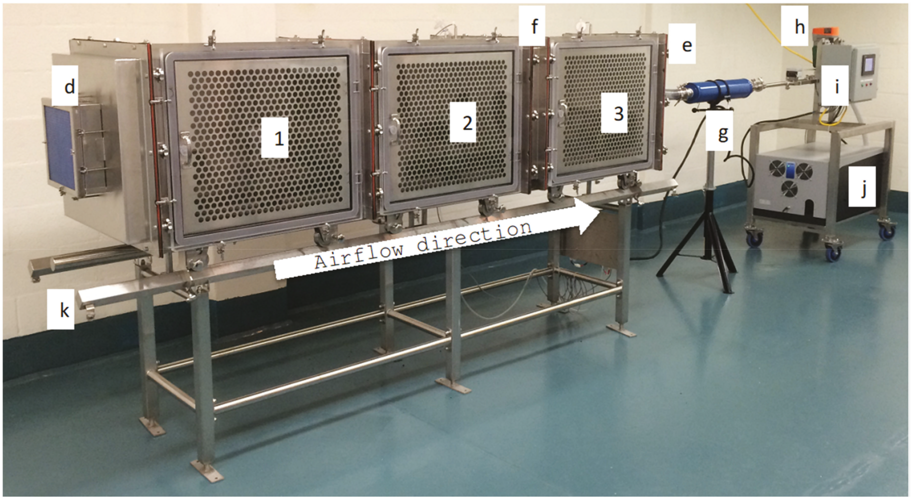


Figure 1. System overview. (**1**) Cage number one, (**2**) cage number two, (**3**) cage number three, (**d**) high-efficiency particulate air (HEPA) filter inlet air, (**e**) HEPA filter exhaust air (not visible in the picture), (**f**) particle separator, (**g**) muffler, (**h**) airflow adjustment valve, (**i**) control panel, (**j**) pump installed in an insulated box, (**k**) support table with rails.

All three stainless steel (grade 316 L) cages are identical. The interior dimensions are 864 mm wide by 610 mm deep and 610 mm high for a total volume of 321 L per cage. The cages have perforated grates on the left and right sides as well as a 102-mm-high excreta pan with a perforated lid, which serves as a floor for the animals (Figure 2). The available space for the animals inside the cages is thus 784 mm wide by 610 mm deep by 508 mm in height, which exceeds Canadian and European guidelines for ferret housing (https://www.ccac.ca, https://www.coe.int/). Sampling ports are located between the side grates and the extremities of the cages on the top, back and bottom walls of the cages, thus making the ports inaccessible to the animals. The cages are assembled together with a 7.5-mm rubber seal between cages (Figure 2). The distance between the grates of cages number one and two is 102 mm.

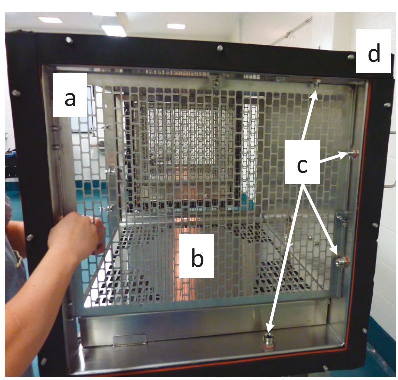


Figure 2. Side view of a cage. (**a**) Perforated grates on each side of the cage, (**b**) excreta pan with a perforated lid, (**c**) sampling ports, (**d**) rubber seal between cages.

The front panel of each cage is composed of a perforated stainless steel door, which is in direct contact with a transparent polycarbonate door (Figure 3A). The polycarbonate door is sealed shut with a 20-mm-thick rubber seal (Figure 3B). The purpose of the stainless steel door is to minimize the electrostatic setting of particles on the polycarbonate door.

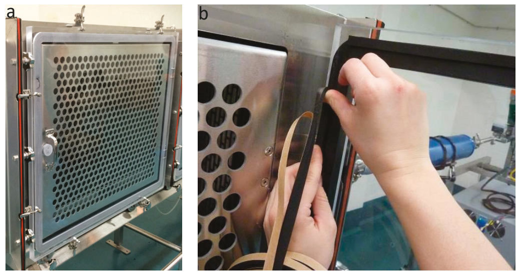


Figure 3. Cage door components. (**a**) Perforated stainless steel door and transparent polycarbonate door. (**b**) Rubber seal of the polycarbonate door.

Airtight feeders and water bottles are connected to the chambers. Butterfly valves in the feeders are used for adding food without disrupting the airflow inside the chamber (Figure 4a). Bars installed inside the feeders prevent the animal from reaching the butterfly valve (Figure 4b).

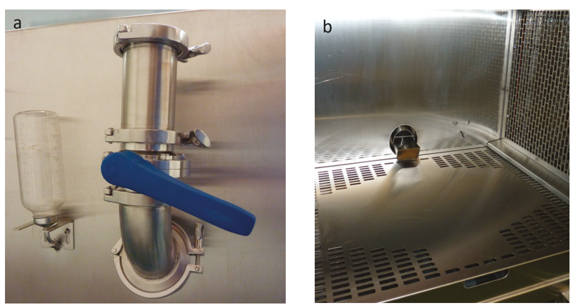


Figure 4. Feeder and water bottle. (**a**) Outside view of the feeder with butterfly valve and water bottle installed on a cage, (**b**) cage inside view with feeder, animal water supply and excreta pan.

A particle separator module was designed to intercept large particles by impaction while letting smaller particles flow through. The separator is composed of a stainless steel plate with 160 orifices distributed in four rows of 40 orifices (Figure 5A). Each orifice is 6.4 mm deep and 5 mm in diameter. An impaction plate is located 5 mm from the outlet of each orifice.

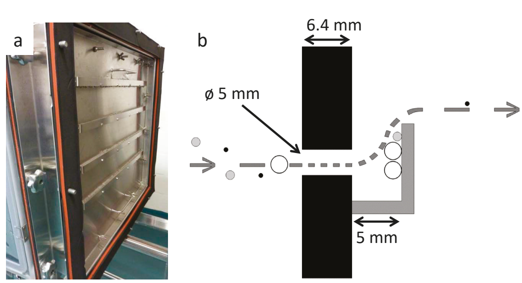


Figure 5. Particles separator. (**a**) Assembly of stainless steel plate with four rows of 40 orifices. On the picture, orifices are covered with impaction plates located 5 mm from the orifice's outlets. (**b**) Schematic representation of the particle separator principle and design.

The velocity of the airborne particles increases as the air from cage number two is forced into cage number three through the orifices of the particle separation module by a suction pump placed upstream of cage number three. Larger particles are impacted on the impaction plates and smaller particles follow the air stream into cage number three (Figure 5b).

The pressure drop across the particle separator is recorded to ensure proper separator function. Temperature and relative humidity are also recorded. All probes are installed on the top of cages, in the inner space between cages one and two and cages two and three. The airflow can be set from 200 L/min to 400 L/min, which correspond to 12 to 25 air changes per hour. Air sampling can be programmed in the three cages, and the system airflow is adjusted automatically to maintain the efficiency of the particle separator.

2.2. Particle Separator D_{50} Measurement

Polydispersed aerosols were produced from Arizona road dust (ISO 12013-1, A2 fine, PTI Powder Technology Inc., Arden Hills, MN, USA) with a powder generator (fluidized bed 3400A, TSI Inc. Shoreview, MN, USA) placed inside cage number one. The powder generator was operated at 25 psi, with a bed purge of 2 L/min, a bed flow of 9 L/min and a chain rotation speed of 40. For each experiment, the aerosol generator was run for 2 h to stabilize the aerosol distribution inside the chamber. A stabilization period of 30 min was also allowed every time the flow rate of the chamber was modified. The aerodynamic distribution of the aerosol was measured with an aerodynamic particle sizer (APS) (model 3321, TSI Inc. Shoreview, MN, USA) equipped with a diluter (model 3302A, TSI Inc. Shoreview, MN, USA) using a dilution factor of 1/20. The APS and diluter were placed under cage number three and connected to a sampling port located at the bottom of the cage at 2” from the particle separator.

The aerosol distribution was measured either with or without the particle separator between cages number two and three at flow rates of 200 L/min and 400 L/min. Measurements were also taken at 400 L/min with 75% of the orifices from the particle separator blocked with masking tape, leaving only 40 orifices open for the passage of air.

For every particle size from the APS, a mean count was calculated from 25 to 40 min of readings. The mean count obtained with the particle separator was divided by the mean count obtained at the same flow rate without the separator, thus giving a ratio of particles passing through the separator for each particle size. These ratios were plotted on graphs as illustrated in Figure 6 to estimate the D_{50} diameter. The experiment was repeated four times and the mean D_{50} diameter for each condition used was extrapolated.

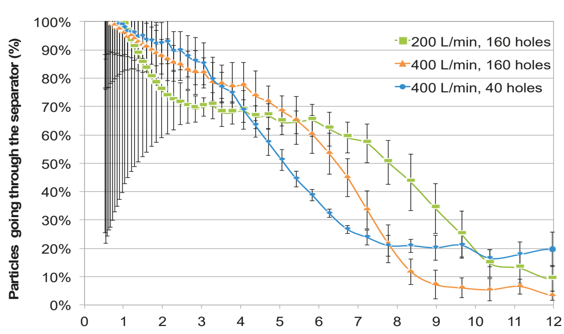


Figure 6. Particle separator D_{50} measurement as a function of airflow and the number of separator orifices used. Comparison of particle distribution in cage three with and without a particle separator, as measured with an aerodynamic particle sizer (APS) located at 2” from the particle separator.

2.3. Nano Particles Concentration in Cages Two and Three

The polydispersed nanometer particle size was generated using a collision 6-jet nebulizer (BGI, Waltham, MA, USA), filled with 50 mL of buffer (20 mM Tris-HCl, 100 mM NaCl, 10 mM MgSO4, pH 7.5) and supplied with 20 psi. Aerosols were passing through a diffusion dryer (model 3062, TSI Inc. Shoreview, MN, USA) and a neutralizer (model 3012A, TSI Inc. Shoreview, MN, USA) before entering the chamber in cage number one. Nebulization was started 30 min before starting measurements to stabilize the particle concentration in the three chambers. Measurements were performed using a NanoScan SMPS (model 3910, TSI Inc. Shoreview, MN, USA). The NanoScan sampled cage two for 10 min and then sampled cage three for 10 min. The experiments were repeated three times in all conditions. Particle size distribution as well as total nanoparticles were compared between cages two and three.

2.4. Experiments with Ferrets

Three groups of three seronegative (800- to 1000-*g*) male ferrets (Triple F Farms Inc., Gillett, PA, USA) were housed consecutively in the system for 7 to 12 days. The ventilation system was set at 200 L/min, with 160 holes of the particle separator for all experiments. Ferrets housed in cage one were infected intra-nasally with 250 μL (125 μL per nostril) containing 4.5 log TCID50/mL of the A/California/7/2009 (H1N1) influenza A virus. Nasal wash was collected every day by instillation of 5 mL Phosphate-Buffered Saline (PBS) into the intranasal cavity.

System ventilation was stopped before opening the cages' sealed doors. Animals were manipulated in the following order: first the ferret from cage three, followed by the ferret from cage two, and then the ferret from cage one. Viral titer from the nasal wash was determined by plaque assay on ST6GalI-MDCK cells.

Air samples were collected every day using NIOSH two-stage bioaerosol cyclone samplers and SKC BioSamplers. Air samplers were connected to sampling ports located in cage two (between the perforated grates of cages one and two) as well as in cage three (between the particle separator and the perforated grate). Air sampling with NIOSH two-stage bioaerosol cyclone samplers was performed at 2 L/min for 24 h. At this flow rate, the cut-off separations of the NIOSH two-stage bioaerosol cyclone sampler were: 4 μm for first stage, 1.7 μm for second stage, and the remaining particles were collected on the backup filter. Air sampling started when ferrets were placed in cages after the infection of the ferret from cage one, and was stopped before shutting down the ventilation system for the daily nasal wash. Samples were eluted from NIOSH two-stage bioaerosol cyclone samplers by vortexing for 1 min in MEM (minimal essential medium; 5 mL in first stage, 500 μL in second stage, 5 ml in backup filter). Air sampling with SKC BioSamplers was performed at 11–14 L/min (determined by critical opening of the instrument) for 20 min and was set before shutting down the ventilation system for daily animal care. SKC BioSamplers were filled with 20 mL of MEM (minimal essential medium) without bovine serum albumin (BSA). After air sampling, 150 μL of BSA was added to the remaining liquid of the SKC BioSampler. Air samples were kept frozen at −80 °C until further quantitation. The virus concentration in NIOSH two-stage bioaerosol cyclone air samples was measured using qPCR [16]. The virus concentration in BioSampler air samples was measured using plaque assays on ST6GalI-MDCK cells and embryonated chicken eggs [17].

Animal procedures were approved by the Institutional Animal Care Committee of Université Laval according to the guidelines of the Canadian Council on Animal Care (protocol 2015031).

3. Results

3.1. System Validation

The D_{50} diameters at a chamber flow rate of 400 L/min were 5.07 ± 0.18 μm and 6.36 ± 0.22 μm with 40 and 160 open orifices, respectively, and 7.8 ± 0.78 μm at 200 L/min with 160 open orifices (Figure 6). To ensure correct air sampling, measurements were also made at 2″ from the air exhaust of cage number three for validation purposes. The D_{50} values measured at 2″ from air exhaust were the same (data not shown).

Data collected from two sampling ports in cage number three (2″ after the particle separator and 2″ before the exhaust) were used to document particle deposition at 200 L/min and 400 L/min without the particle separator. Particle deposition was the same at the two flow rates. The deposition of particles of particles <3.5 μm was less than 30%.

The presence or absence of the particle separator had no significant effect on the nanoparticle size distribution (10 nm to 1 μm) between cage number two and cage number three. Under all conditions the total concentration of nanoparticles in cage number three was 0% to 12% lower than that in cage number two (Table 1).

Table 1. Nanoparticles (10 nm–420 nm) total concentration difference between cages two and three.

Airflow—Particles Separator Configuration	Decrease in Concentration
200 L/min—no separator	0% ± 9%
200 L/min—160 orifices	11% ± 5%
400 L/min—no separator	12% ± 19%
400 L/min—160 orifices	6% ± 5%
400 L/min—40 orifices	11% ± 5%

3.2. Experiments with Ferrets

In all three experiments, influenza A virus was detected from the index ferret (cage one) nasal wash from day 1 to days 6 or 7 (Figure 7). Nasal washes from ferrets housed in cage two were positive from days 3 or 5 in all three experiments. Influenza virus was detected from nasal washes of the ferret housed in cage three 7 days after the infection of the index ferret in experiment 1, and after 4 days in experiment 3. No virus was detected in the nasal wash of the ferret housed in cage three from experiment 2.

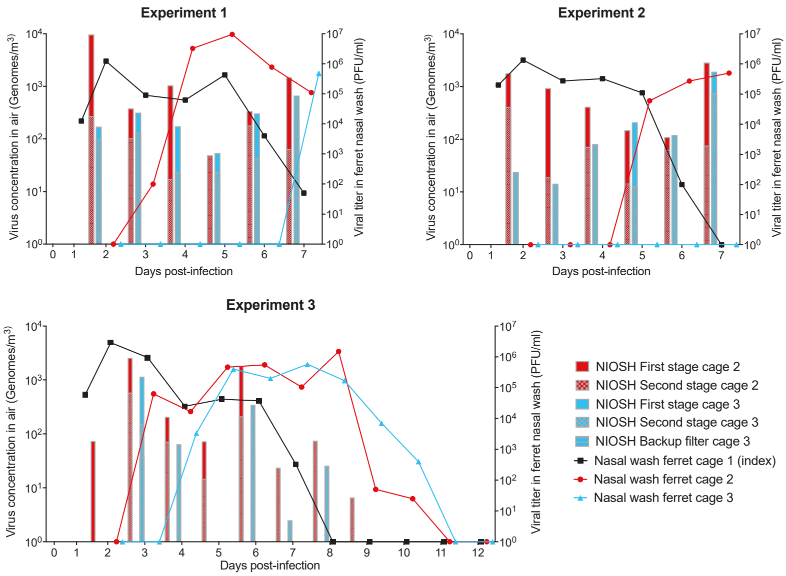


Figure 7. Influenza genome per cubic meter of air, and influenza virus titer in nasal washes of ferrets hosted in cage system for 7 or 12 days from three experiments. The index ferret (cage one) was infected on day 0. Air samples were collected using National Institute for Occupational Safety and Health (NIOSH) two-stage bioaerosol cyclone samplers. Genome concentrations found with the NIOSH first stage, second stage and backup filter are superimposed.

In all experiments, airborne influenza virus genome concentrations up to 10^4 genomes/m^3 were detected from cages two and three using the NIOSH two-stage bioaerosol cyclone sampler (Figure 7) from day 2 until the end of the experiment.

The influenza virus genome concentration was higher in cage two compared to cage three, except for experiment 2 on day 5. Influenza virus genomes were detected in the NIOSH backup filter in only one sample (experiment 3, cage three, day 3). No cultivable viruses were detected from SKC BioSamplers air samples using plaque assay and embryonated chicken eggs (data not shown).

4. Discussion

The particle separator was efficient to prevent the circulation of droplets between cages two and three, as demonstrated with dust experiments. Airflow modulation impacts the particle separator D_{50}. However, airflow must be adjusted in the range of animal comfort. Therefore, the airflow should be maintained between 163 L/min and 400 L/min, which correspond to 10 and 25 air changes per hour, respectively.

Except for one sampling day where viral genomes were detected on the backup filter, influenza virus genomes were detected in the NIOSH first and/or second stages only. This means that most genomes emitted by sick ferrets were carried on particles larger than 1.7 µm. This result is consistent with the results obtained by Zhou et al. [12].

The ferrets housed in cage two were infected with influenza virus in all three experiments. This means that influenza-positive ferrets (such as the index ferret in cage one) can emit airborne particles and/or droplets containing infectious influenza virus in sufficient concentration for disease transmission without direct contact. Ferrets housed in cage three were infected with influenza virus in two out of three experiments. This result indicates that ferrets can be infected by influenza virus carried on airborne particles emitted by influenza positive ferrets, in accordance with the literature [12,14].

Nasal washes of ferrets housed in cage two were positive 2–4 days after the infection of ferrets housed in cage one. Nasal washes of ferrets housed in cage three were positive 1–4 days after washes from ferrets housed in cage two were found to be positive for influenza. The delay between the influenza detection schedule in ferrets in cages two and three can be explained by a lower virus concentration in the air and the infection route. Indeed, the airborne influenza virus genome concentration was lower in cage three compared to cage two in 19 out of 20 sampling days. Large particles eliminated by the separator likely contained high virus concentrations. The airborne influenza virus genome concentration in cage three reached 5×10^2 genomes/m^3 only when influenza virus was detected in the nasal wash of the ferret housed in cage two. Therefore, it is possible that the airborne virus concentration in cage three reached the required concentration to transmit the infection only when the ferret housed in cage two showed flu symptoms. More experiments would be required to elucidate this phenomenon.

5. Conclusions

This paper describes a chamber system that can be used for airborne disease transmission studies. The system is airtight and particle size distribution through the system is satisfactory (no significant difference was found for particles <420 nm). The particle separator D_{50} can be modulated between 5.07 and 7.8 µm by changing the airflow. To further reduce the D_{50}, plates with smaller holes could replace the actual hole plates of the particle separator. This would lead to more accelerated particles and, therefore, more particles being captured by the impaction plates. Ferrets can be used as models for the study of many mammalian viruses, including filovirus [18], respiratory syncytial virus [19] and Morbillivirus [20]. Therefore, the studies that can be conducted in this chamber system are not limited to influenza viruses. Moreover, this cage system can be adapted to accommodate other animals like rats or rabbits. Indeed, smaller, regular cages can be placed inside the ventilated airtight cages.

Author Contributions: Original project idea: C.D. and N.T.; system design, D.V., A.L., M.G., N.T., C.D., C.R., and M.-È.H.; methodology, A.L., D.V., M.G., N.T., and M.-È.H.; formal analysis, N.T., M.È.H., and D.V.; investigation, C.R, N.T., J.C., L.C., M.È.H, D.V.; writing—original draft preparation, N.T., D.V., and M.È.H.; writing—review and editing, D.V., M.È.H, A.L., C.R., J.C., L.C., M.G., G.B., and C.D.; visualization, N.T.; supervision, M.G., C.D., and G.B.; project administration, M.G., C.D., and G.B.; funding acquisition, C.D. and G.B.

Funding: This research was funded by Ministère du Développement Économique, de l'Innovation et des Exportations du Gouvernement du Québec through the "programme de soutien à la recherche volet 2: appui au financement d'infrastructures de recherche" project M-11638 (CD), the Quebec Respiratory Health Network (CD), NSERC discovery grant RGPIN-2014-05900 (CD) and Canadian Institutes of Health Research Foundation grant 148361 (GB).

Acknowledgments: We thank Serge Simard for statistical analysis. We also thank Cédric Morin for software conception and Dan Zegan for his input the design of the chamber system. NIOSH two-stage bioaerosols cyclone samplers were kindly provided by William G. Lindsley.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Roy, C.J.; Milton, D.K. Airborne transmission of communicable infection—The elusive pathway. *N. Engl. J. Med.* **2004**, *350*, 1710–1712. [CrossRef] [PubMed]
2. Yu, I.T.; Li, Y.; Wong, T.W.; Tam, W.; Chan, A.T.; Lee, J.H.; Leung, D.Y.; Ho, T. Evidence of airborne transmission of the severe acute respiratory syndrome virus. *N. Engl. J. Med.* **2004**, *350*, 1731–1739. [CrossRef] [PubMed]
3. Bonifait, L.; Charlebois, R.; Vimont, A.; Turgeon, N.; Veillette, M.; Longtin, Y.; Jean, J.; Duchaine, C. Detection and quantification of airborne norovirus during outbreaks in healthcare facilities. *Clin. Infect. Dis.* **2015**, *61*, 299–304. [CrossRef] [PubMed]

4. Courault, D.; Albert, I.; Perelle, S.; Fraisse, A.; Renault, P.; Salemkour, A.; Amato, P. Assessment and risk modeling of airborne enteric viruses emitted from wastewater reused for irrigation. *Sci. Total Environ.* **2017**, *592*, 512–526. [CrossRef] [PubMed]
5. Uhrbrand, K.; Koponen, I.; Schultz, A.; Madsen, A. Evaluation of air samplers and filter materials for collection and recovery of airborne norovirus. *J. Appl. Microbiol.* **2018**, *124*, 990–1000. [CrossRef] [PubMed]
6. WHO. *Infection Prevention and Control of Epidemic- and Pandemic-Prone Acute Respiratory Infections in Health Care: WHO Guidelines*; World Health Organization: Geneva, Switzerland, 2014; Available online: http://www.who.int/csr/bioriskreduction/infection_control/publication/en/ (accessed on 7 February 2019).
7. Lindsley, W.G.; Blachere, F.M.; Davis, K.A.; Pearce, T.A.; Fisher, M.A.; Khakoo, R.; Davis, S.M.; Rogers, M.E.; Thewlis, R.E.; Posada, J.A.; et al. Distribution of airborne influenza virus and respiratory syncytial virus in an urgent care medical clinic. *Clin. Infect. Dis.* **2010**, *50*, 693–698. [CrossRef] [PubMed]
8. Leung, N.H.; Zhou, J.; Chu, D.K.; Yu, H.; Lindsley, W.G.; Beezhold, D.H.; Yen, H.L.; Li, Y.; Seto, W.H.; Peiris, J.S.; et al. Quantification of Influenza Virus RNA in Aerosols in Patient Rooms. *PLoS ONE* **2016**, *11*, e0148669. [CrossRef] [PubMed]
9. O'Brien, K.M.; Nonnenmann, M.W. Airborne Influenza A Is Detected in the Personal Breathing Zone of Swine Veterinarians. *PLoS ONE* **2016**, *11*, e0149083. [CrossRef] [PubMed]
10. Yang, W.; Elankumaran, S.; Marr, L.C. Concentrations and size distributions of airborne influenza A viruses measured indoors at a health centre, a day-care centre and on aeroplanes. *J. R. Soc. Interface* **2011**, *8*, 1176–1184. [CrossRef] [PubMed]
11. Lednicky, J.A.; Loeb, J.C. Detection and Isolation of Airborne Influenza A H3N2 Virus Using a Sioutas Personal Cascade Impactor Sampler. *Influenza Res. Treat.* **2013**, *2013*, 656825. [CrossRef] [PubMed]
12. Zhou, J.; Wei, J.; Choy, K.T.; Sia, S.F.; Rowlands, D.K.; Yu, D.; Wu, C.Y.; Lindsley, W.G.; Cowling, B.J.; McDevitt, J.; et al. Defining the sizes of airborne particles that mediate influenza transmission in ferrets. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E2386–E2392. [CrossRef] [PubMed]
13. Blachere, F.M.; Cao, G.; Lindsley, W.G.; Noti, J.D.; Beezhold, D.H. Enhanced detection of infectious airborne influenza virus. *J. Virol. Methods* **2011**, *176*, 120–124. [CrossRef] [PubMed]
14. Sutton, T.C.; Finch, C.; Shao, H.; Angel, M.; Chen, H.; Capua, I.; Cattoli, G.; Monne, I.; Perez, D.R. Airborne transmission of highly pathogenic H7N1 influenza virus in ferrets. *J. Virol.* **2014**, *88*, 6623–6635. [CrossRef] [PubMed]
15. Richard, M.; Schrauwen, E.J.; de Graaf, M.; Bestebroer, T.M.; Spronken, M.I.; van Boheemen, S.; de Meulder, D.; Lexmond, P.; Linster, M.; Herfst, S.; et al. Limited airborne transmission of H7N9 influenza A virus between ferrets. *Nature* **2013**, *501*, 560–563. [CrossRef] [PubMed]
16. Ellis, J.; Iturriza, M.; Allen, R.; Bermingham, A.; Brown, K.; Gray, J.; Brown, D. Evaluation of four real-time PCR assays for detection of influenza A(H1N1)v viruses. *Euro Surveill.* **2009**, *14*, 19230. [CrossRef] [PubMed]
17. Turgeon, N.; McNicoll, F.; Toulouse, M.J.; Liav, A.; Barbeau, J.; Ho, J.; Grund, C.; Duchaine, C. Neuraminidase activity as a potential enzymatic marker for rapid detection of airborne viruses. *Aerosol Sci. Technol.* **2011**, *45*, 183–195. [CrossRef]
18. Siragam, V.; Wong, G.; Qiu, X.G. Animal models for filovirus infections. *Zool. Res.* **2018**, *39*, 15–24. [CrossRef] [PubMed]
19. Taylor, G. Animal models of respiratory syncytial virus infection. *Vaccine* **2017**, *35*, 469–480. [CrossRef] [PubMed]
20. Da Fontoura Budaszewski, R.; von Messling, V. Morbillivirus Experimental Animal Models: Measles Virus Pathogenesis Insights from Canine Distemper Virus. *Viruses* **2016**, *8*, 274. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

Risk Factors Associated with Dengue Virus Infection in Guangdong Province: A Community-Based Case-Control Study

Jundi Liu, Xiaolu Tian, Yu Deng, Zhicheng Du, Tianzhu Liang, Yuantao Hao and Dingmei Zhang *

School of Public Health, Sun Yat-sen University, Guangzhou 510080, China; liujd5@mail2.sysu.edu.cn (J.L.); tianxlu@mail2.sysu.edu.cn (X.T.); dengyu9@mail.sysu.edu.cn (Y.D.); duzhch3@mail2.sysu.edu.cn (Z.D.); tianzhu@bu.edu (T.L.); haoyt@mail.sysu.edu.cn (Y.H.)

* Correspondence: zhdingm@mail.sysu.edu.cn; Tel.: +86-208-733-0595

Received: 26 December 2018; Accepted: 14 February 2019; Published: 20 February 2019

Abstract: Dengue fever (DF) is a mosquito-borne infectious disease that is now an epidemic in China, Guangdong Province, in particular and presents high incidence rates of DF. Effective preventive measures are critical for controlling DF in China given the absence of a licensed vaccination program in the country. This study aimed to explore the individual risk factors for the dengue virus infection in Guangdong Province and to provide a scientific basis for the future prevention and control of DF. A case-control study including 237 cases and 237 controls was performed. Cases were defined for samples who were IgG-antibody positive or IgM-antibody positive, and willing to participate in the questionnaire survey. Additionally, the controls were selected through frequency matching by age, gender and community information from individuals who tested negative for IgG and IgM and volunteered to become part of the samples. Data were collected from epidemiological questionnaires. Univariate analysis was performed for the preliminary screening of 28 variables that were potentially related to dengue virus infection, and multivariate analysis was performed through unconditioned logistic regression analysis to analyze statistically significant variables. Multivariate analysis revealed two independent risk factors: Participation in outdoor sports (odds ratio (OR) = 1.80, 95% confidence interval (CI) 1.17 to 2.78), and poor indoor daylight quality (OR = 2.27, 95% CI 1.03 to 5.03). Two protective factors were identified through multivariate analysis: 2 occupants per room (OR = 0.43, 95% CI 0.28 to 0.65) or $\geq$3 occupants per room (OR = 0.45, 95% CI 0.23 to 0.89) and air-conditioner use (OR = 0.46, 95% CI 0.22 to 0.97). The results of this study were conducive for investigating the risk factors for dengue virus infection in Guangdong Province. Effective and efficient strategies for improving environmental protection and anti-mosquito measures must be provided. In addition, additional systematic studies are needed to explore other potential risk factors for DF.

Keywords: dengue fever; Aedes albopictus; living environment; logistic regression analysis

1. Introduction

Dengue fever (DF) is an acute viral disease caused by four distinct serotype dengue viruses, transmitted between humans by the mosquito *Aedes aegypti*. Statistics provided by the World Health Organization (WHO) has shown that only nine countries experienced severe dengue epidemics before 1970. Recent research, however, has revealed that the number of countries with severe dengue epidemics now exceeds 100 and that the actual number of DF cases has reached approximately 390 million, among which 500,000 patients required hospital admission because of severe infection [1]. In Asian and American countries wherein dengue is endemic, the effect of dengue is approximately

1300 disability-adjusted life years per million population; this effect is highly similar to the disease burden of related childhood and tropical diseases, including tuberculosis [2].

In Asia, the coverage of epidemic dengue hemorrhagic fever (DHF) has expanded geographically westward from southeast Asian countries to India, Sri Lanka, the Maldives, and Pakistan and then eastward to China [2]. In China, the first reported DF outbreak due to dengue virus type 4 occurred in Foshan City, Guangdong Province, in 1978; DF then began to spread to southern Chinese provinces from Foshan City [3]. Since then, Guangdong Province has exhibited the highest incidence of DF in mainland China, and more than 65% of all DF cases in the country were reported in this province [4]. In 2014, the number of DF cases increased dramatically to 38,753 in Guangdong province and accounted for 93.83% of DF cases in mainland China [5]. *Aedes albopictus*, an aggressive mosquito species that is also one of the main vectors of DF, is widely distributed with high density in Guangdong Province [6,7]. Therefore, controlling the outbreak of DF in Guangdong Province, which can act as a bridge for DF transmission to other provinces in mainland China, is urgent. Unfortunately, effective drugs and a licensed vaccination program for the treatment or prevention of DF are unavailable in China.

Understanding the risk factors for dengue virus infection is necessary to control this disease effectively. However, most of the current case–control studies on risk factors for DF focused on severe dengue infections, such as dengue shock syndrome and DHF, and variables related to clinical and laboratory indexes [8–11]. Environmental factors, such as heavy rainfall and global warming, and factors based on the awareness and knowledge of dengue prevention measures are also responsible for drastic reductions in dengue transmission [12,13]. Several macroscopic descriptive studies have been performed to explore the risk factors for dengue virus infection and to provide a basis for formulating control strategies in Guangdong Province. These studies have obtained considerable information on the group level and climate factors but limited information on personal protective measures [14,15].

In this case-control study, we evaluated potential risk factors, including personal life activities, environmental sanitation, housing situation, living conditions, mosquito protection status, and residential surroundings to identify additional risk factors for DF on the individual level and to recommend specific approaches for preventing DF.

2. Materials and Methods

2.1. Community Selection and Study Design

Guangzhou City and Zhongshan City is located in the Pearl River Delta Region of Guangdong, which is the main area where DF is highly epidemic [16,17]. Guangzhou City is the capital of Guangdong Province, and the first reported case of autochthonous DF occurred in Zhongshan City [18]. Thus, the prevalence of DF in Guangzhou City and Zhongshan City is a good representation of the prevalence of DF in Guangdong Province.

This case-control study was performed on the basis of the project of Research on the Prevention and Control of Human Immunodeficiency Virus and Hepatitis B Virus in Guangdong Province. This project has constructed a database containing 200,000 samples. The demographic information contained by the database could be seen in our related publication [19]. Approximately 30–35 persons per month were sampled from every age group (<19 years, 19–40 years, 41–65 years, and >65 years) over a 2-year period from September 2013 to August 2015.

2.2. Ethical Statement

This work obtained ethics approval from the Institutional Review Board of the School of Public Health at Sun Yat-sen University (L2017030) in line with the guidelines for the protection of human subjects. All research participants or their guardians provided signed written informed consent after being informed of the research subject matter and were assured that their personal information would be kept private. Each participant had the right to withdraw from this study at any time.

2.3. Enzyme Immunoassay Test

Enzyme-linked immunosorbent assay (ELISA) was used to detect dengue IgG and IgM antibodies. The IgG antibody was measured through indirect ELISA (LOT: 01P20A006, Inverness Medical/Panbio, Windsor, Australia). The IgM antibody was tested via capture ELISA (LOT:01P30A002, Inverness Medical/Panbio, Windsor, Australia). Undefined results were confirmed through the colloidal gold method (LOT: DEN141001, Inverness Medical/Cortez, Calabasas, CA, USA). The details of the assays could be found in a previous publication [19].

2.4. Case Definition and Control Selection

Among the 3136 serum samples, 305 and 103 were identified as IgG-antibody positive and IgM-antibody positive, respectively, and 34 were positive for both antibodies. Thus, 374 individuals with antibody-positive samples were defined as dengue-infected, and those willing to participate in the questionnaire survey were selected as the members of the case group.

In this study, 256 patients with dengue infections opted to fill in the questionnaires, and 19 questionnaires missing most of the important information were eliminated. Eventually, 237 cases were included in the case group.

The controls were selected through frequency matching from individuals who tested negative for IgG and IgM. Specifically, candidate controls were stratified in accordance with the age, gender, and community information of the case group and selected through convenience sampling, wherein participants volunteered to become part of the samples, from each layer. Additionally, age matching was carried out according to $\leq$15 years, 16–30 years, 31–50 years, 51–65 years and $\geq$66 years. Finally, 237 controls were selected.

2.5. Data Collection and Analysis

The phone questionnaire investigation was conducted by trained investigators. Subjects who interrupted the telephone investigation and whose questionnaire information contained logic errors were interviewed face-to-face to verify the integrity and validity of their information. Just as shown in Supplementary Materials File 1, the main contents of the questionnaire included general demographic characteristics (age, gender, blood type, and average household income). It also included personal life activities, such as activities in the park, outdoor sports (such as hiking, mountain climbing, and camping), and outbound tourism experience. Moreover, the questionnaire presented questions related to environmental sanitation (domestic sewage and garbage management and participation in community hygiene management interventions); housing situation, such as the age and area of domiciles and living floor; and living conditions (average numbers of occupants per room, use of air conditioning, quality of indoor daylight, and presence of animals or aquatic plants on property). It also had questions on mosquito protection status (use of mosquito nets and pesticide) and residential surroundings (presence of junk yards, ponds, or construction sites within 200 m).

"Contact with patients with DF" was defined as previous living or working experience with patients with DF in the past years in his/her life. The definition of "outdoor activities in parks" was established as outdoor activity for at least two times per week and for more than half an hour each time. "Participation in outdoor sports" was defined as participation in activities at least twice a year for more than half an hour each time. The definition of "good indoor daylight quality" was established on the basis of the minimum requirement for sunshine in general residences of not less than 2 h on extremely cold days. The definition of "good ventilation" was given to an open-ventilation area of not less than 5% of the floor area of each domicile. The use of mosquito repellent referred to the use of mosquito coils and insecticide vaporizers. "Occasionally" was defined as less than once a week, and "often" was defined as at least once a week.

Epidata 3.1 software (Epidata Association, Odense, Denmark) was used to establish a database of individual risk factors for dengue infection among residents in Guangdong Province. All data were analyzed by SPSS statistics 23.0 software (SPSS Inc., Chicago, IL, USA). The χ^2 test was used to test for differences in demographic characteristics between cases and controls. A univariate unconditioned logistic regression analysis was applied for the preliminary screening of variables identified by using the questionnaire but not for the variables of general demographic characteristics. An unconditioned logistic regression analysis method for multivariate analysis was employed to analyze the relative importance of statistically significant variables in univariate analysis. Additionally, considering the rule of frequency matching design, the age, gender and community information variables were also introduced. Then, stepwise regression was used to establish a regression equation. $p < 0.05$ was set as the significance level of the χ^2 test and multivariate analysis. However, to avoid missing important factors, $p < 0.1$ was set as the significance level in univariate analysis. In addition to odds ratios (ORs), 95% confidence intervals (CIs) were used to express associations.

3. Results

3.1. General Demographic Characteristics of the Samples

A total of 3136 serum samples collected from the residents of Yuexiu District in Guangzhou City (n = 699), Liwan District in Guangzhou City (n = 1386), Torch Development Area in Zhongshan City (n = 180), and Xiaolan Town in Zhongshan City (n = 871) were selected via stratified cluster sampling rooted in the database for serological testing. The study population had a male: female ratio of 1:1.92. Age statistics showed that the age group of ≥66 years old accounted for the largest proportion of the study population (25.86%), followed by 51–65 years old.

Finally, 474 subjects, including 237 cases and 237 controls, were recruited successfully (Figure 1). No statistical difference in gender (p = 0.950) and age (p = 0.127) existed between persons who were willing to receive the questionnaire survey and those who were unwilling to receive the questionnaire survey. The gender ratio was 1:1.66 (male: female) in both the case group and the control group. The demographic characteristics of the two groups were comparable (Table 1).

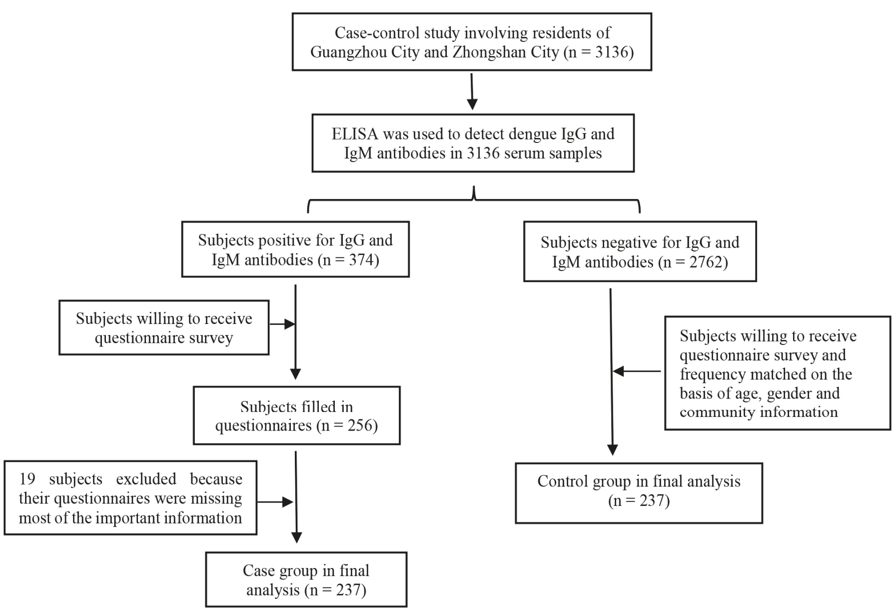


Figure 1. Consort diagram of cases and controls used in the study.

Table 1. Demographic characteristics of cases and controls.

Demographic Characteristics [a]	Cases (*n*/%) (*n* = 237)	Controls (*n*/%) (*n* = 237)	*p*-Value
Age (years)			-
≤15	15 (6.3)	15 (6.33)	-
16–30	13 (5.49)	13 (5.49)	
31–50	33 (13.92)	33 (13.92)	
51–65	43 (18.14)	43 (18.14)	
≥66	133 (56.12)	133 (56.12)	
Gender			
Male	89 (37.55)	89 (37.55)	
Female	148 (62.45)	148 (62.45)	
Residential status			0.149
Permanent residents	225 (94.94)	231 (97.47)	
Floating population [b]	12 (5.06)	6 (2.53)	
Number of residents per household			0.394
1	27 (11.39)	17 (7.17)	
2–3	126 (53.16)	125 (52.75)	
4–5	75 (31.65)	84 (35.44)	
≥6	9 (3.80)	11 (4.64)	
Monthly per capita family income (¥) [c]			0.317
<2000	50 (21.10)	36 (15.19)	
2000–4999	145 (61.18)	162 (68.35)	
5000–7999	35 (14.77)	31 (13.08)	
≥8000	7 (2.95)	8 (3.38)	
Blood type			0.196
A	18 (7.59)	14 (5.91)	
B	15 (6.33)	19 (8.02)	
O	29 (12.24)	46 (19.40)	
AB	9 (3.80)	6 (2.53)	
Unknown	166 (70.04)	152 (64.14)	

[a] Except where otherwise indicated, values are the number (percentage) of patients with the characteristic. [b] Floating population refers to migrants who live locally for 6 months or less. [c] The minimum wage in Guangzhou City is 2100 yuan per month and the minimum wage in Zhongshan City is 1720 yuan per month.

3.2. Univariate Analysis

A total of 28 potential risk factors were analyzed. These factors were further divided into six dimensions: Personal life activities, environmental sanitation, housing situation, living conditions, mosquito protection status, and residential surroundings. As illustrated in Table 2, people who participated in outdoor activities in parks had a significantly higher probability of DF infections than those who did not participate in outdoor activities in parks ($p = 0.049$). People who participated in outdoor sports were more likely to be infected with DF than those who did not participate in outdoor sports ($p = 0.009$). At the same time, there were statistical differences in terms of housing type ($p = 0.040$), housing location ($p = 0.061$), living floor ($p = 0.096$), the average numbers of persons per room ($p < 0.001$), air-conditioner use ($p = 0.026$) and indoor daylight quality ($p = 0.032$) between the case group and the control group.

Table 2. Univariate analysis of risk factors for dengue virus infection.

Variables	Cases (n/%) (n = 237)	Controls (n/%) (n = 237)	OR (95% CI)	p-Value
Contact with patients with dengue fever				0.655
Yes	3 (1.27)	2 (0.84)	1.51 (0.25–9.10)	
No	234 (98.73)	235 (99.16)	Reference	
Outbound tourism experience				0.867
Yes	19 (8.02)	20 (8.44)	Reference	
No	218 (91.98)	217 (91.56)	1.06 (0.55–2.04)	
Outdoor activities in parks				0.049 *
Yes	200 (84.39)	183 (77.22)	1.60 (1.00–2.54)	
No	37 (15.61)	54 (22.78)	Reference	
Participation in outdoor sports				0.009 *
Yes	74 (31.22)	49 (20.68)	1.74 (1.15–2.64)	
No	163 (68.78)	188 (79.32)	Reference	
Domestic sewage disposal frequency				0.655
Daily	125 (52.74)	118 (49.79)	Reference	
2 days	25 (10.55)	20 (8.44)	1.18 (0.62–2.24)	
≥3 days	35 (14.77)	42 (17.72)	0.79 (0.47–1.32)	
No domestic sewage	52 (21.94)	57 (24.05)	0.86 (0.55–1.35)	
Garbage disposal frequency				0.311
Daily	223 (94.09)	228 (96.20)	Reference	
2 days	11 (4.64)	5 (2.11)	2.25 (0.77–6.58)	
≥3 days	3 (1.69)	4 (1.69)	0.77 (0.17–3.47)	
Participation in community hygiene management interventions				0.104
Yes	77 (32.49)	94 (39.66)	Reference	
No	160 (67.51)	143 (60.34)	1.37 (0.94–1.99)	
Location				0.061 *
Rural	15 (6.33)	11 (4.64)	0.34 (0.08–1.51)	
City	210 (88.61)	223 (94.09)	0.24 (0.07–0.85)	
Urban–rural integration	12 (5.06)	3 (1.27)	Reference	
Housing building structure				0.871
Brick–wood structure	2 (0.84)	3 (1.27)	0.67 (0.11–4.05)	
Brick–wood and concrete structure	31 (13.08)	29 (12.24)	1.07 (0.63–1.85)	
Concrete structure	204 (86.08)	205 (86.49)	Reference	
Housing type				0.040 *
Single-family apartment	60 (25.32)	39 (16.46)	Reference	
Commercial residential community	173 (73.00)	196 (82.70)	0.57 (0.37–0.90)	
Villa	4 (1.68)	2 (0.84)	1.30 (0.23–7.44)	
Housing age (year)				0.919
<10	23 (9.71)	24 (10.13)	Reference	
10–20	115 (48.52)	116 (48.95)	1.03 (0.54–1.94)	
20–40	88 (37.13)	89 (37.55)	1.03 (0.54–1.96)	
>40	11 (4.64)	8 (3.37)	1.43 (0.49–4.21)	
Number of floors per residential structure				0.096 *
1–3	110 (46.41)	87 (36.71)	Reference	
4–9	103 (43.46)	124 (52.32)	0.66 (0.45–0.96)	
≥10	24 (10.13)	26 (10.97)	0.73 (0.39–1.36)	
Average numbers of persons per room				<0.001 *
1	91 (38.40)	51 (21.52)	Reference	
2	123 (51.90)	160 (67.51)	0.43 (0.28–0.65)	
≥3	23 (9.70)	26 (10.97)	0.50 (0.26–0.96)	
Housing area (m^2)				0.235
<50	63 (26.58)	52 (21.9)	Reference	
51–100	150 (63.29)	169 (71.31)	0.73 (0.18–1.12)	
101–150	19 (8.02)	11 (4.64)	1.43 (0.62–3.26)	
>150	5 (2.11)	5 (2.11)	0.83 (0.23–3.01)	
Air-conditioner use				0.013 *
Never	26 (10.97)	11 (4.64)	Reference	
Yes	211(89.03)	226 (95.36)	0.40 (0.19–0.82)	

Table 2. *Cont.*

Variables	Cases (*n*/%) (*n* = 237)	Controls (*n*/%) (*n* = 237)	OR (95% CI)	*p*-Value
Indoor daylight quality				0.032 *
Good	215 (90.72)	227 (95.78)	Reference	
Poor	22 (9.28)	10 (4.22)	2.32 (1.08–5.02)	
Ventilation				0.324
Good	221 (93.25)	226 (95.36)	Reference	
Bad	16 (6.75)	11 (4.64)	1.49(0.68–3.28)	
Keeping of pets				0.800
Yes	38 (16.03)	36 (15.19)	Reference	
No	199 (83.97)	201 (84.81)	0.94 (0.57–1.54)	
Raising of poultry				0.589
Yes	6 (2.53)	8 (3.38)	0.74 (0.25–2.18)	
No	231 (97.47)	229 (96.62)	Reference	
Breeding of aquatic plants				0.578
Yes	54 (20.68)	49 (20.68)	Reference	
No	183 (77.22)	188 (79.32)	0.88 (0.57–1.37)	
Use of mosquito nets				0.361
Yes	185 (78.06)	193 (81.43)	Reference	
No	52 (21.94)	44 (18.57)	1.23 (0.79–1.93)	
Use of mosquito repellent				0.212
Never	122 (51.48)	116 (48.95)	Reference	
Occasionally	88 (37.13)	103 (43.46)	0.81 (0.56–1.19)	
Often	27 (11.39)	18 (7.59)	1.43 (0.75–2.73)	
Use of electric mosquito-killing devices				0.150
Never	152 (64.14)	143 (60.34)	Reference	
Occasionally	57 (24.05)	74 (31.22)	0.73 (0.48–1.10)	
Often	28 (11.81)	20 (8.44)	1.32 (0.71–2.44)	
Use of camphor				0.649
Never	177 (74.68)	168 (70.89)	Reference	
Occasionally	42 (17.72)	48 (20.25)	0.83 (0.52–1.32)	
Often	18 (7.60)	21 (8.86)	0.81 (0.42–1.58)	
Existence of garbage collection sites within 200 m around housing				0.681
Yes	32 (13.50)	29 (12.24)	1.12 (0.65–1.92)	
No	205 (86.50)	208 (87.76)	Reference	
Existence of junk yards within 200 m around housing				0.570
Yes	1 (0.42)	2 (0.84)	0.50 (0.05–5.53)	
No	236 (99.58)	235 (99.16)	Reference	
Existence of ponds within 200 m around housing				0.426
Yes	45 (18.99)	52 (21.94)	0.83 (0.53–1.30)	
No	192 (81.01)	185 (78.06)	Reference	
Existence of construction sites within 200 m around housing				0.639
Yes	24 (10.13)	21 (8.86)	1.16 (0.63–2.15)	
No	213 (89.87)	216 (91.14)	Reference	

* Significance difference: $p < 0.1$.

The existence of garbage collection sites ($p = 0.681$), junk yards ($p = 0.570$), ponds ($p = 0.426$), and construction sites ($p = 0.639$) within 200 m of residences did not show statistically significant differences between the control and case groups.

3.3. Multivariate Analysis

Multivariate logistic regression analysis was performed on the basis of the results of univariate analysis and the rule of frequency matching design. In the unconditioned logistic regression model, participation in outdoor sports (OR = 1.80, 95% CI = 1.17 to 2.78) and poor indoor daylight quality (OR = 2.27, 95% CI = 1.03 to 5.03) were significantly associated with an increased risk of dengue virus infection. On the other hand, 2 occupants per room (OR = 0.43, 95% CI = 0.28 to 0.65), ≥3 occupants per room (OR = 0.45, 95% CI = 0.23 to 0.89), and air-conditioner use (OR = 0.46, 95% CI = 0.22 to 0.97) were significantly associated with protection against dengue virus infection (Table 3).

Table 3. Multivariate analysis of risk factors for dengue virus infection.

Risk factors	Odds Ratio	95% CI	*p*-Value
Participation in outdoor sports			0.007 *
Yes	1.80	(1.17–2.78)	
No	Reference		
Average numbers of occupants per room			<0.001 *
1	Reference		
2	0.43	(0.28–0.65)	<0.001 *
≥3	0.45	(0.23–0.89)	0.021 *
Air-conditioner use			0.040 *
Never	Reference		
Yes	0.46	(0.22–0.97)	
Indoor daylight quality			0.043 *
Good	Reference		
Poor	2.27	(1.03–5.03)	

* Significant difference when $p < 0.05$.

4. Discussion

We found that participation in outdoor sports activities and poor indoor daylight quality significantly increased the probability of contracting DF by 1.80- and 2.27-fold, respectively, in Guangdong Province. Our results also suggested that 2 occupants per room, ≥3 occupants per room, and air-conditioner use might decrease the probability of dengue virus infection by 0.43-, 0.45-, and 0.46-fold, respectively.

Our study revealed that people who participated in outdoor sports were at a significantly higher risk of contracting DF than those who did not participate in outdoor sports. This result may be attributed to the preference of residents who participate in outdoor sports to hike and camp in forest margins, tree copses, and natural reserves, which are the original habitats of *A. albopictus* [20]. Therefore, given that participation in outdoor sports increases the risk of exposure to mosquito bites, anti-mosquito measures, such as treating outdoor areas and materials with insecticides, must be adopted.

We also found that human population density was closely associated with dengue transmission. In general, high population density is a risk factor for dengue transmission [21,22]. We found, however, that crowded households comprising ≥2 occupants were at low risk of dengue infection. On the contrary, Velascosalas et al. found that crowded households with more than 1.5 occupants in one room were at risk of dengue infection [23]. We conducted our study in communities wherein individuals lived together in family groups and wherein parents and their young children tended to share one room. Wang et al. reported that 62.48% of Chinese children aged 0–5 years old shared beds with their parents [24]. Our result may be attributed to the following: Parents who share rooms with their children pay additional attention to the use of anti-mosquito measures and the maintenance of good environmental sanitation to protect their children from mosquito bites. Our previous study also showed that married participants had a lower rate of infection than widowed and divorced participants [19]. These results suggest that married groups who reside in one room with ≥2 occupants are at a reduced risk of DF infection. Moreover, high numbers of occupants in one room are associated with the decreased probability of mosquito bites when the number of mosquitoes was fixed.

Shen et al. [25] and Wu et al. [26] reported that yearly average temperatures of more than 18 °C would increase the risk of dengue virus infection. Meanwhile, our study indicated that air-conditioner use was a protective factor against dengue infection by reducing the risk of dengue transmission through cooling the indoor environment. In addition, doors and windows are commonly shut during air-conditioner use; this practice could also reduce the chance that mosquitoes could enter the rooms [27].

We found that the poor indoor daylight quality increased the likelihood of infection with DF by 2.27-fold because adult *A. albopictus* prefer to inhabit poorly lit areas over well-lit areas [28,29]. As a result, environments receiving insufficient daylight encourage the density of mosquitoes to increase because they are suitable for the survival of mosquitoes.

Vanwambeke et al. [30] and Kenneson et al. [31] reported that the use of mosquito nets reduced the risk of dengue virus infection. However, similar to Tsuzuki et al. [32] and Loroñopino et al. [33], we failed to find a relationship between the use of mosquito nets and the likelihood of dengue virus infection. The lack of a relationship between this variable and dengue virus infection may be attributed to the following: Mosquito nets are usually used at night. However, *A. albopictus* is active during the day, especially in the early morning and late afternoon [34]. Other studies found that mosquito nets play a protective role in preventing dengue virus infection in rural settings [29]. However, we recruited our study population from urban areas. The good living environment in urban areas, for example, the popularity of air-conditioning and mosquito killing facilities, reduced the demand for mosquito nets. Nevertheless, on the basis of our local experience and some studies' results [30,35], we recommend using mosquito nets not only at night but also during the day [31].

Andersson et al. [36] and Roberto et al. [11] reported that the government's ability and capacity to control the dengue vector has crucial effects on dengue transmission. Community neighborhood committees and property management departments in Guangdong Province have organized numerous health remediation activities under the supervision of the relevant health agent or the Centers for Disease Control and Prevention [37] given the high incidence of DF in recent years. These activities have considerably improved the residential living environment and reduced mosquito breeding and may account for the lack of the statistical significance of the variables of domestic sewage disposal, garbage management, and residential surroundings in this study. However, we found that 67.51% of the cases and 60.39% of the controls did not participate in the community hygiene management intervention activities organized by neighborhood committees or property management organizations. The public health consciousness of the case and control groups must be strengthened because the government will be unable to establish a sound prevention system against DF despite having a good macro-control system in place if it lacks the support of the masses.

5. Limitations

This study has several flaws that should be overcome in future studies. First, cases and controls were identified in accordance with the results of antibody detection. However, IgG-positive samples may have been derived from patients who were infected with DF several years ago. Thus, the results of the completed questionnaires of these patients may be incompatible with their situations during their initial infection. Second, recall biases stemming from the inaccuracy of memory could also reduce the validity of the questionnaire. Third, newly infected individuals may have been misdiagnosed as controls because IgG and IgM antibody titers are present at undetectable levels during the initial stages of infection with dengue [38]. Thus, misclassification bias may have been introduced. Fourth, volunteer bias might have been caused by participants who volunteered to become part of the samples. Additionally, the information provided by the volunteered participants might have been different from the general population. Fifth, some definitions like the definitions of "good indoor daylight quality" and "good ventilation" were given in the questionnaire, but respondents might have had a subjective understanding of these definitions, which could have led to inaccuracy of the relevant information. Finally, the model's general applicability was difficult to evaluate at present given that the data used in this study were obtained on the basis of serum samples collected from residents in Guangdong Province over the period of 2013 to 2015 without other extrapolated studies. Given that other studies have shown that the community wherein residents live remains an independent factor associated with DF infection [23], the results of this study could be used as the reference for the development of personalized protective measures against DF infection in tropical and subtropical countries with high densities of *Aedes* mosquitoes.

6. Conclusions

Our study focused on residents in communities with mild or asymptomatic dengue virus infection rather than on patients with severe clinical symptoms to explore the risk factors for dengue virus infection in Guangdong Province because the former sample is highly representative. We established the relationship between dengue infection and individual risk factors. This information is beneficial for avoiding infection by the dengue virus. We also provided evidence and a basis for the development of measures for DF prevention and control. Additional variables must be introduced into the logistic regression model, and further research should be conducted to provide a theoretical basis for formulating prevention and control measures for DF.

Supplementary Materials: The following are available online at http://www.mdpi.com/1660-4601/16/4/617/s1, File 1: The contents of the questionnaire.

Author Contributions: L.J. drafted the manuscript and conducted the data-cleansing process and further data analysis. T.X. collected the questionnaires and performed the questionnaire data entry. D.Y. performed laboratory experiments and the follow-up questionnaire survey. D.Z. conducted initial data analysis. L.T. performed laboratory experiments. H.Y. provided the database of serum and financial resources. Z.D. designed the research, supervised all the work and checked the manuscript. All authors re-examined and approved the final manuscript.

Funding: This work was supported by the National Science and Technique Major Project of China (grant number 2018ZX10715004).

Acknowledgments: We are grateful for the assistance offered by the National Science and Technique Major Project in sample collection. We also appreciate the research assistance that was offered by the Health and Family Planning Commission in Yuexiu District and Liwan District, Yuexiu District Centre for Disease Control and Zhongshan Centre for Disease Control. We thank the investigators, such as staff and students in the School of Public Health for providing field and laboratory assistance, and all subjects.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. World Health Organization. Epidemiology. Available online: https://www.who.int/denguecontrol/epidemiology/en/ (accessed on 14 May 2018).
2. Gubler, D.J.; Meltzer, M. Impact of dengue/dengue hemorrhagic fever on the developing world. *Adv. Virus Res.* **1999**, *53*, 35–70. [CrossRef] [PubMed]
3. Wen, J.; Liang, F. An etiological and serological study on Dengue epidemic in Guangdong Province. *Zhonghua Shi Yan He Lin Chuang Bing Du Xue Za Zhi* **1998**, *12*, 165–168. [PubMed]
4. Liu, C.; Liu, Q.; Lin, H.; Xin, B.; Nie, J. Spatial analysis of dengue fever in Guangdong Province, China, 2001–2006. *Asia Pac. J. Public Health* **2014**, *26*, 58. [CrossRef] [PubMed]
5. Zhang, Y.; Wang, T.; Liu, K.; Xia, Y.; Lu, Y.; Jing, Q.; Yang, Z.; Hu, W.; Lu, J. Developing a Time Series Predictive Model for Dengue in Zhongshan, China Based on Weather and Guangzhou Dengue Surveillance Data. *PLoS Neglect. Trop. Dis.* **2016**, *10*, e0004473. [CrossRef] [PubMed]
6. Guo, Y.; Song, Z.; Luo, L.; Wang, Q.; Zhou, G.; Yang, D.; Zhong, D.; Zheng, X. Molecular evidence for new sympatric cryptic species of Aedes albopictus (Diptera: Culicidae) in China: A new threat from Aedes albopictus subgroup? *Parasite Vectors* **2018**, *11*, 228. [CrossRef] [PubMed]
7. Jia, D.S.; Tan, W.L.; Wang, C.J.; Han, Z.J.; Lu, N.H.; Zheng, J. Aedes-borne diseases prevention and control. *Chin. J. Hyg. Insect. Equip.* **2017**, *23*, 1–7.
8. Thein, T.L.; Leo, Y.S.; Fisher, D.A.; Low, J.G.; Oh, H.M.L.; Gan, V.C.; Wong, J.G.X.; Lye, D.C. Risk Factors for Fatality among Confirmed Adult Dengue Inpatients in Singapore: A Matched Case-Control Study. *PLoS ONE* **2013**, *8*, e81060. [CrossRef]

9. Pang, J.; Thein, T.L.; Leo, Y.S.; Lye, D.C. Early clinical and laboratory risk factors of intensive care unit requirement during 2004–2008 dengue epidemics in Singapore: A matched case-control study. *BMC Infect. Dis.* **2014**, *14*, 649. [CrossRef]
10. Moraes, G.H.; De, F.D.E.; Duarte, E.C. Determinants of mortality from severe dengue in Brazil: A population-based case-control study. *Am. J. Trop. Med. Hyg.* **2013**, *88*, 670–676. [CrossRef]
11. Branco, M.R.; Luna, E.J.; Braga Júnior, L.L.; Oliveira, R.V.B.; Rios, L.T.M.; Silva, M.S.; Medeiros, M.N.L.; Silva, G.F.; Nina, F.C.A.F.; Lima, T.J.; et al. Risk factors associated with death in Brazilian children with severe dengue: A case-control study. *Clinics* **2014**, *69*, 55–60. [CrossRef]
12. Van Benthem, B.H.; Khantikul, N.; Panart, K.; Kessels, P.J.; Somboon, p.; Oskam, L. Knowledge and use of prevention measures related to dengue in northern Thailand. *Trop. Med. Int. Health* **2010**, *7*, 993–1000. [CrossRef]
13. Phuanukoonnon, S.; Mueller, I.; Bryan, J.H. Effectiveness of dengue control practices in household water containers in Northeast Thailand. *Trop. Med. Int. Health* **2005**, *10*, 755–763. [CrossRef] [PubMed]
14. Luo, L.; Liang, H.Y.; Hu, Y.S.; Liu, W.J.; Wang, Y.L.; Jing, Q.L.; Zheng, X.L.; Yang, Z.C. Epidemiological, virological, and entomological characteristics of dengue from 1978 to 2009 in Guangzhou, China. *J. Vector Ecol.* **2012**, *37*, 230–240. [CrossRef] [PubMed]
15. Gu, H.; Leung, K.K.; Jing, Q.; Zhang, W.; Yang, Z.; Lu, J.; Hao, Y.; Zhang, D. Meteorological Factors for Dengue Fever Control and Prevention in South China. *Int. J. Environ. Res. Public Health* **2016**, *13*, 867. [CrossRef] [PubMed]
16. Fan, J.C.; Lin, H.L.; Wu, H.X.; Wang, J.; Yang, S.R.; Liu, Q.Y. Spatial and temporal distribution characteristics of dengue fever in Guangdong province, China during 2006-2011. *Chin. J. Vector Biol. Control* **2013**, *24*, 389–391. [CrossRef]
17. Wang, T.; Wang, M.; Shu, B.; Chen, X.; Luo, L.; Wang, J.; Cen, Y.; Anderson, B.D.; Merrill, M.M.; Mwrrill, H.R.; et al. Evaluation of Inapparent Dengue Infections During an Outbreak in Southern China. *PLoS Neglect. Trop. Dis.* **2015**, *9*, e0003677. [CrossRef] [PubMed]
18. Zhang, J.H.; Yuan, J.; Wang, T. Direct cost of dengue hospitalization in Zhongshan, China: Associations with demographics, virus types and hospital accreditation. *PLoS Neglect. Trop. Dis.* **2017**, *11*, e0005784. [CrossRef]
19. Liu, J.; Deng, Y.; Jing, Q.; Chen, X.; Du, Z.; Liang, T.; Yang, Z.; Zhang, D.; Hao, Y. Dengue Infection Spectrum in Guangzhou: A Cross-Sectional Seroepidemiology Study among Community Residents between 2013 and 2015. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1227. [CrossRef]
20. Rezza, G. Aedes albopictus and the reemergence of Dengue. *BMC Public Health* **2012**, *12*, 72. [CrossRef]
21. Gubler, D.J. Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century. *Trends Microbiol.* **2002**, *10*, 100–103. [CrossRef]
22. Jing, Q.L.; Li, Y.L.; Chen, Z.Q. Spatial pattern and driving factors of dengue virus-4 epidemics in Guangzhou. *J. Trop. Med.* **2016**, *16*, 244–248.
23. Velascosalas, Z.I.; Sierra, G.M.; Guzmán, D.M.; Zambrano, J.; Vivas, D.; Comach, G.; Wilschut, J.C.; Tami, A. Dengue Seroprevalence and Risk Factors for Past and Recent Viral Transmission in Venezuela: A Comprehensive Community-Based Study. *Am. J. Trop. Med. Hyg.* **2014**, *91*, 1039–1048. [CrossRef]
24. Wang, H.S.; Huang, X.N.; Jiang, J.X.; Ma, Y.Y.; An, L.; Liu, X.C. Sleep location in Chinese children aged 0~5 years old. *Chin. J. Child Health Care* **2008**, *16*, 420–422.
25. Shen, C.; Luo, L.; Li, L.; Jing, Q.L.; Ou, C.Q.; Yang, Z.C.; Chen, X.G. The Impacts of Mosquito Density and Meteorological Factors on Dengue Fever Epidemics in Guangzhou, China, 2006–2014: A Time-series Analysis. *Biomed. Environ. Sci.* **2015**, *28*, 321–329. [CrossRef]
26. Wu, P.C.; Lay, J.G.; Guo, H.R.; Lin, C.Y.; Lung, S.C.; Su, H.J. Higher temperature and urbanization affect the spatial patterns of dengue fever transmission in subtropical Taiwan. *Sci. Total Environ.* **2009**, *407*, 2224–2233. [CrossRef]
27. Paul, R.; Sarah, L.; Michel, B.; Biggerstaff, B.; Singer, D.; Tiwari, T.; Baber, L.; Amador, M.; Thirion, J.; Hayes, J.; et al. Texas Lifestyle Limits Transmission of Dengue Virus. *Emerg. Infect. Dis.* **2003**, *9*, 86–89. [CrossRef]
28. Gong, D.F.; Zhou, H.N. Progress in Dengue fever important vector Aedes albopictus in China. *Chin. J. Vector Biol. Control* **2009**, *20*, 607–610.
29. Crepeau, T.N.; Healy, S.P.; Bartlett-Healy, K.; Unlu, I.; Farajollahi, A.; Fonseca, D.M. Effects of Biogents Sentinel Trap Field Placement on Capture Rates of Adult Asian Tiger Mosquitoes, Aedes albopictus. *PLoS ONE* **2013**, *8*, e60524. [CrossRef]

30. Vanwambeke, S.O.; van Benthem, B.H.; Khantikul, N.; Burghoorn-Maas, C.; Panart, K.; Oskam, L.; Lambin, E.F.; Somboon, P. Multi-level analyses of spatial and temporal determinants for dengue infection. *Int. J. Health Geogr.* **2006**, *5*, 5. [CrossRef]
31. Kenneson, A.; Beltrán-Ayala, E.; Borbor-Cordova, M.J.; Polhemus, M.E.; Ryan, S.J.; Endy, T.P.; Stewart-Ibarra, A.M. Social-ecological factors and preventive actions decrease the risk of dengue infection at the household-level: Results from a prospective dengue surveillance study in Machala, Ecuador. *PLoS Neglect. Trop. Dis.* **2017**, *11*, e0006150. [CrossRef]
32. Tsuzuki, A.; Thiem, V.D.; Suzuki, M.; Yanai, H.; Matsubayashi, T.; Yoshida, L.-M.; Tho, L.H.; Minh, T.T.; Anh, D.D.; Kilgore, P.E.; et al. Can Daytime Use of Bed Nets Not Treated with Insecticide Reduce the Risk of Dengue Hemorrhagic Fever Among Children in Vietnam? *Am. J. Trop. Med. Hyg.* **2010**, *82*, 1157–1159. [CrossRef]
33. Loroñopino, M.A.; Machainwilliams, C.; Gomezcarro, S.; Loroño-Pino, M.A.; García-Rejón, J.E.; Machain-Williams, C.; Gomez-Carro, S.; Nuñez-Ayala, G.; Nájera-Vázquez, M.R.; Losoya, A.; et al. Towards a Casa Segura: A Consumer Product Study of the Effect of Insecticide-Treated Curtains on Aedes aegypti and Dengue Virus Infections in the Home. *Am. J. Trop. Med. Hyg.* **2013**, *89*, 385–397. [CrossRef]
34. Trpis, M.; Mcclelland, G.A.; Gillett, J.D.; Teesdale, C.; Rao, T.R. Diel periodicity in the landing of Aedes aegypti on man. *Bull. World Health Organ.* **1973**, *48*, 623–629.
35. Lenhart, A.; Orelus, N.; Maskill, R.; Alexander, N.; Streit, T.; McCall, P.J. Insecticide-treated bednets to control dengue vectors: Preliminary evidence from a controlled trial in Haiti. *Trop. Med. Int. Health* **2008**, *13*, 56–67. [CrossRef]
36. Andersson, N.; Navaaguilera, E.; Arostegui, J.; Morales-Perez, A.; Suazo-Laguna, H.; Legorreta-Soberanis, J.; Hernandez-Alvarez, C.; Fernandez-Salas, I.; Paredes-Solís, S.; Balmaseda, A.; et al. Evidence based community mobilization for dengue prevention in Nicaragua and Mexico (Camino Verde, the Green Way): Cluster randomized controlled trial. *BMJ* **2015**, *351*, h3267. [CrossRef]
37. Pan, H.F.; YE, X.X.; Chen, C.J.; Zhao, J.Y.; Lin, Z.Y. Study on prevention and control of Dengue from perspective of public health. *Chin. J. Public Health Manag.* **2015**, *31*, 775–776. [CrossRef]
38. Guzman, M.G.; Harris, E. Dengue. *Lancet* **2015**, *385*, 453–465. [CrossRef]

International Journal of
Environmental Research and Public Health

MDPI

Article

Social Capital and Mental Health in Rural and Urban China: A Composite Hypothesis Approach

Xiaoming Lin [1,†], Ruodan Lu [2,3,†], Liang Guo [1,*,†] and Bing Liu [4,*]

1 Department of Statistics, Shandong University, Weihai 264209, China; xiaoming.lin@mail.sdu.edu.cn
2 Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, UK; rl508@cam.ac.uk
3 School of Architecture, Building and Civil Engineering, Loughborough University, Loughborough LE11 3TU, UK
4 School of Management, Shandong University, Jinan 250100, China
* Correspondence: liangguo@sdu.edu.cn (L.G.); liubing@sdu.edu.cn (B.L.)
† The first three authors equally contribute to this article.

Received: 15 January 2019; Accepted: 19 February 2019; Published: 25 February 2019

Abstract: The objective of this study is to follow the composite theory approach to analyze the effect of social capital on self-rated mental health in rural and urban China. Our nationally representative sample includes 10,968 respondents from 130 county-level communities. Two-level random-coefficient linear regressions, which model individual and community variations in subjective mental health, were estimated by taking the hierarchical structure of the dataset into account. We found that a significant proportion of the total variations in self-rated mental health were explained at the community level. We also found an association between low contextual civic trust and poor self-rated mental health after adjusting for individual social capital and individual socioeconomic-demographic variables. The study also revealed that: (1) in rural areas a positive relationship between civic and political trust and mental health existed both at the individual and the community level, respectively; and (2) in urban areas, only political trust at the individual level contributed to better mental health. In addition, the individual and community level political participation exhibited a positive impact on mental health measures in both rural and urban China. The individual level civic participation was positively associated to the outcome variable. However, the community-level civic participation seemed to negatively impact mental health in urban area. Our findings emphasize the importance of both individual and community-level healthcare interventions in China. Finally, this study also found that human capital covariates remained important predictors of self-rated mental health status even after controlling social capital both at individual and community levels. This study suggested that the composite thesis could provide a more convincing narrative than other theories in explaining the effects of both human and social capital on health.

Keywords: China; human capital; multilevel linear regression; self-rated mental health; social capital

1. Introduction

As proposed by Sartorius [1] (p. 101), mental health is a ‘’the state of balance that individuals establish within themselves and between themselves and their social and physical environment”. As per the report by Global Forum for Health Research and World Health Organization [2], mental and neurological disorders account for 13% of the global burden of disease. According to Zhang [3] and National Health and Family Planning Commission [4], 2.7% of the population in China was afflicted with severe mental illness in 1950s, a figure which rose to 5.4% in the 1970s, 11.1% in the 1980s, 13.47% in the 1990s, and further to 15% in the 2010s. Currently, over 30 million people have a diagnosis of severe mental illness (especially depression) in China [4]. Mental illness has been found to be the prime cause of the high suicide rate in China, which stands at approximately 10 per 100,000 in 2016 [5]. These statistics call for an urgent need to deepen our understanding on the determinants of Chinese mental health [6]. Research indicates that mental health is determined by socioeconomic, environmental, intersectoral, and civic security factors [7]. Mental health policy and practice remain pivotal to the discussion on social capital [8], because China's surge in mental illness problems is probably due to its rapid urbanization, which is usually associated with a range of social health hazards and risks that can lead to the development of neuropsychiatric illness [4]. In fact, prior studies all over the world report that social capital is a crucial determinant of both physical and mental health issues at the individual and the aggregate level and their interactions [9–19]. Social capital can influence an individual's health by opening information channels, promoting collective action, addressing detrimental cultural norms, and fostering the development of support systems serving as a source of self-esteem and mutual respect [14,20]. These mechanisms of social capital also work in the context of mental health [7,8].

However, most prior studies were focused on western developed economies and presented mixed findings [21]. Research in the transition economies alluded to a positive relationship between social capital and improved health [22]. As argued by Ichida et al. [23], the relationship between high social capital and good health indicators can vary owing to the cultural differences among countries, justifying the need for a systematic study in the different cultural contexts. Recently, relevant research has been conducted in East Asia (for example by Fujisawa et al. [24]; Suzuki et al. [25]; Miller et al. [26], Yip et al. [21], Sun et al. [27]; and Yamaoka [28]). Although these works have advanced our understanding of the social capital-general health nexus, little attention has been particularly devoted to compare how the influence of social capital on mental health behaviors. To our best knowledge, there is little research addressing how social capital influences mental health differently in Chinese urban and rural settings. Such study is important, as prior works indicate that there is a huge urban-rural health disparity in China and call for further studies [29–32].

The major objective of this study is to fill the above-mentioned gap. Specifically, we examined empirical relationships of four dimensions of social capital, i.e., civic participation, civic trust, political participation, political trust, on self-rated mental health. We followed the composite hypothesis, which argued that the mental health of an individual was influenced not only by his or her human capital (i.e., education, income and social status) but also by his or her social capital, as people are usually involved in a variety of analytically distinct but nonexclusive social networks [33–35]. We also followed the multi-level framework proposed in [36]. By community, we referred to counties (*xian* in Chinese) or districts (*qu* in Chinese), the third level of China's administrative hierarchy. Our sample individuals were nested within communities and an individual's mental health status was a function of a set of individual covariates and human capital (i.e., gender, age, education, marriage, unemployment, social-economic status, household annual income, household wealth and household size), individual-level social capital (i.e., civic trust, civic participation, political trust, political participation), and community-level social capital. The multilevel linear regression technique was chosen as the most appropriate design tool to explain the variation in mental health status of the population by both within-community and between-community differences. To our best knowledge, the sample size of our study is the largest of its kind at the individual and community levels.

2. Materials and Methods

2.1. Sample

We used the 2015 sociological survey dataset provided by China's National Survey Research Center (the dataset can be found at http://cnsda.ruc.edu.cn/). Data were collected using a stratified clustered sampling design and were representative for the non-institutionalized adult (i.e., 18 and above) population in China. The primary survey unit (PSU) of this survey is at the county level or district level. Using China's latest census as the sampling frame, the survey project included 969 of the most crowded PSUs across China. One hundred and thirty PSUs were randomly chosen and each chosen PSU was assigned a quota of households based on the population of that neighborhood as a proportion of the total population of all 969 PSUs. The researchers then selected households randomly using the residence records in the sampling frame. One adult person was randomly selected from each sampled household to serve as a respondent. Survey administrators visited each household after 18:00 h on weekdays or after 14:00 h during weekends and holidays to maximize the participation rate. The dataset consists of 10,968 face-to-face interviews (6470 respondents from rural and 4498 from urban areas, respectively). The sample includes 5834 female and 5134 male individuals from 130 PSUs (hereafter, "communities") in 28 provinces. Most of the respondents were aged 41–60 (40%). Our samples were nested in two levels: the individual level and the community (i.e., county or district) level. The detailed distributions of sample are summarized in Table 1.

Table 1. Descriptions of the sample.

Provinces/Regions	Urban	Rural	District/County (Community)	Male	Female	Avg. Age (SD)
Shanghai	502	0	8	236	266	54.89(17.93)
Yunnan	93	292	4	190	195	48.70(16.30)
Inner Mongolia	25	74	1	53	46	54.12(15.08)
Beijing	519	28	9	252	295	48.70(17.62)
Jilin	178	287	5	222	243	48.41(16.37)
Sichuan	275	291	6	268	298	55.12(16.04)
Tianjin	288	0	7	135	153	52.40(16.98)
Ningxia Hui	47	47	1	36	58	38.64(14.07)
Anhui	119	278	4	182	215	51.62(16.54)
Shandong	315	260	6	262	313	48.38(17.09)
Shanxi	189	91	3	137	143	44.84(16.63)
Guangdong	531	0	10	254	277	46.24(16.99)
Guangxi Zhuang	194	199	4	206	187	52.32(16.05)
Jiangsu	321	178	5	231	268	50.82(17.77)
Jiangxi	284	192	5	213	263	49.49(17.90)
Hebei	99	196	4	120	175	47.07(15.63)
Henan	216	366	6	259	323	50.97(17.36)
Zhejiang	341	121	5	202	260	52.22(18.09)
Hubei	350	250	6	274	326	51.91(15.41)
Hunan	240	235	5	222	253	52.80(17.24)
Gansu	50	145	2	88	107	49.47(15.52)
Fujian	194	100	3	147	147	50.51(16.80)
Guizhou	177	72	3	110	139	48.13(16.90)
Liaoning	345	50	4	189	206	49.49(17.27)
Chongqing	79	186	3	132	133	52.95(16.31)
Shaanxi	106	263	4	173	196	49.78(15.05)
Qinghai	75	26	1	51	50	49.67(14.30)
Heilongjiang	318	271	6	290	299	50.54(15.28)
Total	6470	4498	130	5134	5834	50.40(16.89)

2.2. Outcome Variable: Self-rated Mental Health

Self-rated mental health status was based on five statements that are similar to the WHO-Five Well-Being Index (WHO-5): "I often feel calm and peaceful; I often feel active and vigorous; I often feel depressive and unhappy; I often feel exhausted and tired; and I often feel that I cannot stand my life". For the first two statements, the responses were given a value from (1) "not at all" to (5) "very true". As for the last three statements, the responses were given a value from (1) "very true" to (5) "not at all". The Cronbach's alpha was 0.86. The eigenvalue of one principal component of this scale is 3.29, which represents the significant amount of variance in each item that can be explained by the principal component.

The items were evaluated by a confirmatory factor analysis (CFA) model. A weighted average mental health score was generated, which was resulting from the multiplication of each statement by its corresponding CFA standardized regression weight. The mental health score was standardized to have a mean of zero and a standard deviation of one. That is, if an individual is in good mental health status, then his or her mental health score is positive. Otherwise, it is close the zero or negative.

2.3. Independent Variables: Social Capital

Social capital is regarded as "connections among individuals in social networks and norms of reciprocity and trustworthiness that arise from them" [37] (p. 19). The literature has identified four dimensions of social capital:

Structural social capital refers to an individual's social network and various forms of civic engagement while cognitive social capital means an individual's subjective perception of level of trust and reciprocity, which can be regarded as the result of structural social capital [38–42]. Horizontal social capital (or bonding/bridging social capital) refers to the relations developed between individuals or groups at the same social-economic hierarchical level, while vertical social capital (or linking social capital) includes the vertical relations between individuals or groups at different levels of formal power or authority [39,43–45]. Finally, social capital is not just an individual's resource [46–48] but also a contextual capital [39]. Community or neighborhood social capital is usually based on day-to-day interaction between individuals [49]. According to [50,51] and [52,53], an individual in a particular area is also exposed to the community-level context, which may have an effect on an individual's health. Therefore, we included contextual social capital variables, which can be constructed by grouping individuals within the same county/district and by aggregating an individual's answers.

In this study, we have adopted the four-dimension framework of [39] to measure both individual and community-level social capital in the Chinese context.

The first dimension is civic participation (CP), a type of structural and horizontal social capital, here assessed by the frequency of an individual's participation in nine different civic activities over the past 12 months (i.e., festivals, kinship networking, sports, entertainments, time with fellow workers, skill development, philanthropy, educational and religious activities). The responses were given a value from (1) "not at all" to (5) "very frequently". The Cronbach's alpha was 0.74. The eigenvalue of one principal component of this scale is 3.37, indicating a substantially large amount of variance in each item that can be explained by the principal component. The items were standardized to have a mean of zero and a standard deviation of one, and were then evaluated by a CFA model. A weighted average CP score was generated, which was resulting from the multiplication of each statement by its corresponding CFA standardized regression weight.

The second dimension is civic trust (CT), a type of cognitive and horizontal social capital, here measured by the degree of a respondent's trust on people, based on 13 different types of relations (i.e., close neighbors, distant neighbors, people living in the same town/village with the same family name, people living in the same town/village with different family name, relatives, acquaintances, townsmen, members of gymnasium club, members of non-political associations, members of volunteer organizations, classmates, colleagues, and total strangers). For each statement, the responses were given a value from (1) "not true at all" to (5) "very true". The Cronbach's alpha was 0.81. The eigenvalue of one principal component of this scale is 4.03. The items were standardized to have a mean of zero and a standard deviation of one, and then evaluated by a confirmatory factor analysis (CFA) model.

The third level is political participation (PP), a type of structural and vertical social capital, here assessed using four binary items—whether the respondent was a member of the Chinese Communist Party (CCP) or any other political party, or a government official or military officer, or a member of People's Congress (PC) or of People's Political Consultative Conference (PPCC), or had participated in any political protests or petitions. The sum of these four items was used as the PP score, which was then standardized to have a mean of zero and a standard deviation of one.

Fourthly, political trust (PT), a type of cognitive and vertical social capital, referred to the extent to which respondents trusted in China's political institutions (i.e., CCP, PC, PPCC, court system, People's Procuratorate, military, police, and local as well as central governments). PT was measured by nine statements in terms of a respondent's satisfaction of political institution's service performance on transparency and accountability, social inequality, healthcare disparity, social services for the elderly/poor households, education service, judiciary's efficiency and fairness, crime crackdown, defense and policing, and environment protection. For each statement, the responses were given a value from (1) "not satisfied at all" to (5) "very satisfied". The Cronbach's alpha was 0.80. The eigenvalue of one principal component of this scale is 4.71. The items were standardized to have a mean of zero and a standard deviation of one, and then evaluated by a confirmatory factor analysis (CFA) model.

Confirmatory factor analysis was performed on the constructs of CP, CT, PT, and mental health using open source R statistics package (see www.r-project.org) to evaluate the factor structures. The results indicated a good fit of each construct yielding a root mean square error of approximation (RMSEA) smaller than 0.05 and goodness-of-fit index (GFI) as well as an adjusted GFI superior to 0.90. All items were significantly loaded on their latent variables. The internal consistency of the scales (i.e., Cronbach's alpha) was greater than 0.70, indicating high reliability of our measures.

Thereafter, we aggregated these four social capital subcomponents into one general social capital index (SCI). Finally, five contextual, district-level social capital variables (CP-Dist, CT-Dist, PP-Dist, PP-Dist, PT-Dist, and SCI-Dist) at the community level were measured on the basis of aggregated individual answers for each county/district.

2.4. Independent Variables: Human Capital Covariates

Following [54] proposition, we have included a series of socioeconomic-demographic variables to control the confounding effects.

Gender was represented by a dummy variable for female. Age was categorized into four categories (1 = 18–25; 2 = 26–40; 3 = 41–60, and 4 = 61 and above). Education was measured as the highest grade completed (0 = no formal education; 1 = primary school; 2 = junior high school; 3 = senior high school; 4 = college; 5 = postgraduate). Marriage was categorized into three groups (1 = never married; 2 = married; 3 = divorced or widows/widowers). Unemployment was a dummy variable for whether one had been unemployed for the last three months. Self-reported social-economic status was assessed by a five-point likelihood scale ranging from 1 (very low) to 5 (very high). The respondents were asked to compare their social-economic status with peers within the same community.

Household income was measured based on the household annual expenditure (see [21]). Wealth index was constructed from a factor analysis of 15 terms related to household dwelling characteristics and ownership of consumer durables (such as, cars, TV sets, telephones, and so on. See [55] for more details. Finally, household size referred to the total number of family members that were living together.

2.5. Research Model

Previous study dichotomized output variable and used binominal logistic regression to estimate the effect of social capital. The benefit of the approach was to facilitate the interpretation of regression coefficient as the likelihood to report better or worse health condition. However, [39,56,57] argued that such a dichotomization of a continuous variable was often arbitrary, which could fail to provide complete information about the actual distribution of the dependent variable or even cause misleading results. Therefore, in this study, we have kept the outcome variable, i.e., self-rated mental health as continuous and used a series of two-level linear regressions to examine the effect of individual and community social capital on an individual's mental health. Following the approaches of [58,59], and [21], we have specified the following basic model:

$$MH_{ij} = \beta_{0j} + \beta_1 \left(SC_{ij} - SC_j\right) + \beta_2 SC_j + \beta_3 X_{ij} + \mu_{0j} + \varepsilon_{ij} \quad (1)$$

where MH is the mental health status for individual i (level 1) in community j (level 2); SC is the set of social capital variables measured at both the individual and community levels; and X is a vector of socioeconomic-demographic variables. This model estimated the fixed β parameters, which represented the overall relationships between individual social capital variables, covariates and mental health. There were also two random parameters μ_{0j} (community level) and ε_{ij} (individual level), which were assumed to follow a normal distribution and represent the differences from the corresponding means at both individual and community levels. The variance between-community and the variance between-individual within a given community were used to calculate the variance partition coefficient (VPC), which represented the deviation in mental health across individual and community levels [60]. That is to say, the proportion of the total residual variation that is due to differences between communities. As the community-level social capital variables were aggregated from those at the individual level, the value of individual social capital was group-centered in order to alleviate the multicollinearity problem (i.e., $SC_{ij} - SC_j$) [21,61,62].

Our research model is illustrated in Figure 1.

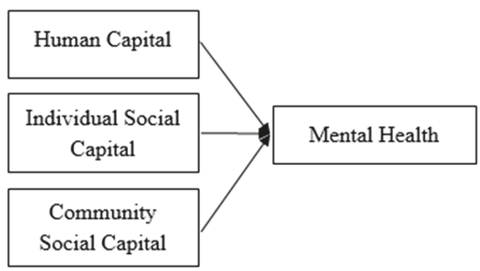


Figure 1. Research model.

We have followed the nested modeling strategy of previous multilevel health studies (e.g., [16,21,42,56,63]) to account for the hierarchical structure of our dataset. We first estimated two variance components models (Model 0-Rural and 0-Urban) that include only random intercepts. Then we estimated Model 1-Rural and 1-Urban to include random intercepts and human capital variables. Finally, we estimated Models 2a–2e that include random intercepts, human capital variables, the five individual-level social capital variables (i.e., the aggregated SCI, CP, CT, PP, and PT) and the five community-level social capital variables (i.e., the aggregated SCI-Dist, CP-Dist, CT-Dist, PP-Dist, and PT-Dist), respectively.

3. Results

Table 1 summarizes the means (standard deviations) and correlations of all variables. We found that people in urban areas reported slightly better mental health status than their rural counterparts as indicated by mean values of 0.05(SD = 0.97) and –0.07(SD = 1.04), respectively. Urban individuals also outperform rural ones in terms of wealth index, household income, and education. However, urban individuals reported lower social-economic status, smaller household size and higher unemployment rate. In terms of social capital, it seems that urban people reported much higher civic participation (0.37 vs. −0.53) but much lower civic trust (−0.18 vs. 0.25), political participation (−0.21 vs. 0.3), and political trust (−0.07 vs. 0.10) than rural ones. In addition, it seems that mental health status is positively associated with all four types of social capitals and married, education, household income, household wealth and social-economic statute. However, mental health status is negatively associated with female and age.

Regression results of Model 0, Model 1, and Model 2a–2e are summarized in Table 2. Both the individual-level variances σ_{ε}^2 and the community-level variance σ_{μ}^2 were all statistically significantly different to zero. The VPCs for the rural subsample ranged from 9.76% to 14.56%. And the VPCs for urban subsample were range from 9.76%. The VPCs indicate that substantial amount of vairance in mental health can be attributed to differences between communities. We then conducted the likelihood tests (see Table 3). We compared the loglikelihoods between Model 1 and 0. The differences (Δ-2ll) with 18 degrees of freedom were statistically significant under the Chi-square tests, suggesting that human capital variables as a whole are important determinants of mental health status in both rural and urban areas. Likewise, we compared Model 2a–2e to Model 1-Rural/1-Urban, respectively. The differences (Δ-2ll) with two degrees of freedom were statistically significant under the chi-squared tests in most models, except the one between Model 2b-Rural and Model 1-Rural. The results suggest that social capital as a whole and CT, PP, and PT are important determinants of mental health for both rural and urban individuals; while CP is an important determinant for urban individuals.

On closer inspection of Model 2a–2e, we noticed that the individual-level aggregated SCI was positively associated with mental health (0.01, $p < 0.01$) in both subsamples. However, the community-level aggregated SCI was positively associated with mental health only in the rural subsample (0.02, $p < 0.05$). In urban areas, the impacts of CP on mental health were significantly positive at the individual level (0.01, $p < 0.01$) but negative at the communitive level (−0.03, $p < 0.01$). The impacts of individual-level CT (0.1, $p < 0.01$) was statistically significant in rural area. The community-level CT had statistically significant impacts in both rural and urban areas (0.06 and 0.07, $p < 0.01$, respectively), which justified the contextual effects in [39].

What is more, the individual-level PP was positively associated (0.03, $p < 0.01$) with mental health in both rural and urban areas. Finally, the individual level PTs (0.02, $p < 0.01$; 0.02, $p < 0.01$) were positively associated with mental health in both rural and urban areas.

Table 2. Results of regression models.

Model	0-Rural	0-Urban	1-Rural	1-Urban	2a-Rural	2a-Urban	2b-Rural	2b-Urban	2c-Rural	2c-Urban	2d-Rural	2d-Urban	2e-Rural	2e-Urban
Beta	−0.04(0.04)	0.10(0.03) **	−0.95(0.11) **	−1.13(0.09) **	−0.93(0.11) **	−1.07(0.09) **	−0.95(0.11) **	−1.10(0.09) **	−1.08(0.12) **	−1.23(0.09) **	−0.88(0.11) **	−1.06(0.09) **	−0.93(0.10) **	−1.13(0.09) **
Female			−0.05(0.03) *	0.08(0.02) **	−0.03(0.03)	0.08(0.02) **	−0.05(0.03) *	0.08(0.02) **	−0.04(0.03) *	0.08(0.02) **	−0.04(0.03)	0.08(0.02) **	−0.05(0.03) *	0.07(0.02) **
SocEco Statue.			0.15(0.01) **	0.16(0.01) **	0.15(0.01) **	0.15(0.01) **	0.15(0.01) **	0.16(0.01) **	0.15(0.01) **	0.16(0.01) **	0.15(0.01) **	0.16(0.01) **	0.15(0.01) **	0.16(0.01) **
Wealth			0.01(0.01) *	0.01(0.00) **	0.01(0.01) *	0.01(0.00) **	0.01(0.01) *	0.01(0.00) **	0.01(0.01) *	0.01(0.00) **	0.01(0.01) *	0.01(0.00) **	0.01(0.01) *	0.01(0.00) **
Relatively low			0.10(0.03) **	0.07(0.04) *	0.09(0.03) **	0.06(0.04) *	0.10(0.03) **	0.07(0.04) *	0.10(0.03) **	0.06(0.04) *	0.09(0.03) **	0.06(0.04)	0.09(0.03) **	0.07(0.04) *
Modest oncome			0.17(0.04) **	0.09(0.04) **	0.19(0.04) **	0.09(0.04) *	0.16(0.04) **	0.09(0.04) **	0.16(0.04) **	0.09(0.04) *	0.17(0.04) **	0.08(0.04) *	0.16(0.04) **	0.09(0.04) **
High oncome			0.17(0.07) **	0.09(0.04) *	0.19(0.07) **	0.09(0.04) *	0.16(0.07) **	0.09(0.04) *	0.17(0.07) **	0.09(0.04) *	0.16(0.07) **	0.08(0.04) *	0.17(0.07) **	0.10(0.04) *
Household size			−0.01(0.01) *	0.02(0.01) **	−0.01(0.01) *	0.02(0.01) **	−0.01(0.01) *	0.02(0.01) **	−0.01(0.01) *	0.02(0.01) **	−0.01(0.01) *	0.02(0.01) **	−0.01(0.01) *	0.02(0.01) **
Age 26–40			−0.29(0.06) **	−0.09(0.05) *	−0.30(0.06) **	−0.08(0.05) *	−0.29(0.06) **	−0.08(0.05)	−0.29(0.06) **	−0.09(0.05) *	−0.31(0.06) **	−0.10(0.05) *	−0.29(0.06) **	−0.09(0.05) *
Age 41–60			−0.38(0.07) **	−0.18(0.06) **	−0.40(0.07) **	−0.17(0.06) **	−0.38(0.07) **	−0.16(0.06) **	−0.39(0.07) **	−0.18(0.06) **	−0.41(0.07) **	−0.21(0.06) **	−0.39(0.07) **	−.18(.06) **
Age 61+			−0.28(0.08) **	−0.10(0.06) *	−0.30(0.07) **	−0.10(0.06) *	−0.28(0.08) **	−0.08(0.06)	−0.29(0.07) **	−0.10(0.06) *	−0.30(0.08) **	−0.14(0.06) *	−0.29(0.07) **	−0.10(0.06) *
Primary school			0.04(0.04)	0.00(0.06)	0.03(0.04)	−0.01(0.06)	0.04(0.04)	−0.00(0.06)	0.04(0.04)	0.01(0.06)	0.03(0.04)	−0.01(0.06)	0.04(0.04)	0.00(0.06)
Junior high			0.08(0.05) *	0.10(0.06) *	0.06(0.05)	0.08(0.06) *	0.07(0.05) *	0.09(0.06) *	0.08(0.05) *	0.11(0.06) *	0.06(0.05)	0.07(0.06)	0.08(0.05)	0.10(0.06) *
Senior high			0.05(0.06)	0.12(0.06) *	0.02(0.06)	0.10(0.06) *	0.03(0.06)	0.11(0.06) *	0.04(0.06)	0.13(0.06) *	0.02(0.06)	0.093(0.057) *	0.05(0.06)	0.12(0.06) *
College			−0.18(0.17)	0.12(0.06) *	−0.24(0.17)	0.08(0.06)	−0.21(0.18)	0.09(0.06) *	−0.20(0.17)	0.12(0.06) *	−0.20(0.17)	0.08(0.06)	−0.18(0.17)	0.12(0.06) *
Postgraduate			0.00(0.00)	0.01(0.20)	0.00(0.00)	−0.03(0.20)	0.00(0.00)	−0.03(0.20)	0.00(0.00)	0.01(0.20)	0.00(0.00)	−0.03(0.20)	0.00(0.00)	0.03(0.20)
Married			0.16(0.07) *	0.01(0.05)	0.16(0.07) *	0.00(0.05)	0.17(0.08) *	0.01(0.05)	0.15(0.07) *	0.00(0.05)	0.14(0.07) *	−0.01(0.05)	0.16(0.07) *	0.00(0.05)
Divorced/Widow			−0.06(0.10)	−0.14(0.07) *	−0.06(0.10)	−0.14(0.07) *	−0.05(0.10)	−0.14(0.07) *	−0.07(0.10)	−0.14(0.07) *	−0.07(0.10)	−0.15(0.07) *	−0.06(0.10)	−0.14(0.07) *
Unemployed			−0.15(0.14)	−0.10(0.03) **	−0.14(0.14)	−0.09(0.03) **	−0.15(0.14)	−0.09(0.03) **	−0.15(0.14)	−0.09(0.03) **	−0.14(0.14)	−0.09(0.03) **	−0.15(0.14)	−0.10(0.03) **
SCI-Indi.					0.01(0.00) **	0.01(0.00) **								
CP-Indi							0.01(0.01)	0.01(0.00) **						
CT-Indi									0.014(0.01) **	0.00(0.00)				
PP-Indi											0.03(0.01) **	0.03(0.01) **		
PT-Indi													0.02(0.01) **	0.01(0.00) **
SCI-Dist					0.02(0.01) *	−0.01(0.01)								
CP-Dist							0.01(0.02)	−0.03(0.01) **						
CT-Dist									0.06(0.03) **	0.07(0.02) **				
PP-Dist											0.00(0.03)	−0.01(0.02)		
PT-Dist													0.02(0.02)	−0.02(0.01)
σ_{μ}^2	0.15(0.03) **	0.12(0.02) **	0.10(0.02) **	0.08(0.01) **	0.09(0.02) **	0.08(0.01) **	0.10(0.02) **	0.08(0.01) **	0.09(0.02) **	0.07(0.01) **	0.10(0.02) **	0.08(0.01) **	0.09(0.02) **	0.08(0.01) **
σ_{ε}^2	0.88(0.02) **	0.88(0.02) **	0.74(0.02) **	0.74(0.01) **	0.74(0.02) **	0.74(0.01) **	0.74(0.02) **	0.74(0.02) **	0.74(0.02) **	0.74(0.01) **	0.74(0.02) **	0.74(0.01) **	0.74(0.02) **	0.74(0.01) **
−2ll	11,758.18	16,806.03	11,014.44	15,677.74	10,978.14	15,651.57	11,013.29	15,662.8	10,999.03	15,658.99	11,000.53	15,650.52	10,997.09	15,669.1
VPC	14.56%	12.00%	11.90%	9.76%	10.84%	9.76%	−11015.59	9.76%	10.84%	8.64%	11.90%	9.76%	10.84%	9.76%

** p <0.01; * p <0.05. Standard errors in parentheses. σ_{ε}^2: individual-level variances; σ_{μ}^2: community-level variance. SCI: social capital index; CP: civic participation; PP: political participation; PT: political trust; VPC: variance partition coefficient. Indi: Individual-level. Dist: District-level.

Table 3. The likelihood chi-squared tests.

Model Comparison	Δ-2ll	df	*p*-Value
1-Rural vs. 0-Rural	−743.74	18	<0.01
1-Urban vs. 0-Urban	−1128.29	18	<0.01
2a-Rural vs. 1-Rural	−36.3	2	<0.01
2a-Urban vs. 1-Urban	−26.17	2	<0.01
2b-Rural vs. 1-Rural	−1.15	2	>0.10
2b-Urban vs. 1-Urban	−14.94	2	<0.01
2c-Rural vs. 1-Rural	−15.41	2	<0.01
2c-Urban vs. 1-Urban	−18.75	2	<0.01
2d-Rural vs. 1-Rural	−13.91	2	<0.01
2d-Urban vs. 1-Urban	−27.22	2	<0.01
2e-Rural vs. 1-Rural	−17.35	2	<0.01
2e-Urban vs. 1-Urban	−8.64	2	<0.05

Δ-2ll = The –2ll value of the first model minus the –2ll value of the second model.

4. Discussion

This study follows the composite hypothesis to simultaneously test the effects of human and social capital on an individual's self-rated mental health at both individual and community levels. The results of our multilevel linear regressions seem to support the composite hypothesis. We will discuss the impacts of social capital and human capital covariates on the outcome variable separately.

4.1. Effect of Social Capital

The study revealed that higher civic trust helped to improve the mental health at the individual (rural areas) and community level (both rural and urban areas). This finding further corroborated prior findings between positive health outcomes and trust [15,63–68], which suggested that higher civic trust may enhance social networking among individuals, thereby reducing the mental stress and then lead to better mental health status.

Our study revealed that higher political trust helped to improve mental health at the individual and community level in rural areas. This finding was consistent with a prior study [69], which indicated that higher political trust could, one the one hand, led to an enhanced feeling of security [69] and created stronger affiliations with the legal system as identified through questions on government functioning; and on the other hand, higher political trust also promoted a better sense of belongingness and sense of responsibility [69].

We found that higher political participation also tended to improve mental health at the individual level in both rural and urban areas. That is probably that political participation may promote a sense of gratification from the social welfare activities and aroused deeper feelings of social and moral responsibility among individuals. It may be also instrumental in warding off socially deviant behavior among individuals, inculcating socially beneficial norms and promoting mental health.

Finally, we revealed that higher civic participation helped to improve the mental health at the individual level in urban areas. That is probably that higher civic participation could be manifested through progressive coaching mechanism or knowledge dissemination process to fellow individuals, thereby leading to enhanced self-esteem among individuals. This could help in promoting social interactions, reducing stress, improving security networks, and thereby improving mental health. However, at the community level our study suggested that enhanced civic participation gave rise to reduced mental health in urban areas, ceteris paribus. An explanation to this could be drawn from the "network resources" approach [70]. As predicted in [71], subjective social status could be considered as a partly mediating factor in explaining the negative association between network resources and psychological stress. We found a positive correlation (0.36, $p < 0.01$, in Table 4) between subjective social-economic status and civic participation in our urban sample, suggesting that a higher level of community civic participation could create extra responsibility or overwhelming burden on an individual with relatively high social-economic status. This is especially true in China in where loss of "face" has been considered as a key deterrent in an individual's access to social capital [20,72]. Individuals will stretch beyond their economic horizon to meet moral obligations and reciprocity norms in order to save "face" in their communities. Non-adherence to these norms would result in a disastrous solitary state leading to a lower mental health. However, as found in a prior study [20], over participation in social events could involve detrimental social norms like alcohol consumption, physical stress and inevitably reduced social family interactions, leading to increased mental stress and worsening mental health.

Table 4. Descriptive statistics and correlation matrix.

Rural\Urban	Rural (mean, SD)	Urban (mean, SD)	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Female	0.51 (0.5)	0.53 (0.5)		−0.01	−0.01	−0.03 **	−0.02	−0.02	−0.15 **	0.02	0.07 **	−0.07 **	−0.08 **	−0.02	−0.08 **	0.00
2. Soc-Economic Status	7.18 (2.15)	6.64 (2.12)	0.01		0.41 **	0.25 **	0.02	−0.13 **	0.26 **	−0.00	−0.17 **	0.36 **	0.28 **	0.11 **	0.09 **	0.08 **
3. Wealth Index	−2.75 (3.21)	1.93 (4.87)	−0.02	0.41 **		0.35 **	0.02	−0.11 **	0.37 **	0.03 *	−0.10 **	0.24 **	0.36 **	−0.01	0.10 **	0.01
4. HouseholdIncome	9211.39 (10,639.28)	25,842.81 (44,648.74)	−0.05 **	0.35 **	0.51 **		0.01	−0.08 **	0.18 **	−0.02	−0.08 **	0.11 **	0.21 **	−0.01	0.05 **	0.01
5. HouseholdSize	4.47 (1.97)	3.47 (1.65)	−0.01	0.08 **	0.13 **	0.10 **		0.00	−0.10 **	0.08 **	0.03 *	0.02	−0.04 **	−0.02	−0.02	0.02
6. Age	44.73 (13.81)	44.68 (15.45)	−0.11 **	−0.07 **	−0.12 **	−0.10 **	0.03		−0.41 **	0.24 **	−0.05 **	−0.34 **	−0.28 **	0.02	0.18 **	−0.00
7. Education	1.33 (0.91)	2.4 (1.11)	−0.22 **	0.22 **	0.30 **	0.23 **	−0.04 **	−0.31 **		−0.07 **	−0.10 **	0.27 **	0.42 **	0.04 **	0.08 **	−0.04 **
8. Married	0.9 (0.31)	0.8 (0.4)	0.10 **	0.10 **	0.06 **	0.04 *	0.07 **	0.05 **	−0.00		0.07 **	−0.02	−0.13 **	0.03 *	0.12 **	−0.02
9. Unemployed	0.01 (0.1)	0.14 (0.34)	−0.02	−0.00	0.04 *	0.03	0.00	−0.03 *	0.04 *	−0.04 *		−0.10 **	−0.14 **	−0.06 **	−0.08 **	−0.03 *
10. MentalHealth	−0.07 (1.04)	0.05 (0.97)	−0.06 **	0.34 **	0.25 **	0.20 **	−0.00	−0.28 **	0.25 **	0.05 **	0.01		0.18 **	0.11 **	0.07 **	0.06 **
11. CP	−0.53 (0.5)	0.37 (1.09)	−0.10 **	0.14 **	0.23 **	0.22 **	0.01	−0.17 **	0.32 **	−0.10 **	0.04 **	0.11 **		0.07 **	0.15 **	0.03
12. CT	0.25 (0.98)	−0.18 (0.97)	−0.06 **	0.05 **	0.04 *	0.04 *	−0.02	0.05 **	0.05 **	0.03	−0.01	0.08 **	0.04 **		0.07 **	0.07 **
13. PP	0.3 (1)	−0.21 (0.95)	−0.16 **	0.13 **	0.07 **	0.08 **	−0.01	0.07 **	0.13 **	0.08 **	−0.02	0.11 **	0.12 **	0.10 **		0.08 **
14. PT	0.1 (0.99)	−0.07 (1)	−0.03	0.05 **	−0.06 **	−0.01	0.01	0.03 *	−0.02	−0.02	−0.01	0.04 *	0.05 **	0.06 **	0.12 **	

** $p < 0.01$; * $p < 0.05$. The upper triangular part of the matrix contains the correlation coefficients of the urban subsample. The lower triangular part of the matrix contains the correlation coefficients of the rural subsample.

4.2. Effect of Human Capital

Overall our findings were in line with those previously reported in developing countries that human capital factors were important predictors of self-rated mental health. The compositional effects remained statistically significant even after taking social capital variables into account.

In particular, we found that an increase in age led to a lower self-reported mental health, which is of paramount importance when people reach their middle-age. That is probably because that responsibilities and senses of isolation or worthlessness may be bound to increase with age. We believe that this is especially true for the elderly in rural areas, when their adult children are away as migrant workers.

Female gender was also found to be related to poorer self-rated mental health in rural areas than in urban areas. That is probably because, on one hand, rural women may suffer from lower social status in their villages and on the other hand, it may be accepted that when husbands are away as migrant workers, the responsibility to take care of younger family members and elderly parents always falls on the women without anyone to share the load.

Secondary education was also found to improve self-rated mental health in both rural and urban areas. This is probably because education may enable individuals to better understand the benefits of a healthy lifestyle and healthcare utilization [56]. We noticed that when both individual and community-level social capital variables were present, college education exhibited a marginally significant, negative effect on self-rated mental health in rural areas. We believed that when compared to their urban counterparts, the college students from rural areas may have lesser chances to find decent jobs in China because their families possessed fewer social networks. This probably led to low economic returns to their college tuition fees and made them anxious and frustrated.

Our study also found that wealth mattered for mental health. Being rich or having relatively high social-economic status was strongly associated with superior self-rated mental health. We argued that this was probably because richer households or individuals may be able to access more medical care resources and experience less stress when compared to the poor, who had to work harder to earn their living.

An increase of household size in rural area was associated with a decrease in self-rated mental health status. Given family size in rural China is usually larger (mean = 4.47, in Table 4), this association may have had its origins in increased competition between siblings for parental attention, resources and extra burden to take care of elderly parents or even grandparents. However, an opposite effect was found in the urban sample. A possible interpretation of the positive family size–mental health relationship in urban China may be due to the strict one-child policy. With the family size being small in urban areas (mean = 3.47, in Table 4), children may enjoy greater parental attention in their early life and as a result, developed greater resilience to maladaptive responses and stressful events in adulthood. Likewise, a larger family size may increase the social contact of urban elderly, reducing their sense of loneliness, a general cause of many mental illnesses.

In addition, our study indicated that, the married individuals portrayed better mental health when compared to unmarried individuals in rural areas. This finding was not supported in urban areas. We found that the mental health of the urban individual was found to deteriorate following a divorce or loss of spouse.

We believe that there is a stark distinction in the levels of aspirations between the individuals hailing from rural and urban areas. In rural areas, the concept of close families, strong bonding, and well-knit relations may exist that foster easy solving of mutual discord and problems. However, in urban areas, individuals live with more autonomy in smaller nuclear families, which could be beneficial to the extent of lesser burdening an individual with monetary and social responsibilities. At the same time, smaller nuclear families could be detrimental since the individuals may face increased mental stress when handling problems since they lack the emotional support of spouse, leading to poor mental health and estranged family ties. Therefore, with the resulting non-successful matrimony or death of beloved partner, an individual in urban area may develop a sense of personal loneliness and dissatisfaction, lack of emotional stability, and a dearth of social interaction develops, resulting in reduced mental health.

Finally, the employment status had no significant impact in rural areas. However, the effect of unemployment tended to be significantly negative in urban areas. We believe that the societal set up in rural and urban areas is starkly different. In rural areas, there are more closely-knit families characterized by large joint families. These families live together and most probably the family earnings are shared for the entire family's sustenance. This involves collective earning and collective spending to supplement the entire family's needs. This could be in contrast to the urban areas, where the concept of nuclear families with fewer members predominates. Financially independent individuals most probably prefer to be socially independent as well. Hence for individuals in urban areas, professional upheavals coupled with lack of emotional support may lead to mental dissatisfaction and reduced mental health.

5. Limitations and Conclusions

5.1. Limitations

Like most cross-sectional analyses, this study may be subject to common method bias as information regarding both social capital and mental-health is self-reported. Since both social capital and mental health can be expressed as an individual's general well-being, there may be a recall bias [16,39]. In addition, the cross-sectional design usually leads to reversed causality problem—poor mental health may reduce one's social capital. For example, participation in political activities requires inherent enthusiasm and good mental situation as a pre-requisite. A longitudinal study is desired to check the robustness of our findings. However, the aggregated measures of social capital and the use of mean-centered individual social capital can mitigate this problem as individual co-variance can be diluted [39]. On the other hand, the problem of reverse causality at the community level may not be an issue. Since the Chinese government restricts people's freedom to choose which district to live in under the Household Registration System, the community-level measures of social capital may be exogenous and random.

5.2. Conclusions

Our results suggest that social capital is reflective of mental health both at the individual and community level. We have employed four cognitive-structural and horizontal-vertical constructs to comprehensively measure social capital in Chinese contexts and empirically test for observations across samples from rural and urban Chinese population segments. The results support the argument that in general, individual social capital holds a positive association with self-rated mental health in China. This study also finds that a significant proportion of the total variation in individual's self-rated mental health status can be explained by differences at the community level. Those living in areas with low civic trust social capital have high excess risk of poor self-rated mental health. On the contrary, urban areas offer disagreement on the structural dimension measured through community-civic participation index, offering a negative association. An explanation for this has been based on the network resources approach. The likelihood ratio tests further support our conclusions that the impact of both individual and contextual social capital is not an artifact. These results offer further dimensions for future research. A more detailed and vigorous investigation may be called upon to study the impact of social capital variation, involving development of more suitable instruments to unearth the pattern of variations more systematically.

Finally, we suggest that policy mechanisms should be set in place to strengthen social networks and create economic upliftment of individuals. The programs should be aimed at inculcating social parity among the community members and at reducing the disparity in the access of medical benefits. Community health and social wellbeing for individuals should to be given attention with special emphasis to women, middle-aged, and elderly citizens. In rural areas, schemes should be set in place to encourage employment and business opportunities for youth through promotion of cottage industries. These schemes will be able to create employment avenues for the large number of college-educated students hailing from rural areas and hence avoid their rapid transition to larger cities. It will also be helpful in curtailing dissatisfaction among their elderly family members who are plagued with issues of staying alone as their children and even grand-children migrate to urban areas for better job prospects.

Author Contributions: Conceptualization, L.G.; methodology, X.L.; software, X.L.; validation, X.L.; formal analysis, X.L. and R.L.; investigation, B.L.; resources, B.L.; data curation, B.L.; writing—original draft preparation, L.G.; writing—review and editing, L.G. and R.L.; supervision, B.L.; project administration, B.L.; funding acquisition, L.G. and B.L.; X.L and L.G. equally contributed to this article.

Funding: This project was funded by the Qilu Early Career Scholarship Project of Shandong University, China.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Sartorius, N. *Fighting for Mental Health*; Cambridge University Press: Cambridge, UK, 2002.
2. Sharan, P.; Levav, I.; Olifson, S.; de Francisco, A.; Saxena, S. *Research Capacity for Mental Health in Low- and Middle-Income Countries: Results of a Mapping Project*; World Health Organization & Global Forum for Health Research: Geneva, Switzerland, 2007.
3. Jiang, Q.; Li, J.; Wang, G. The Relationship between Constitution of Traditional Chinese Medicine in the First Trimester and Pregnancy Symptoms: A Longitudinal Observational Study. *Evidence-Based Complement. Altern. Med.* **2016**, *2016*, 1–8. [CrossRef] [PubMed]
4. *China's Mental Health Action Program 2015–2020 (in Chinese)*; National Health and Family Planning Commission: Beijing, China, 2015.
5. World Health Organization. *Global Health Estimates 2016: Disease Burden by Cause, Age, Sex, by Country and by Region, 2000–2016*; World Health Organization: Geneva, Switzerland, 2018.
6. Lee, S. Mental illness in China. *The Lancet* **2009**, *374*, 1063–1064. [CrossRef]
7. Henderson, S.; Whiteford, H. Social capital and mental health. *Lancet* **2003**, *362*, 505–506. [CrossRef]
8. Almedom, A.M. Social capital and mental health: An interdisciplinary review of primary evidence. *Soc. Sci. Med.* **2005**, *61*, 943–964. [CrossRef] [PubMed]
9. Browning, C.R.; Cagney, K.A. Neighborhood Structural Disadvantage, Collective Efficacy, and Self-Rated Physical Health in an Urban Setting. *J. Health Soc. Behav.* **2002**, *43*, 383. [CrossRef] [PubMed]
10. Cook, K.S. *Trust in Society*; RSF Series on Trust; Russell Sage Foundation: New York, NY, USA, 2001.
11. Franzini, L.; Caughy, M.; Spears, W.; Esquer, M.E.F. Neighborhood economic conditions, social processes, and self-rated health in low-income neighborhoods in Texas: A multilevel latent variables model. *Soc. Sci. Med.* **2005**, *61*, 1135–1150. [CrossRef] [PubMed]
12. Fukuyama, F. *Trust: The Social Virtues and the Creation of Prosperity*; Free Press: New York, NY, USA.
13. Kavanagh, A.M. Does gender modify associations between self rated health and the social and economic characteristics of local environments? *J. Epidemiol. Community Health* **2006**, *60*, 490–495. [CrossRef] [PubMed]
14. Kawachi, I.; Kennedy, B.P.; Glass, R. Social capital and self-rated health: a contextual analysis. *Am. J. Public Health* **1999**, *89*, 1187–1193. [CrossRef] [PubMed]
15. Poortinga, W. Social capital: An individual or collective resource for health? *Soc. Sci. Med.* **2006**, *62*, 292–302. [CrossRef] [PubMed]
16. Poortinga, W. Social relations or social capital? Individual and community health effects of bonding social capital. *Soc. Sci. Med.* **2006**, *63*, 255–270. [CrossRef] [PubMed]
17. Schwarzer, R.; Leppin, A. Social support and health: A meta-analysis. *Psychol. Health* **1989**, *3*, 1–15.
18. Wang, H.-H.; Wu, S.-Z.; Liu, Y.-Y. Association Between Social Support and Health Outcomes: A Meta-analysis. *Kaohsiung J. Med. Sci.* **2003**, *19*, 345–350. [CrossRef]
19. Wen, M.; Browning, C.R.; Cagney, K.A. Poverty, affluence and income inequality: Neighbourhood economic structure and its implications for health. *Soc. Sci. Med.* **2003**, *57*, 843–860. [CrossRef]

20. Nie, X.; Zhu, Y.; Fu, H.; Dai, J.; Gao, J. The "Dark Side" Effects of Social Capital on Harmful Drinking among Chinese Community Residents: A Multilevel Study. *IJERPH* **2018**, *15*, 2249. [CrossRef] [PubMed]
21. Yip, W.; Subramanian, S.; Mitchell, A.D.; Lee, D.T.; Wang, J.; Kawachi, I. Does social capital enhance health and well-being? Evidence from rural China. *Soc. Sci. Med.* **2007**, *64*, 35–49. [CrossRef] [PubMed]
22. Skrabski, A.; Kopp, M.; Kawachi, I. Social capital and collective efficacy in Hungary: Cross sectional associations with middle aged female and male mortality rates. *J. Epidemiol. Community Health* **2004**, *58*, 340–345. [CrossRef] [PubMed]
23. Ichida, Y.; Kondo, K.; Hirai, H.; Hanibuchi, T.; Yoshikawa, G.; Murata, C. Social capital, income inequality and self-rated health in Chita peninsula, Japan: A multilevel analysis of older people in 25 communities. *Soc. Sci. Med.* **2009**, *69*, 489–499. [CrossRef] [PubMed]
24. Fujisawa, Y.; Hamano, T.; Takegawa, S. Social capital and perceived health in Japan: An ecological and multilevel analysis. *Soc. Sci. Med.* **2009**, *69*, 500–505. [CrossRef] [PubMed]
25. Suzuki, E.; Takao, S.; Subramanian, S.; Komatsu, H.; Doi, H.; Kawachi, I. Does low workplace social capital have detrimental effect on workers' health? *Soc. Sci. Med.* **2010**, *70*, 1367–1372. [CrossRef] [PubMed]
26. Miller, D.L.; Scheffler, R.; Lam, S.; Rosenberg, R.; Rupp, A. Social capital and health in Indonesia. *World Dev.* **2006**, *34*, 1084–1098. [CrossRef]
27. Rehnberg, C.; Sun, X.; Meng, Q. How are individual-level social capital and poverty associated with health equity? A study from two Chinese cities. *Int. J. Equity Health* **2009**, *8*, 2.
28. Yamaoka, K. Social capital and health and well-being in East Asia: A population-based study. *Soc. Sci. Med.* **2008**, *66*, 885–899. [CrossRef] [PubMed]
29. Zhang, S.; Dang, R.; Yang, N.; Bai, Y.; Wang, L.; Abbey, C.; Rozelle, S. Effect of Caregiver's Mental Health on Early Childhood Development across Different Rural Communities in China. *IJERPH* **2018**, *15*, 2341. [CrossRef] [PubMed]
30. Liu, H.; Fang, H.; Zhao, Z. Urban–rural disparities of child health and nutritional status in China from 1989 to 2006. *Econ. Hum. Biol.* **2013**, *11*, 294–309. [CrossRef] [PubMed]
31. Wu, Q.; Lu, D.; Kang, M. Social capital and the mental health of children in rural China with different experiences of parental migration. *Soc. Sci. Med.* **2015**, *132*, 270–277. [CrossRef] [PubMed]
32. Cao, W.; Li, L.; Zhou, X.; Zhou, C. Social capital and depression: Evidence from urban elderly in China. *Aging Ment. Health* **2014**, *19*, 418–429. [CrossRef] [PubMed]
33. Muntaner, C.; Lynch, J.; Smith, G.D. Social Capital, Disorganized Communities, and the Third Way: Understanding the Retreat from Structural Inequalities in Epidemiology and Public Health. *Int. J. Health Serv.* **2001**, *31*, 213–237. [CrossRef] [PubMed]
34. Rojas, Y.; Carlson, P. The stratification of social capital and its consequences for self-rated health in Taganrog, Russia. *Soc. Sci. Med.* **2006**, *62*, 2732–2741. [CrossRef] [PubMed]
35. Rose, R. How much does social capital add to individual health? A survey study of Russians. *Soc. Sci. Med.* **2000**, *51*, 1421–1435. [CrossRef]
36. Kawachi, I.; Kim, D.; Coutts, A.; Subramanian, S. Commentary: Reconciling the three accounts of social capital. *Int. J. Epidemiol.* **2004**, *33*, 682–690. [CrossRef] [PubMed]
37. Bowling alone: The collapse and revival of American community. *Choice Rev. Online* **2000**, *38*, 38.
38. Baum, F.; Ziersch, A. Social capital. *J. Epidemiol. Community Health* **2003**, *57*, 320–323. [CrossRef] [PubMed]
39. Engström, K.; Mattsson, F.; Järleborg, A.; Hallqvist, J. Contextual social capital as a risk factor for poor self-rated health: A multilevel analysis. *Soc. Sci. Med.* **2008**, *66*, 2268–2280. [CrossRef] [PubMed]
40. Harpham, T.; Grant, E.; Rodriguez, C. Mental health and social capital in Cali, Colombia. *Soc. Sci. Med.* **2004**, *58*, 2267–2277. [CrossRef] [PubMed]
41. Krishna, A.; Shrader, E. The social capital assessment tool. In *Understanding and Measuring Social Capital: A Multidisciplinary Tool for Practitioners*; Grootaert, C., Van Bastelaer, T., Eds.; The World Bank: Washington, DC, USA, 2002.
42. Subramanian, S.V.; Kim, D.J.; Kawachi, I. Social Trust and Self-Rated Health in US Communities: A Multilevel Analysis. *J. Heredity* **2002**, *79*, S21–S34. [CrossRef]
43. Kaplan, G.A.; Lynch, J.W. Whither studies on the socioeconomic foundations of population health. *Am. J. Public Health* **1997**, *87*, 1409–1411. [CrossRef]
44. Szreter, S. A new political economy for New Labour: The importance of social capital. *Renewal* **1999**, *7*, 30–44.
45. Woolcock, M. Social capital and economic development: Toward a theoretical synthesis and policy framework. *Theor. Soc.* **1998**, *27*, 151–208. [CrossRef]

46. Bourdieu, P. The forms of capital. In *Handbook of Theory and Research for the Sociology of Education*; Richardson, J., Ed.; Greenwood Press: New York, NY, USA, 1986; pp. 241–258.
47. Portes, A. Social Capital: Its Origins and Applications in Modern Sociology. *Annu. Rev. Sociol.* **1998**, *24*, 1–24. [CrossRef]
48. Portes, A. The two meanings of social capital. *Sociol. Forum* **2000**, *15*, 1–12. [CrossRef]
49. Lochner, K. Social capital: a guide to its measurement. *Health Place* **1999**, *5*, 259–270.
50. Coleman, J.S. Social Capital in the Creation of Human Capital. *Am. J. Sociology* **1988**, *94*, S95–S120. [CrossRef]
51. Coleman, J. *Foundations of Social Theory*; Harvard University Press: Cambridge, MA, USA, 1990.
52. LaPalombara, J.; Putnam, R.D. Making Democracy Work: Civic Traditions in Modern Italy. *Political Sci. Q.* **1993**, *108*, 549. [CrossRef]
53. Putnam, R.D. Bowling Alone: America's Declining Social Capital. *Cult. Politics* **2000**, *6*, 223–234.
54. Oakes, J. The (mis)estimation of neighborhood effects: Causal inference for a practicable social epidemiology. *Soc. Sci. Med.* **2004**, *58*, 1929–1952. [CrossRef] [PubMed]
55. Filmer, D.; Pritchett, L. Estimating Wealth Effects Without Expenditure Data–Or Tears: An Application to Educational Enrollments in States of India. *Demography* **2001**, *38*, 115–132. [PubMed]
56. Habibov, N.N.; Afandi, E.N. Self-rated health and social capital in transitional countries: Multilevel analysis of comparative surveys in Armenia, Azerbaijan, and Georgia. *Soc. Sci. Med.* **2011**, *72*, 1193–1204. [CrossRef] [PubMed]
57. Dahlin, J.; Härkönen, J. Cross-national differences in the gender gap in subjective health in Europe: Does country-level gender equality matter? *Soc. Sci. Med.* **2013**, *98*, 24–28. [CrossRef] [PubMed]
58. Blakely, T.; Subramanian, S.V. Multilevel studies. In *Methods for Social Epidemiology*; Oakes, M., Kaufman, J., Eds.; Jossey Bass: San Francisco, CA, USA, 2006; pp. 316–340.
59. Goldstein, H. *Multilevel Statistical Models London*; Oxford University Press: New York, NY, USA, 2003.
60. Horton, N.J. Multilevel and Longitudinal Modeling Using Stata. *Am. Statistician* **2006**, *60*, 293–294. [CrossRef]
61. Kemp, F. Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. *Statistician* **2003**, *52*, 691. [CrossRef]
62. Longford, N.T.; Bryk, A.S.; Raudenbush, S.W. Hierarchical Linear Models: Applications and Data Analysis Methods. *Contemp. Sociol.* **1993**, *22*, 293. [CrossRef]
63. Snelgrove, J.W.; Pikhart, H.; Stafford, M. A multilevel analysis of social capital and self-rated health: Evidence from the British Household Panel Survey. *Soc. Sci. Med.* **2009**, *68*, 1993–2001. [CrossRef] [PubMed]
64. Fujiwara, T.; Kawachi, I. A prospective study of individual-level social capital and major depression in the United States. *J. Epidemiol. Community Health* **2008**, *62*, 627–633. [CrossRef] [PubMed]
65. Giordano, G.N.; Lindstrom, M. The impact of changes in different aspects of social capital and material conditions on self-rated health over time: A longitudinal cohort study. *Soc. Sci. Med.* **2010**, *70*, 700–710. [CrossRef] [PubMed]
66. Hyyppä, M.T.; Mäki, J.; Impivaara, O.; Aromaa, A. Individual-level measures of social capital as predictors of all-cause and cardiovascularmortality. *Eur. J. Epidemiol.* **2007**, *22*, 589–597. [CrossRef] [PubMed]
67. Nieminen, T.; Martelin, T.; Koskinen, S.; Aro, H.; Alanen, E.; Hyyppä, M.T. Social capital as a determinant of self-rated health and psychological well-being. *Int. J. Public Health* **2010**, *55*, 531–542. [CrossRef] [PubMed]
68. Tampubolon, G.; Subramanian, S.V.; Kawachi, I. NEIGHBOURHOOD SOCIAL CAPITAL AND INDIVIDUAL SELF-RATED HEALTH IN WALES. *Health Econ.* **2011**, *22*, 14–21. [CrossRef] [PubMed]
69. De Silva, M.J.; Huttly, S.R.; Harpham, T.; Kenward, M.G. Social capital and mental health: A comparative analysis of four low income countries. *Soc. Sci. Med.* **2007**, *64*, 5–20. [CrossRef] [PubMed]
70. Ferlander, S. The Importance of Different Forms of Social Capital for Health. *Acta Sociol.* **2007**, *50*, 115–128. [CrossRef]
71. Song, L. Social Capital and Psychological Distress. *J. Health Soc. Behav.* **2011**, *52*, 478–492. [CrossRef] [PubMed]
72. Yang, L.H.; Kleinman, A. 'Face' and the embodiment of stigma in China: The cases of schizophrenia and AIDS. *Soc. Sci. Med.* **2008**, *67*, 398–408. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

Start-Up of a Biofilter in a Full-Scale Groundwater Treatment Plant for Iron and Manganese Removal

Huiping Zeng [1,*], Can Yin [1], Jie Zhang [1,2] and Dong Li [1,*]

[1] Key Laboratory of Water Science and Water Environment Recovery Engineering, Beijing University of Technology, Beijing 100124, China; 18811715867@163.com (C.Y.); 6282031@163.com (J.Z.)
[2] State Key Laboratory of Urban Water Resource and Environment, Harbin Institute of Technology, Harbin 150090, China
* Correspondence: zenghuiping@bjut.edu.cn (H.Z.); lidong2006@bjut.edu.cn (D.L.); Tel.: +86-152-0122-7561 (H.Z.); +86-131-2121-8006 (D.L.); Fax: +86-106-739-2099 (H.Z.); +86-106-739-2579 (D.L.)

Received: 28 December 2018; Accepted: 23 February 2019; Published: 27 February 2019

Abstract: In recent years, biological purification technology has been widely developed in the process of iron and manganese removal from groundwater. The cultivation and maturation of the biological filter layer are key for biological iron and manganese removal processes. The time needed for maturation varies significantly with the water quality, filter and filter media conditions and operation parameters; sometimes it takes only one or two months, sometime more than half a year. In this paper, the feasibility of adopting an intermittent operation for the cultivation of biofilter was investigated with productive filters in a groundwater treatment plant, and the comparative test of the filter column was conducted. The results showed that the intermittent operation had little effect on the cultivation of the biofilter because dissolved oxygen would be gradually exhausted during the filter-suspension process, making the filter layer anaerobic, thus possibly inhibiting the growth and reproduction of IMOB (Iron and Manganese Oxidizing Bacteria). At the same time, the test shows that when the mature biological filter needs the suspension operation, the emptying method should be considered to avoid the destruction of the biological layer.

Keywords: groundwater; iron and manganese removal; biofilter

1. Introduction

Groundwater usually contains iron and manganese in Northeast China, above the allowed maximum concentration levels 0.3 mg/L and 0.1 mg/L for iron and manganese, respectively, according to the Chinese Standards for Drinking Water Quality [1]. Excessive iron and manganese can affect human health and industry production [2,3]. Together with dissolved iron, manganese could form sediments in drinking water distribution lines and incidents of "black" or "brown" water have occurred with the fluctuation of the water supply.

Many methods have been developed since the end of 19th century, mainly divided into chemical oxidation and biological oxidation. Chemical oxidation is usually adopted by traditional water treatment plants [4–7], including oxidation with atmospheric oxygen assisted by aeration and oxidation by chemical agents such as potassium permanganate, chlorine, chlorine dioxide and ozone. Fe (II) can be oxidized by oxygen at natural pH easily, while the condition of manganese oxidation is stringent. Excessive aeration is often needed to raise the pH to 9–10 to achieve the abiotic oxidation of soluble Mn (II) to insoluble Mn (III/IV) (oxyhydr)oxides. At natural pH, manganese can only be oxidized by the stronger oxidants: ozone, chlorine dioxide, chlorine and potassium permanganate. However, adverse effects could be caused with excessive oxidants. Excessive potassium permanganate will make the color of the filtrate pink. Excessive chlorine dioxide can generate chloric organics such as chlorobenzene, and chlorophenol which are thought to be carcinogens [8–10].

Biological iron and manganese removal has shown much promise as an effective low-cost way to treat iron- and manganese-contaminated groundwater, becoming more and more common [11–14] in Europe, Asia, and North America. It was noted that biotic methods could substantially increase treatment capacity and reduce the cost of operation compared with abiotic methods [15]. However, the long time start-up period of the biofilters is a main drawback for their popularization and application. Chemical oxidation of iron and manganese occurs in several minutes or even several seconds by strong oxidants. Therefore, almost no time is needed for the start-up of a water treatment plant for iron and manganese removal adopting the abiotic oxidation process. For the plants adopting the biotic oxidation process, 1–2 months or even more time is needed for the maturation of the biofilter [16,17] because time is needed for the attachment of inoculated bacteria to the filter media, and time is needed for the acclimation and enrichment of functional bacteria.

Recently, several biological groundwater treatment plants have been built in Northeast China, which have been operated successfully and the filtrate quality is better than the Chinese Standard, but because of the difference in water quality, filter material and operation parameters, the start-up time varies from one month to eight months [18–22].

The establishment of a biological layer is the first step for the start-up of the biological iron and manganese removal process. The efficient and stable ability of iron and manganese removal depends on the establishment of a stable and complex ecosystem with iron- and manganese-oxidizing bacteria as dominant bacteria [23], but these bacteria are microorganisms with a low metabolic rate and growth rate. Therefore, the biofilter for iron and manganese removal is very different from the biofilter for carbon and nitrogen removal.

So far, there are not many references on the start-up of the biological iron and manganese removal filter except that when ammonia nitrogen is contained in raw water, the removal of manganese requires complete nitrification of ammonia nitrogen, so the maturation of the filter takes 3–4 months [24,25].

This paper provides a specific example of the start-up of biological groundwater treatment removal for iron and manganese removal. Because of the external pipe network problems, the water plant can only be operated intermittently. Therefore, the effect of the cultivation of an intermittent water supply biological filter layer was analyzed with the aid of a filter column test, in order to provide more references and accumulate more experience for the start-up of the biological iron and manganese removal filter layer.

2. Materials and Methods

2.1. Water Treatment Plant and Groundwater

Shenyang is the capital of Liaoning Province in Northeast China. Groundwater is an important water supply source in Shenyang. The groundwater temperature is 10–12 °C all year round due to the latitude. In this paper, the groundwater came from 16 deep wells (depth 60–100 m). The water outputs were between 100 to 150 m^3/h; the groundwater characteristics are shown in Table 1. The total treatment capacity of the WTP (Water Treatment Plant) is 3×10^4 m^3/d; the flow chart is shown in Figure 1, with the aim that the simultaneous removal of iron and manganese occurs in one filter. After the raw water was extracted from the ground, it first entered the cascade aeration tank for oxygen filling, which provided sufficient oxygen for the subsequent bio-oxidation of iron and manganese in the biochemical filter, where iron and manganese ions were oxidized into solid iron and manganese oxides, which were intercepted and removed by the filter layer. Finally, the purified water entered into the clear water reservoir and then the distribution pipe network. With the extension of filtration time, the interception impurities of the filter layer increased, and the filtration resistance increased sharply, and the filter backwashing process was needed. The filter had a depth of 1000 mm, using sand as the filter media, with a diameter of 0.5–1.2 mm.

Table 1. Characteristics of the groundwater.

Parameter	Value	Parameter	Value
Temperature (°C)	12–14	Total dissolved solids (mg/L)	240–260
pH	6.7–6.9	DO (mg/L)	0
Turbidity (NTU)	32.3–39.0	Conductivity (Us/cm)	268–280
Mn^{2+} (mg/L)	0.8–2.0	Fe^{2+} (mg/L)	0.15–0.20

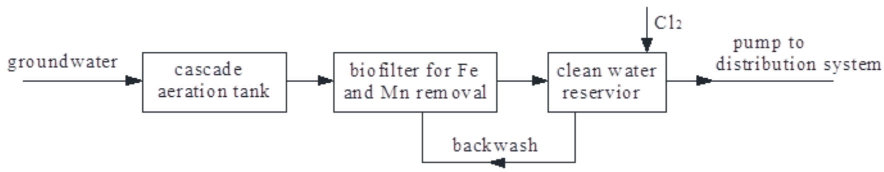


Figure 1. Flow chart of Water Treatment Plants.

Raw water from waterworks was basically used as the influent for the filter column test, and chemical agents can be added to change the content of iron, manganese and other substances according to the need.

2.2. Start-Up Process

Due to the construction of the external pipeline network, water plants could only be operated intermittently at the initial stage of biofilter cultivation, and then gradually transit to continuous operation. The whole operation process is shown in Table 2.

Table 2. Operation process during the biofilter cultivation.

Stage	Frequency of Operation	Velocity	Time
Intermittent Operation	Once three days, 6 h at a time	1.5 m/h	Two months
	Once a day, 7 h at a time	1.5 m/h	Two months
Continuous operation	Continuous running	1.5 m/h	Two months
speed-up	Continuous running	1.5 m/h to 5 m/h step by step	One month

2.3. Filter Column Test

The test device is shown in Figure 2. The filter column was made of Plexiglas with a height of 3 m and an inner diameter of 250 mm. A sampling port was set every 10 cm along the direction of the filter layer. In order to simulate the actual productive filter, the filling of filter media and the inoculation amount were consistent with those in filter 1#. The filter column adopted a downward flow, and the filter speed was controlled by a flowmeter to ensure uniform filtration. The backwashing strength was controlled by the expansion rate of filter media, using about 15–20%. The influent and backwash water were both from the aeration tank.

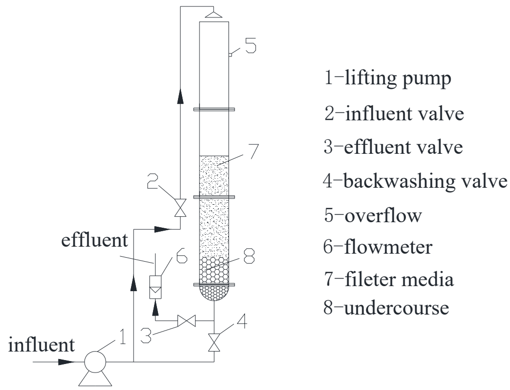


Figure 2. Flow chart of the filter column.

2.4. Analysis Methods

The influent was taken from the reserved sampling port of the inlet pipe of the cascade aeration tank and the effluent was taken from the outlet pipe of the biofilter. All samples were analyzed immediately in the laboratory of the Water Treatment Plant, so no storage measures were adopted. All iron and manganese analyses were performed by a spectrophotometer in accordance with standard methods [26]. Dissolved oxygen was measured by a Dissolved Oxygen Meter (oxi340i, WTW, Munich Germany), pH was determined using an electrode (pH/oxi340i, WTW, Munich Germany)

2.5. Filter Inoculation

Source of bacteria: From the engineering practice of a biological iron and manganese removal water treatment plant, it was found that autochthonous IMOB were contained in groundwater containing iron and manganese all over different regions, and they were adaptable for the local water environment. In the process of biofilter establishment, the local bacteria was added to the filter after amplification and cultivation, which showed strong adaptability and promoted the maturity of the biofilter as soon as possible and greatly shorten the cultivation period [23]. Therefore, in this study, the ferromanganese oxidizing bacteria in the mud of the aeration tank wall of the local groundwater treatment plant for iron and manganese removal were collected and cultured, and then inoculated into the filter.

Methods: A high concentration of bacterial liquid was used to inoculate the filter at one time, and a large number of bacteria were brought into the filter through some operation methods, then the filter was cultured at a low filtration rate. Under the condition of no medium and low nutrients, the bacteria were directly cultured by raw water. The bacteria proliferated slowly in the initial stage of culture, it was beneficial to further expand and cultivate the biological population suitable for the characteristics of raw water in the filter layer [23]. In order to study the effect of inoculation amount on the start-up of biofilter, filter 3# was selected to inoculate twice the amount of bacterial liquid among these six filters. When the effluent was not up to the standard, the biological filter was considered to be mature.

3. Results and Discussion

3.1. Full-Scale Filter Test

3.1.1. Initial Intermittent Operation Stage

The results are shown in Figure 3. The filters had a certain manganese removal capacity due to the inoculation at the initial stage of operation, and the manganese removal rate was about 50%; there were some differences between these 6 filters due to the initial difference of the filters and the difference of operation which could not be completely the same, but the trend was similar. After five months of operation, the manganese concentration in the effluent of the filter did not show a downward trend with time, it just fluctuated up and down, indicating that the number of ferromanganese oxidizing bacteria in the filter has not been proliferated and the activity was in an unstable stage [23].

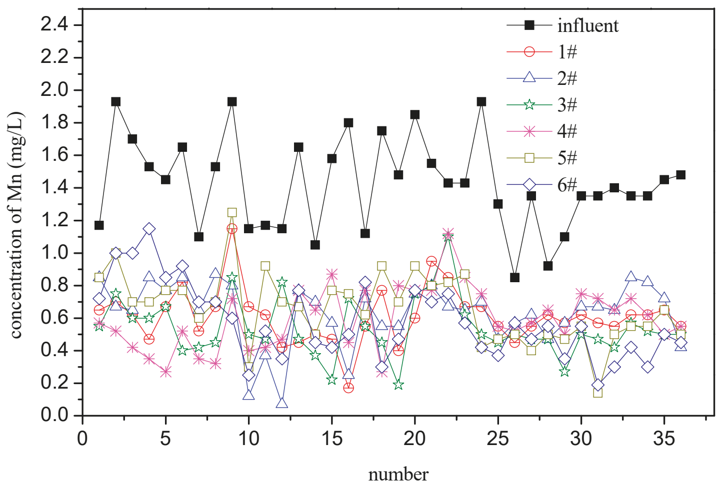


Figure 3. Mn in the influent and effluent (during the stage of intermittent operation; samples are taken once every three days of operation).

3.1.2. Continuous Operation Stage

The effluent quality of the filters showed in Figure 4 has been significantly improved in the continuous operation stage, but was not the same as the intermittent operation stage. The effluent of filter 3#, 5#, 6# came close to the standard after one month of continuous operation, and tended to reach the standard steadily in the next one month. Also, the effluent of filter 1#, 2#, 4# came close to the standard after 50 days of continuous operation, and then it took about 10 days to reach the standard. But at that time, the filters were not fully mature, so the effluent of filter1#, 2#, 4# fluctuated slightly.

Generally, after inoculation, microorganisms will undergo a period of adaptation under stable and suitable conditions so as to achieve a logarithmic growth period of rapid growth and rapid propagation. In this stage, the quality of effluent water will quickly improve with the mass propagation of microorganisms, and finally achieve the effluent standard [23]. But the appearance of the logarithmic growth phase depends entirely on the operation condition. The actual operation results showed that the intermittent operation in the first five months greatly delayed the microorganisms entering the logarithmic growth period, and thus delayed the maturity of the biological filter. In the subsequent filter column tests, it was verified that under the same conditions, the filter column did achieve maturity of the biological layer in about 40 days after continuous operation.

According to the figure, the continuous operation process could be divided into two stages. The first month was a period of rapid decline of the manganese concentration in the effluent, and the second month was a period of fluctuation of the effluent in a low manganese concentration. Although the inoculation amount of filter 3# was twice as much as that of the other five filters, and the removal efficiency of filter 3# was better than that of filters 1#, 2#, 4# in the first month. However, filter 6# has almost the same effect as filter 3#, and filter 5# did even slightly better; these differences might stem from the fact that it is impossible for the filter and the operation to be completely the same. Finally in the second month, the six filters all entered the same fluctuation stage with a low concentration of manganese, i.e., basically no difference.

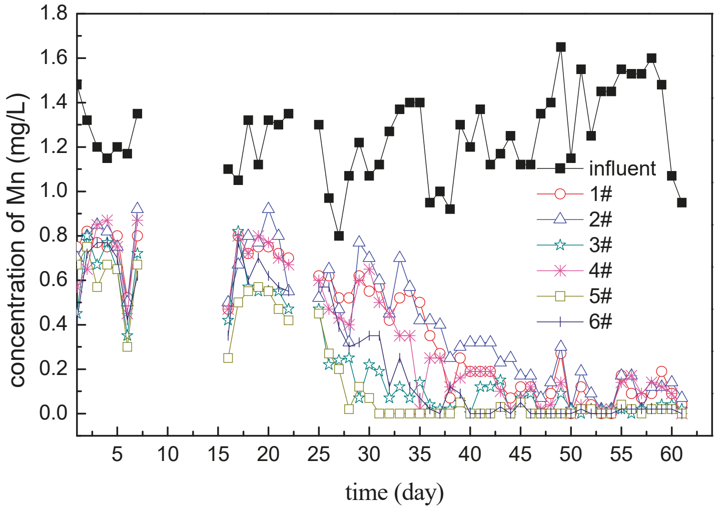


Figure 4. Mn in the influent and effluent (during the stage of continuous operation; samples are taken once every day of operation).

It can be inferred from the experimental results that the maturation of filter layer was not directly related to the amount of inoculation completely because filter 3# was better than 1#, 2#, 4# in the early stage, but 5#, 6# was similar to 3#, and even 6# was better than 3#. Therefore, in the future inoculation process, the appropriate amount is enough, not the more the better.

3.1.3. Speed-Up Stage

In the granular media filtration process, the shear force on the surface of the filter material rises with the increase in flow rate [27]. Therefore, in the beginning of the cultivation, a low filtration rate was usually adopted [28], in case of reducing the shear force to promote the adhesion and fixation of bacteria, which were in the mud or free in the water, onto the surface of the filter media. Therefore, 1.5 m/h was the initial filtration rate for the culture of biofilter in this plant, and the speed-up was carried out after the effluent of the filter reached the standard for one week. In the beginning, the increase for the speed-up was only 0.5 m/h. After the effluent reached the standard, the next speed-up started, and the filtration rate ranged from 1.5 m/h to 2 m/h to 2.5 m/h and then to 3 m/h. In this process (Figure 5), the effluent quality of individual filters fluctuated slightly for a short time, but soon the effluent quality recovered. Therefore, the speed-up was accelerated from 3 m/h to 4.5 m/h to 4.7 m/h and then to 5 m/h; in the first two days, the effluent of 6# fluctuated, but on the third day, it recovered to below the standard again. This indicates that the biological filter has been fully cultivated.

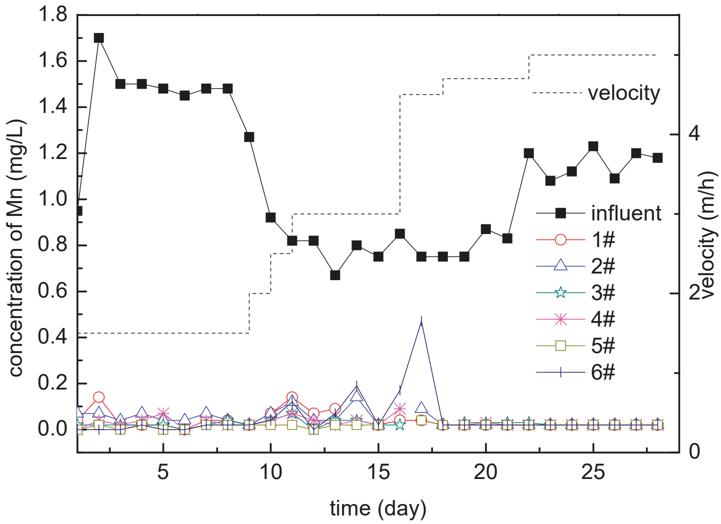


Figure 5. Mn in the influent and effluent (during the stage of speed-up; samples are taken once every day of operation).

3.2. Pilot-ScaleFilter Column Test

3.2.1. Start-Up of the Filter Column

The concentration of manganese in the effluent in Figure 6 fluctuated around 0.6 mg/L after the operation of the filter column, which is similar to the actual production filter, but the concentration of manganese in the effluent decreased gradually, and was finally below 0.1 mg/L within 40 days; the removal rate was as high as 95%, and all these phenomena indicate that the biological filter layer was mature. It can be found that the manganese variation curve of the effluent from the filter column was similar to that of the effluent from the production filter after continuous operation. Therefore, in the first four months of intermittent operation of the water plant, the cultivation of the biological filter almost had no effect. The reason could be that the intermittent operation could not provide a good environment for the microorganisms in the filter layer, so the microorganisms have been staying in the adaptive stage and have not entered the logarithmic growth stage [23].

3.2.2. Suspension and Rerun of the Filter Column

As has been noted, the intermittent operation had little effect on the cultivation of the biofilter. In order to find the difference between intermittent operation and continuous operation, the suspension and rerun tests of filter column were carried out, under the water quality of low iron and manganese and high iron and manganese with micro-pollution.

(1) Groundwater quality of low concentrations of iron and manganese

After several days of stopping operation, the biofilter column restarted, and the manganese removal capacity of the filter column decreased significantly (Figure 7). The manganese removal capacity of the first 50 cm filter layer decreased from 1.18 mg/L to 0.61 mg/L, only 50%. At the same time, it was found that the filter column was in an anoxic state with a DO of zero. Without the basic need of oxygen for the manganese oxidizing bacteria, their functions would decrease [29], so the manganese removal rate had a sharp decrease.

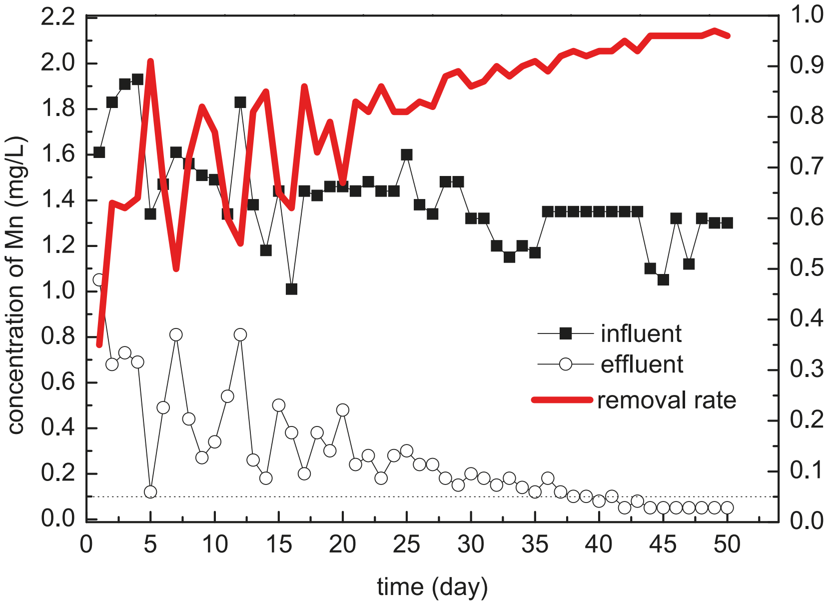


Figure 6. Mn removal and rate of the filter column.

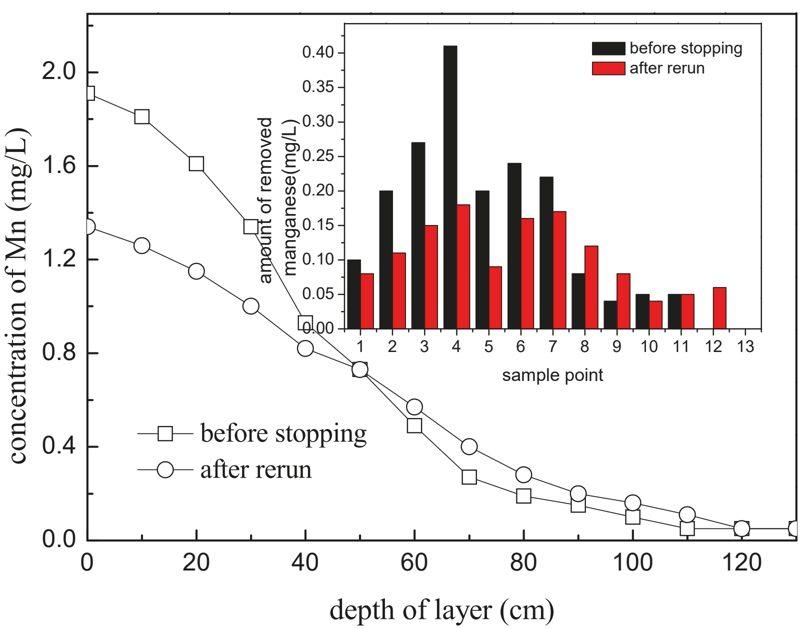


Figure 7. Mn concentration at the different height layers of the filter column before stopping and after rerun.

(2) Groundwater quality with a high concentration of iron and manganese

Two biofilter columns were used for the suspension and rerun test; during the suspension stage, one column was kept empty, and the other was full of raw water. Two distinct phenomena appeared as shown in the figure, when the two columns restarted to operate. For the column kept full of raw water (Figure 8), the manganese content in the effluent exceeded that of the influent, and it recovered slightly the next day, costing four days for the completely recovery (the data are not shown). The results showed that the dissolution of manganese occurred during the suspension process, while it was found that the dissolved oxygen in the water was zero. This two phenomena could be explained that due to the suspension of filter column with the situation of full of water, there was no way for the supplementation of oxygen for the filter layer. Therefore, when the dissolved oxygen in the layer was gradually depleted by reducing substances, it became anaerobic, which led to the dissolution of high valence manganese (Mn^{4+}), and the concentration of manganese in the effluent was higher than that in the influent. This phenomenon is extremely harmful to the cultivation of the biofilter layer, because most of the Manganese Oxidizing microorganisms construct their living space with manganese oxides batched on the filter material [23]. With the dissolution of manganese oxides, the biofilm of the filter layer will be greatly destroyed. Its recovery will also take a long time.

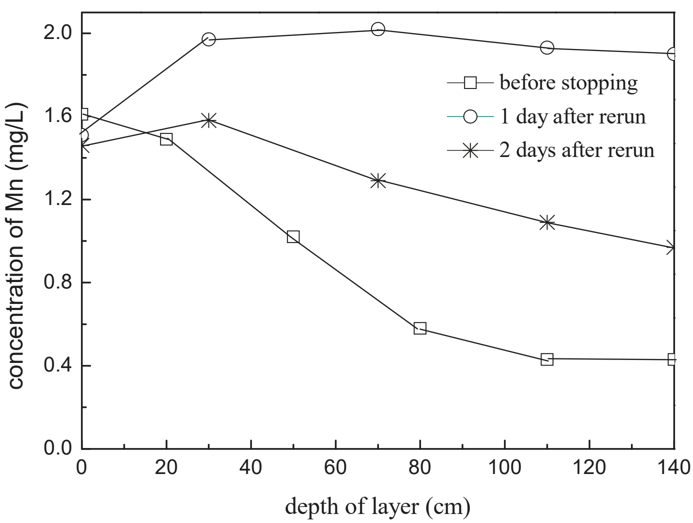


Figure 8. Mn concentration in the filtrate along filter depth (soaked state).

The air in the filter column kept empty (Figure 9) was in a circulation state, which made the filter column an aerobic environment, and provided oxygen for the microorganisms. Yet due to no supply of raw water, the microorganisms in the filter column stayed in a long-term starvation state. So when the filter restarted, the manganese in the raw water was utilized by microorganisms in large quantities, which led to the phenomenon of low manganese in the effluent water at the initial stage of the rerun operation of the filter column. After a few days of operation, the concentration of manganese in the effluent returned to its original state (data not given).

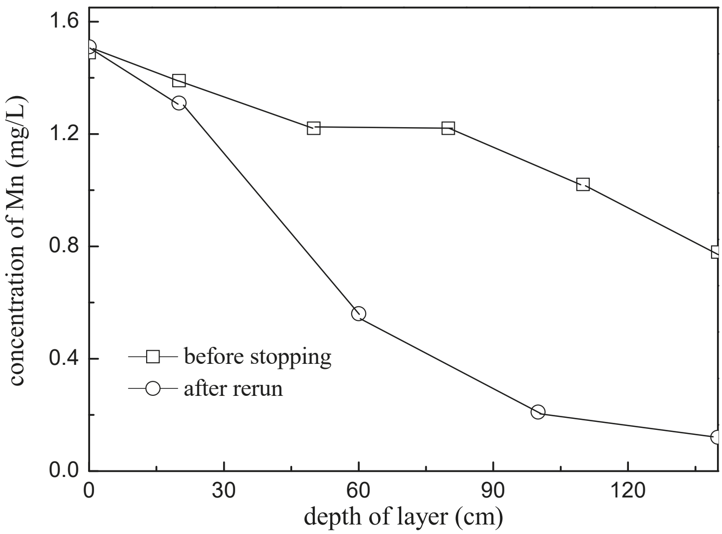


Figure 9. Mn concentration in the filtrate along filter depth (empty state).

(3) Reflections of the test of Suspension and Rerun

Generally, substrate and oxygen are two basic demands for the cultivation of microorganisms in the biological manganese removal filters. The shutdown process of filter without emptying will deplete the residual dissolved oxygen in the filter layer, which will eventually lead to the anaerobic state, thus affecting the activity and metabolism of microorganisms. Especially when there are some reducing substances in raw water, anaerobic conditions may cause the re-dissolution of manganese oxides which are an important structural component of biofilm, finally causing serious damage to the manganese removal biofilter. Therefore, it is suggested that in the process of cultivation of the biological filter or operation of mature filter, if it is necessary to shut down the tank, it should be kept empty, not full of water.

4. Conclusions

1. With the initial intermittent operation and the late continuous operation, the plant took 7 months to complete the start-up of biological iron and manganese removal filters.

2. By comparing the production filter and the filter column test, it can be inferred that although the intermittent operation reduced the backwashing frequency, no contribution was made to the cultivation of the biological manganese removal layer.

3. For the stopping operation of the biological filter in the start-up or stable operation stage, it is suggested to use filter emptying to ensure that the filter can be in an aerobic state, which is beneficial to the aerobic microorganisms and the biofilter layer.

Author Contributions: Formal analysis, H.Z.; Funding acquisition, J.Z. and D.L.; Investigation, H.Z.; Methodology, H.Z.; Project administration, J.Z.; Supervision, J.Z.; Writing—original draft, C.Y.; Writing—review & editing, H.Z.

Funding: We would like to gratefully acknowledge the National Natural Science Foundation of China (No. 51308009 and No. 51678006) and the Scientific and Technological Research Program of Beijing Municipal Education Commission project (KM201510005021) for financial support.

Conflicts of Interest: The authors declare no competing financial interest.

References

1. *Standards for Drinking Water Quality (GB5749-2006)*; Ministry of Health of the PRC: Beijing, China, 2007.
2. Civardi, J.; Tompeck, M. *Iron and Manganese Removal Handbook*, 2nd ed.; AWWA: Denver, CO, USA, 2015.
3. Dietrich, A.M. *EPA Secondary Maximum Contaminant Levels: A Strategy for Drinking Water Quality and Consumer Acceptability*; United States Environmental Protection Agency: Washington, DC, USA, 2015.
4. Wong, J.M. Chlorination-filtration for iron and manganese removal. *JAWWA* **1984**, *76*, 76–79. [CrossRef]
5. Knocke, W.R.; Hamon, J.R.; Thompson, C.P. Soluble manganese removal on oxide-coated filter media. *JAWWA* **1988**, *80*, 65–70. [CrossRef]
6. Ellis, D.; Bouchard, C.; Lantagne, G. Removal of iron and manganese from groundwater by oxidation and microfiltration. *Desalination* **2000**, *130*, 255–264. [CrossRef]
7. Aziz, H.A.; Smith, P.G. Removal of manganese from water using crushed dolomite filtration technique. *Water Res.* **1996**, *30*, 489–492. [CrossRef]
8. Frederick, W.P. Legislation/Regulation—Nitrate and Cancer: Is There a Link? *JAWWA* **1993**, *85*, 12–14.
9. Dominguez, H.L.; Andrea, T.; Manuela, A. Disinfection by-products formation and ecotoxicological effects of effluents treated with peracetic acid: A review. *Chemosphere* **2018**, *213*, 25–40. [CrossRef] [PubMed]
10. Hu, J.L.; Chu, W.H.; Sui, M.H. Comparison of drinking water treatment processes combinations for the minimization of subsequent disinfection by-products formation during chlorination and chloramination. *Chem. Eng. J.* **2018**, *335*, 352–361. [CrossRef]
11. Katsoyiannis, I.A.; Zouboulis, A.I. Biological treatment of Mn(II) and Fe(II) containing groundwater: Kinetic considerations and product characterization. *Water Res.* **2004**, *38*, 1922–1932. [CrossRef] [PubMed]
12. Sharma, S.K.; Petrusevski, B.; Schippers, J.C. Biological iron removal from groundwater: A review. *J. Water Supply Res. Technol. Aqua.* **2005**, *54*, 239–245. [CrossRef]
13. Yang, H.; Li, D.; Zeng, H.P.; Zhang, J. Autotrophic nitrogen conversion process and microbial population distribution in biofilter that simultaneously removes Fe, Mn and ammonia from groundwater. *Int. Biodeterior. Biodegrad.* **2019**, *135*, 53–61. [CrossRef]
14. Yang, H.; Li, D.; Zeng, H.P.; Zhang, J. Impact of Mn and ammonia on nitrogen conversion in biofilter coupling nitrification and ANAMMOX that simultaneously removes Fe, Mn and ammonia. *Sci. Total Environ.* **2019**, *648*, 955–961. [CrossRef] [PubMed]
15. Nuratiqah, M.; Hassimi, A.H.; Siti Rozaimah Sheikh, A. A review of biological aerated filters for iron and manganese ions removal in water treatment. *J. Water Proc. Eng.* **2018**, *23*, 1–12.
16. Breda, I.L.; Ramsay, L.; Roslev, P. Manganese oxidation and bacterial diversity on different filter media coatings during the start- up of drinking water biofilters. *J. Water Supply Res. Technol. Aqua.* **2017**, *66*, 641–650. [CrossRef]
17. Cai, Y.; Li, D.; Liang, Y.W.; Zeng, H.P. Operational parameters required for the start-up process of a biofilter to remove Fe, Mn, and NH_3-N from low-temperature groundwater. *Desalin. Water Treat.* **2016**, *57*, 3588–3596. [CrossRef]
18. Yang, H.; Li, D.; Zhang, J. Design of biological filter for iron and manganese removal from water. *J. Environ. Sci. Health Part A* **2004**, *39*, 1447–1454. [CrossRef]
19. Li, D.; Zhang, J.; Wang, H.T.; Yang, H.; Wang, B.Z. Operational Performance of Biological Treatment Plant for Iron and Manganese Removal. *J. Water Supply Res. Technol. Aqua.* **2005**, *54*, 15–24. [CrossRef]
20. Li, D.; Zhang, J.; Wang, H.T.; Chen, L.X.; Wang, H. Application of biological process to treat the groundwater with high concentration of iron and manganese. *J. Water Supply Res. Technol. Aqua.* **2006**, *55*, 313–320. [CrossRef]
21. Li, D. The Study on the Theory and Application in Engineering for Biological Removal of Iron and Manganese. Ph.D. Thesis, Beijing University of Technology, Beijing, China, 20 January 2005. (In Chinese)
22. Li, X.; Chu, Z.; Liu, Y.; Zhu, M.; Yang, L.; Zhang, J. Molecular characterization of microbial populations in full-scale biofilters treating iron, manganese and ammonia containing groundwater in Harbin, China. *Bioresour. Technol.* **2013**, *147*, 234–239. [CrossRef] [PubMed]
23. Zhang, J.; Li, D.; Yang, H.; Chen, L.X.; Gao, J. *Mechanism and Engineering Technology of Biological Manganese Fixation and Manganese Removal*; China Architecture & Building Press: Beijing, China, 2005. (In Chinese)
24. Tamara, S.; Marinko, M.; Felicita, B.; Laszlo, S. Rapid startup of biofilters for removal of ammonium, iron and manganese from ground water. *J. Water Supply Res. Technol. Aqua.* **2004**, *53*, 509–518.

25. Cai, Y.; Li, D.; Liang, Y.H.; Luo, Y.H.; Zeng, H.P.; Zhang, J. Effective start-up biofiltration method for Fe, Mn, and ammonia removal and bacterial community analysis. *Bioresour. Technol.* **2015**, *176*, 149–155. [CrossRef] [PubMed]
26. SEPA (State Environmental Protection Administration). *Monitoring and Analytical Methods for Water and Wastewater*, 4th ed.; China Environmental Science Press: Beijing, China, 2002. (In Chinese)
27. Edzwald, J.K. *Water Quality & Treatment: A Handbook on Drinking Water*, 6th ed.; American Water Works Association, McGraw-Hill: New York, NY, USA, 2011.
28. Li, D.; Cao, R.H.; Yang, H.; Wang, Y.J.; Lv, S.S.; Zhang, J. Start-up and operation of biofilter coupled nitrification and CANON for the removal of iron, manganese and ammonia nitrogen. *Huanjing Kexue* **2018**, *39*, 1264–1271. (In Chinese) [PubMed]
29. Cheng, Q.F. Competitive mechanism of ammonia, iron and manganese for dissolved oxygen using pilot-scale biofilter at different dissolved oxygen concentrations. *Water Sci. Technol. Water Supply* **2016**, *16*, 766–774. [CrossRef]

International Journal of
Environmental Research and Public Health

MDPI

Article

Effects of Food Contamination on Gastrointestinal Morbidity: Comparison of Different Machine-Learning Methods

Qin Song [1,*], Yu-Jun Zheng [2] and Jun Yang [3]

[1] Scientific Research Institute, Hangzhou Normal University, Hangzhou 311121, China
[2] Institute of Service Engineering, Hangzhou Normal University, Hangzhou 311121, China; yujun.zheng@computer.org
[3] School of Public Health, Zhejiang University, Hangzhou 310009, China; gastate@zju.edu.cn
* Correspondence: songqin@hznu.edu.cn

Received: 29 January 2019; Accepted: 3 March 2019; Published: 7 March 2019

Abstract: Morbidity prediction can be useful in improving the effectiveness and efficiency of medical services, but accurate morbidity prediction is often difficult because of the complex relationships between diseases and their influencing factors. This study investigates the effects of food contamination on gastrointestinal-disease morbidities using eight different machine-learning models, including multiple linear regression, a shallow neural network, and three deep neural networks and their improved versions trained by an evolutionary algorithm. Experiments on the datasets from ten cities/counties in central China demonstrate that deep neural networks achieve significantly higher accuracy than classical linear-regression and shallow neural-network models, and the deep denoising autoencoder model with evolutionary learning exhibits the best prediction performance. The results also indicate that the prediction accuracies on acute gastrointestinal diseases are generally higher than those on other diseases, but the models are difficult to predict the morbidities of gastrointestinal tumors. This study demonstrates that evolutionary deep-learning models can be utilized to accurately predict the morbidities of most gastrointestinal diseases from food contamination, and this approach can be extended for the morbidity prediction of many other diseases.

Keywords: food contamination; public health; gastrointestinal diseases; morbidity; deep neural networks; evolutionary learning

1. Introduction

In recent decades, industrial emissions, domestic waste, and the overuse of pesticides and fertilizers have caused serious environmental pollution, which has been confirmed as an important factor causing alarming deterioration in public health [1–5]. In particular, food contamination arising from soil and water pollution has been reported to be involved in almost all types of gastrointestinal diseases [6–8]. However, modeling the effects of food contamination on gastrointestinal morbidity is still a challenging task because the pathogenic mechanisms of gastrointestinal diseases are very complex, the number of contaminants is large, and the pathogenic roles of contaminants in the diseases are often unknown or uncertain.

There are numerous studies on the effects of environmental pollution on public health. A majority of studies have been devoted to the relationships between air pollution and respiratory diseases. Using logistic regression and weighted linear regression, Zhang et al. [9] examined the association between children's respiratory morbidity prevalence and district-specific ambient levels of main air pollutants in four Chinese cities, and their results evidenced that morbidity prevalence was positively associated with the levels of NO_x, SO_2, and coarse particles. Jayaraman and Nidhi

[10] used a generalized additive Poisson regression model to evaluate the association between air pollutants and daily variations in respiratory morbidity in Delhi in 2004–2005. Based on a log-linear Poisson regression model, Sousa et al. [11] performed time-series analysis to assess the impact of air pollution on emergency hospitalization for respiratory disease in Rio de Janeiro, Brazil, in 2000–2005. Zhao et al. [12] used a time-series model with a quasi-Poisson link to examine the association between PM pollution and respiratory morbidities in Dongguan City, China, in 2013–2015. Qiu et al. [13] used a similar approach to estimate the short-term effects of ambient air pollutants (PM_{10}, $PM_{2.5}$, NO_2, and SO_2) on hospital admissions of overall and cause-specific respiratory diseases in 17 cities of Sichuan Province, China, during 2015–2016. Although such regression models can demonstrate the associations between pollution and diseases, they are often incapable of providing sufficiently accurate morbidity prediction for healthcare management.

To overcome the limitation of classical linear and logistic models with multiple variables to handle the multifactorial effect, Bibi et al. [14] used an artificial neural network (ANN) to predict the effect of atmospheric changes on emergency department visits for respiratory symptoms. The results showed that the average prediction error of the ANN was much less than the classical models on the test set. Wang et al. [15] applied the Granger causality method to identify the main air pollutants correlated with the mortality of respiratory diseases, and then constructed an ANN model for respiratory mortality prediction in Beijing during 2005–2008, which also achieved higher accuracy than classical correlation-analysis methods. Junk et al. [16] used an ANN to predict the mortality rates of respiratory diseases associated with air pollution under different weather conditions in Western Europe. Moustris et al. [17] developed an ANN model to predict the weekly number of childhood asthma admission at the greater Athens area in Greece from ambient air-pollution data during 2001–2004. Zhu et al. [18] studied the effects of air pollutants on lower respiratory disease in Lanzhou City, China, during 2001–2005, and constructed an ANN based on a group method of data handling to forecast the number of patients in a hospital. Sundaram et al. [19] developed an Elman neural network to predict respiratory mortality and cardiovascular mortality from a set of air-pollution indicators, and the results showed that the dynamic ANN showed good performance on time-series prediction. Recently, Liu et al. [20] employed long short-term memory recurrent neural networks to forecast influenza trends from multiple data sources, including virologic surveillance, influenza geographic spread, Google trends, climate and air pollution; their results also exhibited high prediction accuracy.

Although it is known that many diseases are related to food contamination, studies on their correlations are relatively few, mainly because the number of food contaminations is much larger than the number of air pollutants, and thus classical regression methods and shallow ANNs become inefficient in handling complex correlations in such a high-dimensional feature space. Recently, deep neural networks (DNNs) are a powerful tool for modeling complex probabilistic distributions over a large number of influence factors by automatically discovering intermediate abstractions, layer by layer. Song et al. [21] developed a DNN based on a denoising autoencoder [22] to predict gastrointestinal-infection morbidity from food-contamination data in four counties in China during 2015–2016, and the results showed that the deep-learning model had significantly higher prediction accuracy than shallow ANNs. However, their work only concerned the morbidity of all acute gastrointestinal infections, i.e., it neither considered other gastrointestinal diseases such as chronic gastritis and gastrointestinal tumors, nor did it differentiate the morbidities of different gastrointestinal infections, such as acute gastritis and dysentery.

This study investigates the effects of food contamination on six main gastrointestinal diseases, acute gastroenteritis, chronic gastroenteritis, gastrointestinal ulcers, gastrointestinal tumors, food poisoning, and other acute gastrointestinal infections. We employed five methods, multiple linear regression (MLR), a three-layer feed-forward ANN, a deep belief network (DBN) [23], a deep autoencoder (DAE), and a deep denoising autoencoder (DDAE) [22], for correlation analysis and gastrointestinal-morbidity prediction. For each of the last three deep-learning methods, we respectively constructed two models, one using the basic gradient-based training algorithm and the other using an evolutionary training algorithm. Results showed that the deep-learning models achieved significantly higher accuracies than the MLR and shallow ANN models, and the DDAE with evolutionary training exhibited the highest prediction accuracy.

2. Materials and Methods

2.1. Materials

We collected data from ten cities/counties in central China, Yichun City (Yuanzhou Municipal District), Gao'an City, Wanzai County, Tonggu County, Pingxiang City (Anyuan Municipal District), Shangli County, Ji'an County, Xingan County, Liling City, and Chaling County, from May 2015 to September 2018 (178 weeks). These cities/counties have similar dietary habits and levels of health services. The dataset consists of two parts:

- Weekly food-contamination data from food-supervision departments. They include 119 types of food (given in Table 1) and 227 types of contaminants (given in Table 2). Therefore, the total number of contaminant indicators was at most 27,013. However, in practice, it is impossible to inspect so many contaminants, and thus the data tuples contain a large portion of missing values, and the average number of indicators per tuple is only approximately 4955.
- Weekly gastrointestinal-morbidity data from hospitals and healthcare-management departments. As aforementioned, these involve six general types of gastrointestinal diseases.

We constructed a data tuple per week for each city/county; the total number of tuples is 1780. If an indicator was measured more than once in a week, we took the mean value in the tuple.

Table 1. Types of food for morbidity prediction [21].

Class	Food
Cereals	Rice, wheat, barley, corn, millet, black rice, sticky rice
Beans	Soybean, mung soybean, red bean, black bean, broad bean, pea, cow pea, hyacinth bean, kidney bean, sword bead
Vegetables	Cabbage, pak choi cabbage, baby cabbage, celery cabbage, celery, lettuce, broccoli, Chinese broccoli, mustard leaf, leaf lettuce, okra, rape, spinach, water spinach, potherb mustard, amaranth, cauliflower, purslane, yam, carrot, celtuce, summer radish, loofah, tomato, cucumbers, lappa, radish, potato, sweet potato, pumpkin, bitter gourd, white gourd, chilli pepper, bell pepper, green pepper, sweet pepper, pod pepper, pea sprout, soybean sprout, mung bean sprout, Chinese toon sprout, shiitake, button mushroom, oyster mushroom, needle mushroom, agaric, day lily, tremella, spring onion, Chinese onion, ginger, caraway, garlic, fragrant-flowered garlic, garlic sprouts
Fruits	Apple, gala apple, bergamot pear, snow pear, mili pear, pineapple, orange, navel orange, vibrio mimicus, pomelo, peach, nectarine, melon, watermelon, Hami melon, apricot, plum, cherry, bayberry, grape, longan, lychee, winter jujube, red jujube, sugarcane, pitaya
Meals and eggs	Pork, beef, mutton, chicken, duck, egg, duck egg, quail egg
Aquatic	Kelp, laver, carp, grass carp, yellow croaker, perch, crucian, prawn, river prawn, crab, river crab, river snail

Table 2. All 227 contaminants used for morbidity prediction [21].

Class	Subclass	Contaminants
Inorganic contaminants	Heavy metals	Pb, Cd, Hg, Cu, Ni, As, Be, Bi, Sb, Tl, Cr, Mo, Ni, Zn, F, V
	Others	cyanide, nitrate, nitrite, sulfate, carbonate
Organic contaminants	Hydrocarbons	benzene series, polycyclic aromatic hydrocarbons, total petroleum hydrocarbon
	Halogenated	hydrochlorofluorocarbons, chlorinated solvents, polychlorinated biphenyls, dioxin
	Oxygenated	alcohols, phenols, ethers, esters, phthalate
	Dyes	Azo, quaternary ammoniun compounds, benzidine, naphthylamine
	Plastics	polypropylene, polyphenyl ether, polystyrene, phthalic acid esters
	Pesticides	66 commonly used pesticides [24]
	Herbicides	18 commonly used herbicides [25]
	Endocrine disruptors	68 chemicals [26]
	Others	trichloroethylene, organochlorine pesticide
Pathogenic organisms	Bacteria	salmonella, shigella, dysentery bacillus, plague bacillus, tubercle bacillus, typhoid bacillus, diphtheria bacillus, Francisella tularensis, Brucella, vibrio parahaemolyticus, vibrio cholerae, vibrio mimicus, vibrio fluvialis, clostridium tetani, clostridium botulinum, clostridium perfringens, staphylococcus aureus, Bacillus anthraci, Escherichia coli, Yersinia, helicobcter pylori, campylobacter jejuni, aeromonas hydrophila, roundworm eggs, hookworm eggs
	Fungi	candida albicans, aspergillus fumigatus, mucor racemosus
	Virus	rotavirusm, norovirus, sapovirus, astrovirus

2.2. Methods

We used eight machine-learning models for gastrointestinal-morbidity prediction based on food contamination. The aim of model training was to minimize the root mean squared error (RMSE) between the actual model outputs and the expected outputs over the training set:

$$\min \mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i|^2} \tag{1}$$

where N is the number of tuples in the training set, y_i is the model actual output of the i-th tuple, and $\hat{y}_i$ is the expected (labeled) output of the i-th tuple. In this study, the output morbidity y_i is calculated as the ratio of the incidences to the resident population in the investigated region (the floating population is not taken into account because of the difficulty of data collection).

A model is evaluated based on its prediction accuracy over the test set. We used fivefold cross-validation, i.e., we partitioned the dataset into five equal-size pieces, and ran the validation five times, each using four pieces as the training set and the remaining piece as the test set. Prediction accuracy was averaged over the five validations.

2.2.1. Multiple Linear Regression (MLR)

The MLR method calculates an output y from an n-dimensional input $\mathbf{x}$ as:

$$y = a_0 + \sum_{i=1}^{n} a_i x_i \tag{2}$$

where a_i are the regression coefficients ($i = 1, 2, \ldots, n$). Here, n = 27,013; if a value x_i is missing, it is filled by the mean value of those nonmissing x_i of training tuples.

2.2.2. Shallow Neural Network

We used a three-layer feed-forward ANN trained by the back-propagation algorithm. Each neuron in the input layer directly accepts an input component x_i, while each neuron j in the hidden layer calculates an inner output z_j as:

$$z_j = s\left(\sum_{i=1}^{n} w_{ij}x_i - \theta_j\right) \tag{3}$$

where θ_j is the threshold of the neuron, w_{ij} is the connection weight between the i-th input neuron to the neuron j, and s is the sigmoid activation function:

$$s(u) = \frac{1}{1 + e^{-u}} \tag{4}$$

Similarly, the output neuron calculates the final output y as:

$$y = s\left(\sum_{j=1}^{m} w_j z_j - \theta_0\right) \tag{5}$$

Empirically, we set number of neurons m in the hidden layer to $\sqrt{n}$.

2.2.3. Deep Belief Network (DBN)

A DBN [23] consists of a stack of Restricted Boltzmann Machines (RBMs) [27]. An RBM, consisting of a visible input layer and a hidden layer, is an energy-based probabilistic model that defines a joint probability distribution over an input vector $\mathbf{x}$ and a hidden vector $\mathbf{z}$ as:

$$P(\mathbf{x}, \mathbf{z}) = \frac{1}{\sum_{i=1}^{N} \exp(-E(\mathbf{x}_i, \mathbf{z}_i))} \exp(-E(\mathbf{x}, \mathbf{z})) \tag{6}$$

where $E(\mathbf{x}, \mathbf{z}) = -\mathbf{x}^{\mathrm{T}}\mathbf{b}\mathbf{x} - \mathbf{z}^{\mathrm{T}}\mathbf{c}\mathbf{z} - \mathbf{x}^{\mathrm{T}}\mathbf{w}\mathbf{z}$, and $\mathbf{b}$, $\mathbf{c}$, and $\mathbf{w}$ are the parameter vectors representing visible-to-visible, hidden-to-hidden, and visible-to-hidden interaction weights, respectively. Note that a basic RBM learns distributions over binary vectors, but we can use Gaussian-Bernoulli energy function to transform a real vector into a binary one [28], and then use DBN to learn distributions over the transformed binary vector [29].

After fine-tuning the structural parameters of the DBN on the training sets, we set the number of hidden layers to four, and set the numbers of neurons in the hidden layers to 3860, 550, 80, and 12, respectively. A Gaussian mixture model was added to the topmost RBM of DBN to produce output morbidity y from topmost hidden vector $\mathbf{z}$. DBN training consists of two stages. The first stage is pretraining, which tries to maximize the joint distribution of each RBM over the training set layer-by-layer:

$$\arg\max_{\mathbf{b},\mathbf{c},\mathbf{w}} \mathcal{J} = \frac{1}{N}\sum_{i=1}^{N} \log P(\mathbf{x}_i, \mathbf{z}_i) \tag{7}$$

The second stage is to minimize the RMSE of the whole DBN over the training set.

2.2.4. Evolutionary Deep Belief Network (EvoDBN)

A classical DBN is trained by a gradient-based, layerwise training algorithm [30], which is easily trapped in local optima, especially when the dimension is high. This issue can be tackled by using evolutionary training algorithms, which evolve populations of solutions to simultaneously explore multiple regions in the solution space to increase the chances of jumping out of local optima [31]. Here, we employed a recent efficient evolutionary algorithm called water wave optimization (WWO) [32], which has exhibited competitive performance compared to many other popular evolutionary algorithms in neural-network training [33].

To solve an optimization problem, WWO evolves a population of candidate solutions by mimicking wave propagation and breaking in shallow water. In WWO, each solution X is analogous to a wave. The higher the energy (fitness) $f(X)$, the smaller the wavelength λ_X, and thus the smaller the range that the wave propagates. λ_X is initially set to 0.5, and then updated at each generation as:

$$\lambda_X = \lambda_X \cdot \alpha^{-(f(X)-f_{\min}+\epsilon)/(f_{\max}-f_{\min}+\epsilon)} \tag{8}$$

where $f_{\max}$ and $f_{\min}$ are the maximum and minimum fitness among the population, respectively, α is the wavelength-reduction coefficient suggested set to 1.0026, and ϵ is a very small number to avoid division by zero. At each generation, X is propagated by adding an offset proportional to λ_X to each dimension X_i as follows:

$$X_i' = X_i + \lambda_X \cdot rand(-1,1) \cdot L_i \tag{9}$$

where L_i is the length of the i-th dimension of the solution space. Whenever a propagation produces a new best solution X^*, it is broken into several solitary waves, each of which moves a small distance from X^* in a random dimension i:

$$X_i' = X_i^* + \mathcal{N}(0,1) \cdot \beta L_i \tag{10}$$

where β is the breaking coefficient, and $\mathcal{N}$ denotes a normal distribution. The best solitary wave, if better than X^*, replaces X^* in the population.

The EvoDBN uses the same architecture as DBN, and also employs a Gaussian mixture model to produce output morbidity. When training EvoDBN, WWO is first applied to optimize the $\{\mathbf{b}, \mathbf{c}, \mathbf{w}\}$ parameters of each RBM layer by layer, where $f(X)$ corresponds to the objective function in Equation (7). After pretraining, WWO is applied to optimize the parameters of the DBN as a whole, where $f(X)$ is inversely proportional to RMSE.

2.2.5. Deep Autoencoder (DAE)

An autoencoder also consist of a visible input layer (called an encoder) and a hidden layer (called a decoder). It first transforms (encodes) an input vector $\mathbf{x}$ to a hidden representation $\mathbf{z}$ through affine mapping

$$\mathbf{z} = s(\mathbf{wx} + \mathbf{b}) \tag{11}$$

and then maps (decodes) $\mathbf{z}$ back to a reconstructed vector $\mathbf{x}'$ in the input space:

$$\mathbf{x}' = s(\mathbf{w}'\mathbf{z} + \mathbf{b}') \tag{12}$$

The aim of autoencoder training is to minimize the average reconstruction error over the training set:

$$\arg\min_{\mathbf{w},\mathbf{b},\mathbf{w}',\mathbf{b}'} \mathcal{L} = \frac{1}{N}\sum_{i=1}^{N} \|\mathbf{x}, \mathbf{x}'\|^2 \tag{13}$$

A DAE [23] consists of a stack of autoencoders. Its training consists of two stages. The first stage is to train each autoencoder layer by layer, and the second stage is to train the whole DAE to minimize the RMSE over the training set.

For the morbidity-prediction problem, we used a DAE with four hidden layers, and tuned the numbers of neurons in the hidden layers to 4500, 640, 80, and 12, respectively. It also employed a Gaussian mixture model to produce output morbidity.

2.2.6. Evolutionary DAE (EvoDAE)

Similarly, we implemented a DAE trained by the WWO evolutionary algorithm, which is first applied to minimize the reconstruction error in Equation (13) of each autoencoder layer by layer, and then applied to minimize the RMSE of the whole DAE. The EvoDAE uses the same structure (including the top-level Gaussian mixture model) as DAE.

2.2.7. Deep Denoising Autoencoder (DDAE)

A denoising autoencoder is a variant of the basic autoencoder. It first randomly adds some noise to an initial input vector $\mathbf{x}$ to form a corrupted $\widetilde{\mathbf{x}}$, and then encodes $\widetilde{\mathbf{x}}$ to a hidden representation $\mathbf{z}$, which is then decoded to a reconstructed $\mathbf{x}'$. The aim of denoising-autoencoder training is to reconstruct a clean "repaired" $\mathbf{x}'$ from a corrupted $\widetilde{\mathbf{x}}$, which can still be represented by Equation (13). The key difference is that $\mathbf{z}$ is deterministic mapping of $\widetilde{\mathbf{x}}$ and thus the result of a stochastic mapping of $\mathbf{x}$.

Similarly, a DDAE [22] consists of a stack of denoising autoencoders. Its training consists of two stages. The first stage is to train each denoising autoencoder layer by layer, and the second stage is to train the whole DDAE to minimize the RMSE over the training set. For our prediction problem, the DDAE model uses the same structure (including the top-level Gaussian mixture model) as DAE.

2.2.8. Evolutionary DDAE (EvoDDAE)

Similarly, we implemented a DDAE trained by the WWO evolutionary algorithm, which is first applied to minimize the reconstruction error of each denoising autoencoder layer by layer, and then applied to minimize the RMSE of the whole DDAE. The EvoDDAE model uses the same structure as DDAE.

3. Results

According to historical experience, the weekly morbidities of acute gastroenteritis, food poisoning, and other acute gastrointestinal infections are predicted based on food-contamination data one week before. However, the time-lag effects of food contamination on chronic gastroenteritis, gastrointestinal ulcers, and gastrointestinal tumors are unknown. Therefore, we first tested the RMSE of the models for predicting the morbidities of the three types of diseases with a time lag of 1–8 weeks, respectively. Results are given in Figure 1, from which we can observe that:

- For chronic gastroenteritis, ANN and EvoDAE achieved the best RMSE when the lag was 2–3 weeks; DBN, EvoDBN, and DDAE achieved the best RMSE when lag was 3–4 weeks; DAE achieved the best RMSE when lag was 5–6 weeks; EvoDDAE achieved the best RMSE when lag was 3–5 weeks; and MLR showed good performance when lag was 3, 5, or 8 weeks (more irregular than other models).
- For gastrointestinal ulcers, ANN, DDAE, and EvoDDAE achieved the best RMSE when lag was 3–4 weeks; DAE and EvoDAE achieved the best RMSE when lag was 2–3 weeks; DBN achieved the best RMSE when lag was 4–5 weeks; EvoDBN achieved the best RMSE when lag was 3–5 weeks; and MLR showed good performance when lag was 4 or 6 weeks.
- For gastrointestinal tumors, the time-lag effect greatly varied among the models.

Consequently, we chose a time lag of three weeks for predicting the morbidities of both chronic gastroenteritis and gastrointestinal ulcers. For gastrointestinal tumors, because we could not determine an appropriate time lag for most models, we determined a different time lag for each model that resulted in the best RMSE for the model (6, 2, 6, 2, 1, 6, 5, and 1 week(s) for MLR, ANN, DBN, EvoDBN, DAE, EvoDAE, DDAE, and EvoDDAE, respectively).

Figure 2a–f presents the prediction accuracies of the models for the six gastrointestinal diseases, respectively. Results show that the traditional MLR exhibits the worst prediction performance on all diseases, the shallow ANN exhibits significantly better performance than MLR, and all deep-learning models exhibited much better performance than the MLR and shallow ANN. Among the six deep models, EvoDDAE exhibited the best performance on five diseases except gastrointestinal tumors. The average prediction accuracy of EvoDDAE was over 80% on acute gastroenteritis and food poisoning, close to 80% on other gastrointestinal infections, and approximately 72%–73% on chronic gastroenteritis and gastrointestinal ulcers. For gastrointestinal tumors, except that EvoDBN obtained an average perdition accuracy of approximately 52%, the accuracies of all other models were less than 50%, which indicates that the gastrointestinal-tumor morbidity is difficult to predict using these models. We also observed that, in most cases, the performance of a deep model could be significantly improved by using evolutionary training to replace traditional gradient-based training.

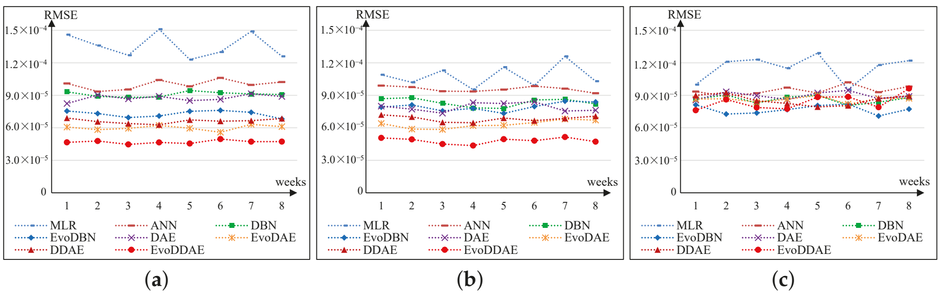


Figure 1. Prediction root mean squared error (RMSE) of the eight models using different time lags. x-axis denotes the time lag in weeks, and y-axis denotes the RMSE. (**a**) Chronic gastroenteritis; (**b**) gastrointestinal ulcers; (**c**) gastrointestinal tumors. MLR: multiple linear regression; ANN: artificial neural network; DBN: deep belief network; EvoDBN: evolutionary DBN; DAE: deep autoencoder; EvoDAE: evolutionary DAE; DDAE: deep denoising autoencoder; EvoDDAE: evolutionary DDAE.

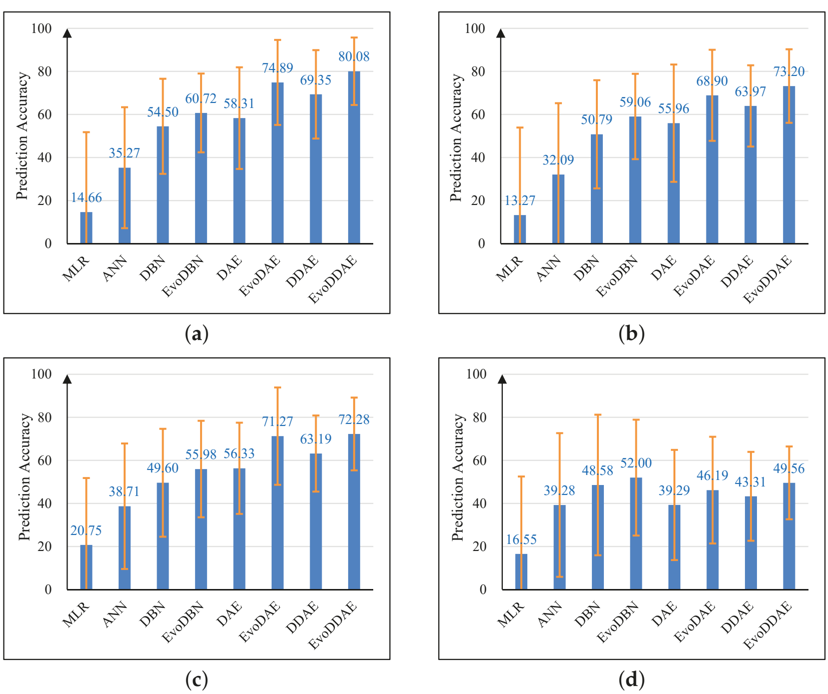


Figure 2. *Cont.*

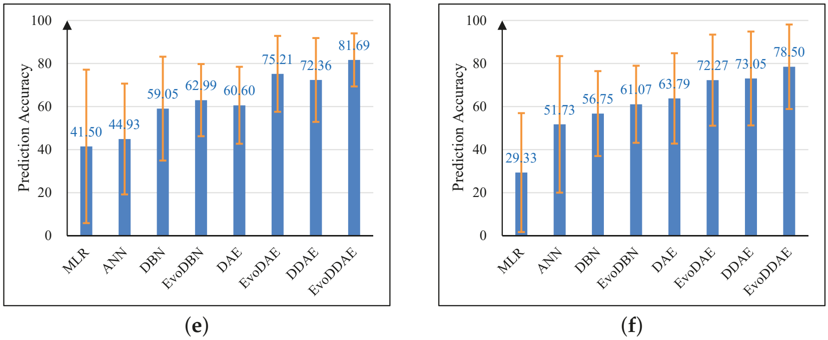


Figure 2. Accuracies of the models for gastrointestinal morbidity prediction. (**a**) Acute gastroenteritis; (**b**) chronic gastroenteritis; (**c**) gastrointestinal ulcers; (**d**) gastrointestinal tumors; (**e**) food poisoning; (**f**) other gastrointestinal infections.

4. Discussion

This study constructed and compared eight models for predicting the morbidities of six main gastrointestinal diseases from food contamination. Results demonstrate that some deep-learning models can achieve relatively high prediction accuracy. However, this does not mean that gastrointestinal diseases are mainly caused by food contamination, or that gastrointestinal morbidities in a region are mainly determined by the levels of food contamination. In fact, the relationships between food contamination and gastrointestinal morbidities can be highly complex and probabilistic, and morbidities are also affected by many other factors, such as the dietary habits and working pressures of inhabitants, and the levels of health services of that society. Our study reveals that, given a large number of historical data of food contamination and gastrointestinal morbidities in a region, we could use deep neural networks to learn such highly complex and probabilistic relationships. After sufficient training, we could obtain models that embed other influencing factors into model parameters, and thus output relatively accurate morbidities from food-contamination inputs. Consequently, the prediction results would be very useful to improve healthcare services.

In general, the traditional MLR model is incapable of learning complex relationships for morbidity prediction. According to our results, its average prediction accuracy is below 20% on most diseases. For food poisoning, MLR achieves the highest prediction accuracy of 41.5%, which is also significantly less than the seven other models. The low performance of MLR indicates that relationships between food contamination and gastrointestinal morbidities are highly nonlinear and probabilistic, which is beyond the capability of the linear model.

The shallow ANN model performs much better in approximating nonlinear relationships. However, its average prediction accuracy is only between 30% and 40% in most cases, which is still too low for medical management. This is mainly because the number of food-contamination indicators is large, and the generalization ability of the classical three-layer structure of ANN decreases dramatically with increasing dimension.

DNN models can effectively overcome the limitations of the MLR and shallow ANN models, as they can learn complex probabilistic distributions over a large number of influence factors by automatically discovering intermediate abstractions layer by layer. Comparing DBN and DAE, two of the most widely used DNNs, DAE achieved higher accuracies than DBN on five gastrointestinal diseases, while DBN only achieved higher accuracy on gastrointestinal tumors. This indicates that the energy-based probabilistic model of DBN is less effective than the reconstruction-error minimization model of DAE in morbidity prediction. By introducing the denoising learning mechanism into DAE, DDAE achieved significantly higher accuracies than DBN and DAE on all gastrointestinal diseases.

This is because the food-contamination data inevitably contain much noise, which can often mislead the learning process of DAE, while DDAE is much more robust in handling noisy inputs.

It was also observed that the prediction performance of all three DDNs could be significantly improved by equipping them with evolutionary training algorithms, because gradient-based training algorithms are easily trapped in local optima. An evolutionary algorithm uses a population of candidate solutions to simultaneously explore the search space; if some solutions are trapped in local optima, others can still explore other regions and help the trapped solutions jump out of local optima. Consequently, evolutionary DNNs can effectively suppress premature convergence and exhibit high learning abilities. Among the eight models, EvoDDAE that combines DDAE with evolutionary learning exhibited the best performance for morbidity prediction.

Among the six main types of gastrointestinal diseases, the prediction accuracies on three types of acute diseases are generally higher than other diseases, because the pathogenic mechanisms of acute diseases are relatively simpler, and their time-lag effects are easier to determine. That is why all models achieved the highest prediction accuracies on food poisoning, which is considered as "the most acute" disease. Among the diseases, each DNN model achieved the lowest prediction accuracy on gastrointestinal tumors, mainly because the pathogenic mechanisms of tumors are more complex than other diseases, and thus their correlation with food contamination is much weaker or is much difficult to learn.

5. Conclusions

This study compared eight machine-learning models for predicting the morbidities of six main gastrointestinal diseases from food-contamination data. Experiments on the datasets from ten cities/counties in central China demonstrate that the DNN models achieved significantly higher accuracies than the classical MLR and shallow ANN models, and the DDAE model with evolutionary learning exhibited the best prediction performance. Results also indicate that model accuracies are generally higher on acute gastrointestinal diseases than on other diseases, but it is difficult to predict the morbidities of gastrointestinal tumors. Moreover, a drawback of DNN models is that it takes significant effort to tune the structural parameters of the networks.

The studied deep-learning models could be utilized for the morbidity prediction of many other diseases whose influencing factors are large and complex. However, DNNs typically need to be trained on a large amount of labeled data, but disease- and health-related data are often very limited. Thus, we are currently studying unsupervised and transfer-learning technologies [34] for adapting the models from some well-known diseases to other diseases with insufficient data. Our future work also includes integrating the deep-learning models with fuzzy systems to handle uncertain information in the data [35,36], and utilizing the morbidity-prediction results for improving medical services, such as for medical-resource preparation and drug-procurement planning [37]. We believe that the combination of emerging deep-learning and intelligent decision-making technologies can significantly improve our society's healthcare services.

Author Contributions: Conceptualization, Q.S. and Y.-J.Z.; methodology, Y.-J.Z.; software, Y.-J.Z.; validation, Q.S.; formal analysis, J.Y.; investigation, Q.S.; resources, J.Y.; data curation, Q.S. and Y.-J.Z.; writing—original draft preparation, Q.S.; writing—review and editing, Y.-J.Z.; visualization, Q.S.; supervision, J.Y.; project administration, J.Y.; funding acquisition, Y.-J.Z.

Funding: This research was funded by the National Natural Science Foundation of China under grants number 61872123 and 61473263.

Acknowledgments: The authors would like to thank the Institute of Yichun Agricultural Science and Yichun University, China, for their help in data acquisition and processing.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial neural network
DNN	Deep neural network
MLR	Multiple linear regression
RMSE	Root mean squared error
DBN	Deep belief network
RBM	Restricted Boltzmann machine
DAE	Deep autoencoder
DDAE	Deep denoising autoencoder
WWO	Water wave optimization

References

1. Kiessling, K.L. Conference on the Aral Sea: Women, Children, Health and Environment. *Ambio* **1998**, *27*, 560–564.
2. Kelly, F.J.; Fussell, J.C. Air pollution and public health: Emerging hazards and improved understanding of risk. *Environ. Geochem. Health* **2015**, *37*, 631–649. [CrossRef] [PubMed]
3. Landrigan, P.J.; Fuller, R. Global health and environmental pollution. *Int. J. Public Health* **2015**, *60*, 761–762. [CrossRef] [PubMed]
4. Lu, Z.N.; Chen, H.; Hao, Y.; Wang, J.; Song, X.; Mok, T.M. The dynamic relationship between environmental pollution, economic development and public health: Evidence from China. *J. Clean. Prod.* **2017**, *166*, 134–147. [CrossRef]
5. Mannucci, P.M.; Franchini, M. Health Effects of Ambient Air Pollution in Developing Countries. *Int. J. Environ. Res. Public Health* **2017**, *14*. [CrossRef] [PubMed]
6. Lu, Y.; Song, S.; Wang, R.; Liu, Z.; Meng, J.; Sweetman, A.J.; Jenkins, A.; Ferrier, R.C.; Li, H.; Luo, W.; et al. Impacts of soil and water pollution on food safety and health risks in China. *Environ. Int.* **2015**, *77*, 5–15. [CrossRef] [PubMed]
7. Pei, X.; Li, N.; Guo, Y.; Liu, X.; Yan, L.; Li, Y.; Yang, S.; Hu, J.; Zhu, J.; Yang, D. Microbiological Food Safety Surveillance in China. *Int. J. Environ. Res. Public Health* **2015**, *12*, 10662–10670. [CrossRef] [PubMed]
8. Song, Q.; Zhao, M.R.; Zhou, X.H.; Xue, Y.; Zheng, Y.J. Predicting gastrointestinal infection morbidity based on environmental pollutants: Deep learning versus traditional models. *Ecol. Ind.* **2017**, *82*, 76–81. [CrossRef]
9. Zhang, J.J.; Hu, W.; Wei, F.; Wu, G.; Korn, L.R.; Chapman, R.S. Children's respiratory morbidity prevalence in relation to air pollution in four Chinese cities. *Environ. Health Perspect.* **2002**, *110*, 961–967. [CrossRef] [PubMed]
10. Jayaraman, G.N. Air pollution and associated respiratory morbidity in Delhi. *Health Care Manag. Sci.* **2008**, *11*, 132–138, doi:10.1007/s10729-007-9050-7. [CrossRef] [PubMed]
11. Sousa, S.; Pires, J.; Martins, E.; Fortes, J.; Alvim-Ferraz, M.; Martins, F. Short-term effects of air pollution on respiratory morbidity at Rio de Janeiro—Part II: Health assessment. *Environ. Int.* **2012**, *43*, 1–5. [CrossRef] [PubMed]
12. Zhao, Y.; Wang, S.; Lang, L.; Huang, C.; Ma, W.; Lin, H. Ambient fine and coarse particulate matter pollution and respiratory morbidity in Dongguan, China. *Environ. Pollut.* **2017**, *222*, 126–131. [CrossRef] [PubMed]
13. Qiu, H.; Yu, H.; Wang, L.; Zhu, X.; Chen, M.; Zhou, L.; Deng, R.; Zhang, Y.; Pu, X.; Pan, J. The burden of overall and cause-specific respiratory morbidity due to ambient air pollution in Sichuan Basin, China: A multi-city time-series analysis. *Environ. Res.* **2018**, *167*, 428–436. [CrossRef] [PubMed]
14. Bibi, H.; Nutman, A.; Shoseyov, D.; Shalom, M.; Peled, R.; Kivity, S.; Nutman, J. Prediction of emergency department visits for respiratory symptoms using an artificial neural network. *Chest* **2002**, *122*, 1627–1632. [CrossRef] [PubMed]
15. Wang, Q.; Liu, Y.; Pan, X. Atmosphere pollutants and mortality rate of respiratory diseases in Beijing. *Sci. Total Environ.* **2008**, *391*, 143–148. [CrossRef] [PubMed]
16. Junk, J.; Krein, A.; Helbig, A. Mortality rates and air pollution levels under different weather conditions: An example from Western Europe. *Int. J. Environ. Waste Manag.* **2009**, *4*, 197–212. [CrossRef]

17. Moustris, K.P.; Douros, K.; Nastos, P.T.; Larissi, I.K.; Anthracopoulos, M.B.; Paliatsos, A.G.; Priftis, K.N. Seven-days-ahead forecasting of childhood asthma admissions using artificial neural networks in Athens, Greece. *Int. J. Environ. Health Res.* **2012**, *22*, 93–104. [CrossRef] [PubMed]
18. Zhu, W.; Wang, J.; Zhang, W.; Sun, D. Short-term effects of air pollution on lower respiratory diseases and forecasting by the group method of data handling. *Atmosph. Environ.* **2012**, *51*, 29–38. [CrossRef]
19. Sundaram, N.M.; Sivanandam, S.; Subha, R. Elman neural network mortality predictor for prediction of mortality due to pollution. *Int. J. Appl. Eng. Res.* **2016**, *11*, 1835–1840.
20. Liu, L.; Han, M.; Zhou, Y.; Wang, Y. LSTM Recurrent Neural Networks for Influenza Trends Prediction. In *Bioinformatics Research and Applications*; Zhang, F., Cai, Z., Skums, P., Zhang, S., Eds.; Springer: Cham, Switzerland, 2018; pp. 259–264.
21. Song, Q.; Zheng, Y.J.; Xue, Y.; Sheng, W.G.; Zhao, M.R. An evolutionary deep neural network for predicting morbidity of gastrointestinal infections by food contamination. *Neurocomputing* **2017**, *226*, 16–22. [CrossRef]
22. Vincent, P.; Larochelle, H.; Bengio, Y.; Manzagol, P.A. Extracting and Composing Robust Features with Denoising Autoencoders. In Proceedings of the 25th International Conference on Machine Learning, Helsinki, Finland, 5–9 June 2008; pp. 1096–1103.
23. Hinton, G.E.; Osindero, S.; Teh, Y.W. A fast learning algorithm for deep belief nets. *Neural Comput.* **2006**, *18*, 1527–1554. [CrossRef] [PubMed]
24. Li, H.; Huang, K. *Mannual of Organic Pesticides and Intermediate Mass Spectrum*; Chemical Industry Press: Beijing, China, 2009.
25. Zhou, M.; Li, L.; Dai, H.; Zhang, X. *Heterocyclic Pesticides: Herbicides*; Since Press: Beijing, China, 2014.
26. Keith, L.H. Environmental endocrine disruptors: An overview of the analytical challenge. In Proceedings of the 13th Annual Waste Testing and Quality Assurance Symposium, Washington, DC, USA, 6–9 July 1997.
27. Taylor, G.W.; Hinton, G.E. Factored Conditional Restricted Boltzmann Machines for Modeling Motion Style. In Proceedings of the 26th Annual International Conference on Machine Learning, Montreal, QC, Canada, 14–18 June 2009; pp. 1025–1032.
28. Hinton, G.E.; Salakhutdinov, R.R. Reducing the dimensionality of data with neural networks. *Science* **2006**, *313*, 504–507. [CrossRef] [PubMed]
29. Zheng, Y.J.; Sheng, W.G.; Sun, X.M.; Chen, S.Y. Airline passenger profiling based on fuzzy deep machine learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *28*, 2911–2923. [CrossRef] [PubMed]
30. Bengio, Y.; Lamblin, P.; Popovici, D.; Larochelle, H. Greedy layer-wise training of deep networks. In Proceedings of the 19th International Conference on Neural Information Processing Systems, Cambridge, MA, USA, 4–7 December 2006; pp. 153–160.
31. Zheng, Y.; Ling, H.; Xue, J.; Chen, S. Population classification in fire evacuation: A multiobjective particle swarm optimization approach. *IEEE Trans. Evol. Comput.* **2014**, *18*, 70–81. [CrossRef]
32. Zheng, Y.J. Water wave optimization: A new nature-inspired metaheuristic. *Comput. Oper. Res.* **2015**, *55*, 1–11. [CrossRef]
33. Zhou, X.H.; Xu, Z.G.; Zhang, M.X.; Zheng, Y.J. Water wave optimization for articial neural network parameter and structure optimization. In Proceedings of the 13th International Conference, Beijing, China, 2–4 November 2018.
34. Bengio, Y. Deep learning of representations for unsupervised and transfer learning. In Proceedings of the ICML Workshop on Unsupervised and Transfer Learning, Bellevue, WA, USA, 2 July 2012; Volume 27, pp. 17–37.
35. Zheng, Y.J.; Ling, H.F.; Chen, S.Y.; Xue, J.Y. A hybrid neuro-fuzzy network based on differential biogeography-based optimization for online population classification in earthquakes. *IEEE Trans. Fuzzy Syst.* **2015**, *23*, 1070–1083. [CrossRef]
36. Zheng, Y.J.; Chen, S.Y.; Xue, Y.; Xue, J.Y. A Pythagorean-type fuzzy deep denoising autoencoder for industrial accident early warning. *IEEE Trans. Fuzzy Syst.* **2017**, *25*, 1561–1575. [CrossRef]
37. Song, Q.; Zheng, Y.; Huang, Y.; Xu, Z.; Sheng, W.; Yang, J. Emergency drug procurement planning based on big-data driven morbidity prediction. *IEEE Trans. Ind. Inform.* **2018**. [CrossRef]

International Journal of
Environmental Research and Public Health

MDPI

Article

Youth Perspectives of Healthcare in Central Mexico: An Application of Massey's Critical Health Literacy Framework

Steven Hoffman [1,*], Heidi Adams Rueda [2] and Lauren Beasley [3]

1 School of Social Work, Brigham Young University, Provo, UT 84660, USA
2 Department of Social Work, University of Texas at San Antonio, San Antonio, TX 78207, USA; heidi.rueda@utsa.edu
3 Knoxville Department of Kinesiology, Recreation, and Sport Studies, University of Tennessee, Knoxville, TN 37996, USA; lauren.beasley@utk.edu
* Correspondence: steven_hoffman@byu.edu; Tel.: +01-1801-422-3211

Received: 28 December 2018; Accepted: 1 March 2019; Published: 12 March 2019

Abstract: Attention to health literacy is essential more now than ever given the recognition, attention, and resources being dedicated to addressing health disparities throughout the world. Unfortunately, health literacy research is scarce in many parts of the world, particularly among youth. Using focus group discussions with junior high school students ($N = 98$) in a rural town of Central Mexico, we sought to learn about their experiences utilizing healthcare services at a local health clinic. The themes that naturally emerged from focus group discussions aligned with Massey's framework on critical health literacy among US youth, and included problems navigating the health system, embarrassment speaking to doctors about sensitive issues, and minimal importance being placed on preventative care. This suggests that Massey's framework may be appropriate to use when seeking to understand and promote health literacy among youth in Mexico. Furthermore, the challenges faced by adolescent participants in this study suggest that additional research is needed to assess how youth in other areas of Mexico are faring in efforts to understand and access their new and evolving universal healthcare system.

Keywords: health literacy; adolescent health; public health; Mexico; health disparities

1. Introduction

Health literacy (HL) is an individual's ability to gain, understand, process, and apply health information to make appropriate health decisions [1]. Health literacy has been found imperative for improved individual health. Many studies have shown a positive relationship between HL and health knowledge, such as knowledge of the negative effects of smoking and drinking and how to control chronic diseases such as HIV and hypertension [2]. Low HL, conversely, is associated with poor health, limited use of preventative care, and higher hospitalization rates [3]. Moreover, a recent literature review found that low HL is linked to low medication compliance and misunderstanding medical information, which helps explain findings linking low HL to decreased health among certain at-risk populations [4]. Unfortunately, the majority of HL measures, interventions, and research has been focused in the United States among primarily White populations [1,5,6]. Global HL research is essential more now than ever given the recognition, attention, and resources being dedicated to addressing health disparities throughout the world on an individual, community, and societal level [7–9]. A global effort to understanding HL is necessary when seeking to address the needs of migrant populations—particularly the growing migration networks connecting the U.S., Mexico, and Central America [8].

A focus on HL in Mexico is of particular importance at this time due to the country's recent implementation of universal healthcare [10]. Through a series of nationwide health reforms and financial restructuring between 2003 and 2012, universal healthcare in Mexico, known as Seguro Popular, is now utilized by over 60 million Mexican citizens. Prior to the beginning of reforms in 2003, the Mexican healthcare system was decentralized, and it fell on specific states to provide healthcare to the uninsured population. Often this meant high out-of-packet expenses for low-quality care, while insured Mexican citizens received healthcare from well-funded federal hospitals and clinics. This created inequalities in health services between the insured and uninsured. The goals of the healthcare reforms were to offer the equivalent of the federal health services to the uninsured population, reduce out-of-pocket-expenses, and take some of the financial responsibilities off the states. There were a large number of changes under the healthcare reform (for a full overview of the policy and further statistics, see Knaul et al. [11]); the current system (Seguro Popular) now works as a public fund under the Ministry of Health that the government, the states, and individual citizens (based off income, with lower-income families exempt) all pay into. The fund covers essential healthcare services of enrolled citizens who do not have health insurance through employment, creating an option of health coverage for all Mexican citizens [11]. While citizens have the choice of whether or not to enroll, many states have elected to automatically enroll all their citizens as government funding for healthcare distributed to each state is based on enrollment numbers [11]. Since the implementation of Seguro Popular, use of health services by Mexican citizens of all ages has steadily increased [12], and the dramatic improvements in healthcare coverage and service delivery have been heralded by health watchdogs and governments throughout the world [13]. However, several limitations of healthcare access in Mexico still persist, especially in isolated rural areas where there are a lack of clinics, doctors, and relatively low patient–provider ratios [13–15]. Furthermore, vulnerable groups in Mexico such as women, the elderly, and low-income populations are still facing insurance gaps, and preventative care practices are not available for large portions of the population [16], as any services not deemed "essential" by the government are not covered under Seguro Popular [11].

Unfortunately, an improved health system does not necessarily lead to improved HL among consumers. In fact, a changing healthcare system could make navigating services more difficult for individuals [17], particularly those with low HL. For example, Gazmararian et al. [18] found that among those with access to universal healthcare in the U.S. (i.e., Medicare for the elderly), 23.5% of English-speakers and 34.2% of Spanish-speakers still had inadequate HL. Furthermore, citing statistics from a Kaiser Permanente survey, Levitt [19] discussed the implications of the Affordable Care Act (ACA) on newly enrolled Americans. Only 57% of the surveyed population understood the term "provider network" and only 53% could define "a deductible", demonstrating the challenges facing new consumers trying to find their way within a complex health bureaucracy. Since individuals with limited HL inherently struggle to obtain care and make informed health decisions, assessing the HL of healthcare recipients should be a priority whenever there is a major shift in the structure of a national healthcare system. Unfortunately, at the time of this literature review, to our knowledge only one study has assessed HL in Mexico after the full implementation of Seguro Popular. Verastegui, De La Garza, and Allende-Perez [20] surveyed adults in an outpatient Mexican cancer clinic and found low HL in 15.4% of their sample, suggesting the vast majority had at least adequate HL. While their study provides important information for a unique subset of the population (adult patients with cancer), HL research is nonexistent among the majority of the Mexican population, including groups with historically low HL, such as youth [21].

An understanding of HL among youth is important because adolescence is a key development stage where children foster skills for adulthood [22]. Research suggests that low HL in youth—which can lead to the misinterpretation or misuse of health information—is correlated with adolescent obesity [23], less positive health habits [17], and an increase in risky behaviors, such as unsafe sex [24]. In developing nations, it is particularly important for youth to be able to discern between correct and erroneous health information as they typically have less access to the internet, fewer health professionals in their communities, and poorer health outcomes compared to youth in more developed regions of the world [25]. Although a few studies on adolescent HL have been conducted outside the US, the majority of HL studies among adolescents have focused on children in the U.S. [26].

One of the few studies that collected data outside of the U.S. was conducted by Hoffman and Marsiglia [25]. Using a sample of 230 adolescents, they explored the link between HL and substance use among youth in Guatemala. They found that those who went through a substance abuse prevention program scored higher on HL assessments than those who did not. Their results suggest that in more rural, developing parts of the world, utilizing existing health promotion programs and resources could be a viable, and economically beneficial, way to target and improve HL. Another global study offered a preliminary look at adolescent HL in Taiwan [27]. In her sample of over 1600 Taiwanese high school students, Chang found that adolescents with lower HL also had lower health status and poor health behaviors. Studies such as these are important foundational efforts, and similar efforts—within the context of critical HL—are needed in countries such as Mexico where information about the impact of healthcare changes is crucial for identifying systematic strengths and shortcomings in their new system.

In response to the growing recognition of the value of studying adolescent HL, frameworks with attention to adolescents specifically have been developed [6,28]. One such framework was developed by Massey, Prelip, Calimlim, Quiter, and Glik [29], who conducted 12 focus groups with 137 publicly insured teenage youth in the U.S., and completed 36 key informant interviews with primary care physicians. The authors' goals were first, to clearly operationalize the growing definition of adolescent HL, which had lacked a clear and succinct definition in the literature [22], and second, to develop a framework that can be used to develop specific measures and scales of adolescent HL. They identified five prominent domains, each of which serves to operationalize an adolescent's knowledge, attitudes, and practices in a healthcare setting: (1) their ability to appropriately navigate the health system (e.g., knowing how to make an appointment); (2) their knowledge and implementation of their own rights and responsibilities; (3) their knowledge of the need for and the implementation of preventative care practices; (4) their demonstration of appropriate and accurate health information seeking practices; and (5) their ability to form positive and effective patient–provider relationships. Competence in each of these domains can help to define the extent to which an adolescent is "health-literate", as utilizing an operationalization of adolescent HL with multiple domains can better capture the varying definitions of HL [29]. With this expanded framework, policy makers and stakeholders may be able to better identify where gaps in services exist.

This model in particular serves as a useful lens for analyzing adolescent HL as it focuses on many aspects of critical HL, specifically on how adolescents actually interact with health systems. This is becoming more imperative as more and more adolescents gain access with the expansion of universal healthcare worldwide [29]. Critical HL differs from other aspects of HL (e.g., document, quantitative, and functional HL) in that it centers on the ability of an individual to apply their health knowledge in a manner that allows them to exert control over their health [30–32]. Chinn [31] defines three unique aspects to critical HL: information analysis and appraisal, social/structural aspects of health, and collective action. While HL addresses an individual's ability to obtain health knowledge through reading, listening, or searching out information, critical HL is application and action-focused, and thus is a more robust framework for assessing the HL of populations or groups as it points to the ability to actually use information to improve one's health [33]. Within Massey et al.'s [29] model for adolescent HL, there is an underlying focus on actions adolescents can take to gain control over their health; specifically, asking questions of their doctors, taking appropriate steps to access their healthcare, and making preventative health decisions, which may ultimately improve their health. This focus on action also incorporates the growing efficacy and control adolescents have over their healthcare [22]. Massey et al.'s model [29] may also serve as a guide for cross-cultural research because it focuses on actual health behaviors, which allows for more attention to variation across cultures, versus solely health knowledge of adolescents, which many times is measured through comprehension only [22] and, subsequently, with stakeholders in policy considerations aimed narrowly at increasing adolescent knowledge. Considering the changing landscape of healthcare in Mexico, use of Massey et al.'s framework, which specifically focuses on interaction with the health system, may be able to provide valuable insights concerning the assessment, promotion, and actions taken by adolescents with regard to HL. Therefore, the purpose of this study was to assess the fit of Massey et al.'s framework of critical HL among Mexican youth by exploring their experiences utilizing healthcare services at a community health center. In doing so, we recognize the inherent limitations of applying a US derived model to Mexican youth, namely that the US and Mexico have very different health systems and distinctive norms and expectations surrounding healthcare utilization. Rather than comparing findings across cultures, we focus on the extent to which Massey et al.'s model fits, which may provide a useful framework for understanding Mexican youth experiences in their own right. Expanding our understanding of adolescent HL and health-seeking behaviors in Mexico is a foundational step towards understanding the potential impact of a changing healthcare system on their ability to meet their personal health goals.

2. Materials and Methods

2.1. Study Context

Procedures for this study were approved by the governing Institutional Review Board (UTSA Approval Number: 14-241N). We worked closely with two school psychologists from a middle school in a rural town of Central Mexico to recruit youth to participate in a two-pronged study involving a short survey and a focus group about health. Prior to conducting the study, we trained the school psychologists remotely using Skype in all data collection procedures. All youth in the school were eligible for recruitment and were recruited by the psychologists. As migration (both legal and illegal) to the US is integral to the financial stability of the community, the disclosure of personal information to a binational research team was concerning to members of the community. Thus, the use of active and written consent was determined to be unethical and unnecessary by the research team. Alternatively, the school psychologists spoke with parents and youth to answer questions about the risks and benefits of study participation and to discuss confidentiality. Specifically, parents and youth were told that researchers from a university in the U.S. were interested in youth's health and health-related experiences. After discussing the study, any parents or youth could opt to not participate further.

2.2. Sample

Of the 116 potential student participants, four chose not to participate in either the demographic survey or the focus groups. These students dropped out of school during the recruitment process, which was reported as not uncommon among students in the area. As shown in Table 1, 112 students completed the demographic survey. Of the 112 students, 98 also participated in a focus group (61 males and 37 females).

Table 1. Descriptive statistics.

Variables	*N*	Mean	SD	Range
Gender (female = 48%)	112	0.52	0.5	0–1
Age	110	13.20	0.99	11–16
Ever taken a health calls (0 = no)	109	0.64	0.48	0–1
Year in school	91	8.15	0.93	7–10
Grade point average	110	3.22	0.67	0–4
Parent education level [a]	104	0.73	1.49	0–6

[a] 0 = less than high school; 1 = high school or equivalent; 2 = some college; 3 = associate's degree; 4 = bachelor's degree; 5 = master's degree; 6 = doctorate or professional degree.

2.3. Procedures

Eight focus group (five male and three female) took place during normal school hours. The school psychologists worked with teachers to remove students by grade and sex from classrooms to attend the focus groups. One psychologist was male and the other female and each moderated their respective same-gender focus groups. Each discussion lasted approximately 30–45 min. Focus groups were large, averaging just over 12 students. Youth were asked a variety of health-related questions, including where they received health information, their experiences at a local health clinic, and who they spoke to about their health (see Appendix A for a full list of discussion questions).

2.4. Data Analysis

Data from the focus groups were audio recorded, which were later translated from Spanish to English and transcribed by a bilingual member of the U.S. research team. Translations were then checked by a second bilingual member of the research team, an independent bilingual reviewer (not part of the research team), and a bilingual member of the data collection team in Mexico. The English language transcripts were entered into NVIVO, a qualitative software program (QRS International, Melbourne, Australia) [34]. We followed the steps outlined by Braun and Clark [35] in first familiarizing ourselves with the data, open coding the transcripts inductively to note salient and reoccurring patterns, and categorizing these patterns into collapsed and meaningful themes via a preliminary codebook. That is, the codebook was developed in order to operationalize themes which had arisen from the data in a manner of frequency, specificity of discussion (e.g., examples and elaboration), and emotionality. This was in alignment with Krueger and Casey's [36] best practice recommendations for analysis of focus group data. As is common in qualitative research (see Crabtree & Miller [37]; Padgett [38]), we were sensitized through our preliminary coding process to the overlap between initial themes and those outlined by Massey et al.'s [29] framework for studying adolescent HL (e.g., our theme pertaining to "barriers to positive relationships with doctors" with theirs of "patient–provider relationships"), and thus moved from an inductive to a deductive codebook in order to assess fit of our data with this useful framework and to outline findings in a salient context to healthcare literature. (For an example of this inductive to deductive template approach, see Linton & Rueda [39].) The fit of our preliminary codebook to Massey et al.'s framework was excellent to the extent that only the titles and order of themes were changed. An inter-rater reliability kappa was calculated on the final codebook with a fourth and independent researcher who coded the data in its entirety. We had excellent reliability between the initial coding and that of the independent researcher who was unfamiliar with the study or its aims ($\kappa = 0.93$). Throughout the analysis process, the rigor of the study was further enhanced by use of observer triangulation and member checking. That is, three separate researchers corroborated to analyze the data [38] and the final themes of the study were sent to the school psychologists in Mexico for their feedback. Their feedback resulted in some minor changes to wording within the results section to capture accurate translations stemming from local colloquialisms. We report the following themes from our data in alignment with Massey's [29] framework of adolescent HL: (1) navigating the health system, (2) patients' rights and responsibilities, (3) preventative care, (4) information seeking, and (5) patient–provider relationships.

3. Results

The purpose of our study was to explore the experiences of youth in a rural Central Mexican town concerning their health experiences including those at a local healthcare clinic. Here, we outline themes in accordance with Massey and colleagues' [29] framework for critical health literacy. Direct Spanish quotations are provided alongside examples. Students chose their own pseudonyms at the start of each focus group.

3.1. Navigating the Health System

Although students knew the basics of how their local health clinic worked, they expressed frustration in not being able to get appointments or appropriate medications due to systematic issues with the center. The center utilized a token system that determined the order in which patients were seen by the doctors. Many students felt that the token system was ineffective, particularly as it disallowed them from seeing a doctor despite the fact that they had visited the clinic. Roxy (age 15, female) gave a response indicative of the frustration of many other students, stating " ... at eight in the morning they give tokens. If you make it, good! If not, oh well (O que a las ocho de la mañana dan las fichas, si alcanzaste bien si no, ni modo)." Gordito (age 14, male) explained that they do not give out many tokens a day, "they give out very few tokens, thirteen or ten (Dan muy poquitas fichas, trece o diez)." As Natalia (age 12, female) explained, if you do not get one of those ten tokens you have to "come (again) tomorrow (Venga hasta mañana)." The frustration with the token system was exasperated by the limited hours of the health center. The hours of the pharmacy were also limited, closing at noon, and making it difficult for students to get their medicine without missing school or their caretakers missing work. Halo (age 13, male) elaborated, "or sometimes they have closed, but you ring the bell and they do not open (O hay veces que tienen cerrado y tocas el timbre y no te abren)." Furthermore, even if they were able to see a doctor and make it to the pharmacy, many students expressed that the center did not have the medication they needed available: "there are almost no medications (Que no hay casi medicamentos)" (Isabela, age 14, female).

3.2. Patients' Rights and Responsibilities

Patient responsibility was brought up by the youth in the context of asking clarifying questions of the doctor if information was not readily understood. Students expressed that many times they did not understand the medical language doctors used: "they talk about things we do not understand (O que hablan de cosas que nosotros no entendemos)" (Pancracio, age 15, male); "it's that you do not understand how they explain things (Es que no entiendes como te explica por sus palabras)"(Bart, age 13, male). Estrella (age 14, female) said that patients have to ask the doctors to "can you please tell us more clearly what we have, or whatever?' (Nos puede decir por favor más claro lo que tenemos o ¿equis coas?)." Axel (age 13, male) stated that when one goes to the doctor the patient has to "ask them [the doctor] what sickness you have and to please explain it. Why did I get sick? (Tú le vas decir qué enfermedad es esa y me la puede explicar por favor y porque esa enfermedad me dio)." The youth clearly understood that, as patients, they have the right to be spoken to at a level and in a way that they can understand. However, as will be discussed further in the theme of patient–provider relationships, barriers to communication with health professionals included fear, mistrust, and embarrassment, which seemed to stem mostly from discussion of sexual health. It may be that the youth would feel more comfortable asking these questions of their doctor if not related to their sexual health. Even so, although many students spoke about barriers to communication with their doctor, some of the same students still spoke to the importance of asking these clarifying questions.

3.3. Preventative Care

In response to being asked when they go to the doctor the participants' answers overwhelmingly supported a nonpreventative approach to healthcare, with many saying they only go to the doctor "as soon as you're feeling pain from the sickness (En cuánto te da el dolor de la enfermedad)" (Axel, age 13, male); "well, only when we have the flu (Pues solo cuando tenemos gripa)" (Aureli, age 16, male); and "when you feel bad (Cuando te sientes mal)" (Lena, age 14, female; Shakira, age 16, female). A few students demonstrated a somewhat more preventative mindset, stating that one should go to the doctor either "when you have some doubt (Cuando tienes alguna duda)" (Ron, age 13, male) or "when you start to show symptoms (Cuando empiezas a presenter síntomas)" (Panfilo, age 15, male). Only one student (Teo, age 13, male) stated that one should go to the doctor "every month, every two months, or at least every year (Cada mex, cada dos meses o máximo cada año)."

3.4. Information Seeking

Although the participants overwhelmingly indicated that they would go to the health center for health information, when asked who they talk to about their medical concerns only a few of the youth said they would seek out their doctor. The youth responded that they mainly go to their mom or extended family, including grandparents, aunts, uncles, and/or cousins. Other responses included their dad, siblings, teachers, or the internet.

3.5. Patient–Provider Relationship

Many participants described the health center's medical staff as rude and not attending to patients, specifically pointing to the perceived lack of professionalism of the staff. Bart (age 13, male) offered his own experience as an example: " . . . when a (female) doctor is there and you say, 'oh my, that doctor is going to be there', and you go and they see you and say, 'I'm busy' and she isn't busy and is actually on her phone and like, what? She's not busy? (Cuando está una doctora y dices 'chin, va a estar esa doctora', y vas y te atiende y dice 'estoy ocupada' y en realidad está en el celular y ah, ¿cómo? ¿No está ocupada?)". Gordito (age 14, male) sarcastically expressed his perspective of the doctor's attitudes: "Let's see if the doctor is not tired and wants to see you (A ver si el doctor no esta cansado y te quiere atender)." Rosita (female, age 13) mirrored this sentiment, expressing that she believed doctors at the health center "sometimes they do it with a bad attitude (A veces hasta con mala gana)." Aside from not attending to the patients, some students suggested the doctors were incompetent because they would not prescribe the correct medication, or because they would give shots incorrectly or in a way that hurt patients. Regardless of the accuracy of such statements, it is important to note how the doctors' professionalism and competence were perceived by youth in the community.

A follow-up question seeking additional information about the challenges youth faced while speaking with the doctors led to a broad range of response including being scared of the doctor, not trusting the doctor, and simply not wanting to talk to the doctor. Several youth directly stated they were afraid of the doctor: "It scares me (me da miedo)" (Mayte, age 12, female); "I am afraid of the doctors (A mi me dan miedo los doctores)" (Victoria, age 12, female). Others stated that they do not talk openly with the doctors "because I do not trust them (Porque no les tengo confianza)" (Komander, age 13). The distrust theme was especially pervasive among female students, who were uncomfortable with male doctors. The female moderator summarized a conversation among the girls in one focus group: "Ah, yeah. There is no trust because they are doctors and in the case of the women [female youth], they are always hoping for a female doctor (Ah, ya. Que no hay confianza porque son doctores y en el caso de las mueres siempres buscan como una doctora)." Kalel (age 14, male) brought up how being scared and uncomfortable at the health center affected his communication with the doctor, stating, "we are really nervous, so we do not understand much (Estamos tan nerviosos que no entendemos mucho)."

When asked about their level of comfort speaking with adults, including doctors, about their health, the most common response was that it was difficult because they were embarrassed: "I am embarrassed with doctors (Me da verguenza con los doctores)" (Natalia, age 12, female); " . . . we get embarrassed (Que nos da pena)" (Lena, age 14, female); "we even get embarrassed to answer [their questions] (Hasta nos da verguenza contester)" (Gordito, age 14, male). For both males and females, the embarrassment many times stemmed from discussions of sexual health. Axel (age 13, male) explained "for example, sometimes your intimate [private] part hurts and you do not want to tell the doctor so they don't check it (Por ejemplo. A veces te duele tu parte íntima y no le quiere decir al doctor para que no te la revise)," while Victoria (age 12, female) suggested that she was embarrassed because "they ask you . . . if you have had sexual relations (Igual te pregunta si has tenido relaciones sexuales)." Both male and female students expressed that having a doctor of the opposite sex contributed to their embarrassment. Aureli (age 16, male) commented "if a female doctor is chosen you get nervous (luego si te toca un . . . una doctora, te dan como nervios)," while Natalia (age 12, female) recounted her experience with a male doctor, "you feel a lot of embarrassment because, um, they tell you 'raise your blouse a little bit so I can check your stomach,' and you say like . . . ay . . . thinking 'oh, no, how embarrassing' (sientes como mucha pena porque, este, te dice: 'Alzate la blusa tantito para checar tu estómago', y te quedas como . . . ay . . . pensando 'ay, no, que verguenza.'"

4. Discussion

The stories and experiences of youth seeking health information and care are integral to appraising their critical HL skills and health-seeking behaviors. This type of research is imperative to discussions on reducing health disparities, particularly in rural towns where health outcomes are often worse than urban or suburban areas due to the lower socioeconomic status of residents, poor service availability, and more hazardous living conditions [40]. Although research has suggested that Mexicans from low socioeconomic backgrounds have benefited in some ways from the universal healthcare system (i.e., Seguro Popular [12]), the present study offers a qualitative exploration of youth perspectives in the context of a rural Mexican town and details the lived experiences of struggle with the local health clinic provided under the Seguro Popular healthcare model. Theoretically, these results offer support for using Massey et al.'s [29] framework for analysis of adolescent critical HL in Mexico. On the one hand, the youth indicated they were able to navigate the system and understood the basic rights of the patient to be understand and applied these rights by asking clarifying questions of their doctors. However, they also spoke to a lack of basic preventative care practices, information seeking outside of the family, barriers to a productive relationship with healthcare professionals, and preferred family and other nonformal sources when they sought health information. An understanding of the strengths and limitations of the youth's critical HL can guide interventions to increase competence in each of the framework's five domains.

Participants in this study mentioned several barriers to developing a positive relationship with doctors including embarrassment, fear, mistrust, and doctor's use of medical language. These are barriers to critical health literacy as outlined by Massey et al.'s [29] model and common concerns for adolescents across cultures. Two recent qualitative studies in the US found that the use of non-age appropriate medical language was one of the biggest barriers in effective communication between adolescent cancer patients and their doctor [41,42]. Embarrassment is also a common barrier to healthcare utilization by adolescents [43], and one of the main reasons youth turn to the internet to research health problems rather than making an appointment with their doctor [44]. As the youth in the present study suggest, sexual health specifically seems to lead to feelings of embarrassment and hesitation to speak to a healthcare professional [45]. Feedback from member checking with the Mexican psychologists reiterated sexual topics as particularly difficult for adolescents to broach with healthcare providers and their families, despite the developmental importance of these to youth. While fear and mistrust do tend to impact the patient-doctor relationship and healthcare utilization among youth across cultures [45–47], these two themes are particularly prevalent within samples of Hispanic adolescents [47,48]. For example, in a sample of immigrant Latino adolescents in the U.S., qualitative results suggested that mistrust of doctors stemmed from fear of deportation, cultural misunderstandings, and negative experiences with other authority figures, such as teachers [49].

Although youth in the sample overall understood how to utilize the health clinic in their community, there was less understanding of how to navigate the barriers inherent in the system, which is a key element of an adolescent competent in the service navigation domain of Massey et al.'s [29] model. Even though their understanding of how to navigate the barriers inherent in the healthcare system may be somewhat limited, their ability to identify system limitations is a strength that should not be overlooked. This skill is certainly more aligned with critical HL than functional HL, although it is not directly accounted for in Massey et al.'s model [29]. The ability to identify limitations coupled with the apparent inability to navigate barriers suggests that adolescence may be the ideal time to promote critical HL skills. While it may be tempting to think that functional HL skills should be developed at a young age followed by the development of critical HL skills in early adulthood, it may be that critical HL skills would be most beneficial if taught simultaneously alongside functional HL at a young age.

The systematic barriers of the healthcare clinic—specifically the poor customer service, unprofessionalism of clinical staff, and the token system—mirror many of the same issues that were targeted by the Mexican Ministry of Health prior to implementation of Seguro Popular [50]. In a 2001 report based on samples of Mexican healthcare consumers, the Mexican Ministry of Health touted low emergency room wait times, high satisfaction with care, high levels of staff kindness, and positive experiences scheduling appointments. However, only 25% of the hospitals and clinics surveyed were public entities, suggesting that the findings were, in essence, a report on the private healthcare sector. The public health system, which has dramatically grown since the 2001 report due to the implementation of Seguro Popular, is an altogether different entity, and the experiences of youth in this study seem to suggest that healthcare utilization issues may be of concern in rural Mexico.

Preventative care practices that involved seeing a doctor were very rare among youth in our sample. In fact, only one person mentioned going to the doctor when they were not sick. Low use of preventive care practices in Mexico is not unique to youth in this study. An Organisation for Economic Cooperation and Development [13] report cites the lack of preventative care as a major issue facing the current Mexican healthcare system, especially in rural parts of Mexico. One study, analyzing data from a public Mexican health survey, found that over 30% of patients requiring inpatient medical services did not have access to their needed health service [15]. As rural sates in Mexico have some of the lowest patient–provider, patient–nurse, and patient–paraprofessional ratios in the country [14], there is limited access to primary care providers in these rural towns, which makes preventative care visits difficult [13]. Decreasing barriers to access and increasing knowledge of the importance of preventative care practices is key to increasing adolescent HL [29].

Aside from issues around access, expectancy theory can also shed light on how the barriers to positive patient–provider relationships as the systemic issues present at the clinic may relate to a lack of preventative care practices by adolescents. Vroom's [51] expectancy theory postulates that the motivations driving behaviors are linked to expected values and outcomes. Academic healthcare literature commonly uses expectancy theory to interpret patient expectations, perceived service quality, and satisfaction in both adult and adolescent populations [52–56]. Comments throughout our focus groups suggested that the local healthcare clinic has a bad reputation, and youth have had various negative interactions with the doctors and nurses. Expectancy theory suggests that this may lead to a diminished motivation to utilize the local healthcare clinic, with youth perhaps more likely to turn to less reliable sources for information with their healthcare questions and concerns (i.e., family members and the internet).

There are clear systemic changes that could improve the experiences of patients at the health clinic and increase service navigation and utilization. Changes to the token system (e.g., implementing an appointment-based system) could decrease long wait times. Changes to the operating hours of the clinic would better meet the schedules of adolescent patients in school (i.e., extending pharmacy hours beyond 12:00 pm). Frenk, Gomez-Dantes, and Knaul [57] identified long wait times at both outpatient clinics and hospitals, and lack of care hours during evenings and weekends as some of the major challenges facing the Mexican healthcare system after the implementation of Seguro Popular. The qualitative results of this study support the need for these areas, within the service navigation domain of Massey et al.'s [29] model, to continue to be explored by healthcare professionals and lawmakers in Mexico, particularly among rural healthcare providers.

On a community level, there are a variety of changes that could reduce the negative expectations of adolescents towards the local healthcare clinic, increase critical HL, and improve the utilization of health services. According to the short demographic survey completed by study participants, 64% of the students had attended a health class. To target the population who had not yet attended a health class, policy changes could require health education classes for all students in primary school, similar to what many states have already done in the U.S. [58]. Within the health class curriculum, a specific section or focus on patients' rights could also increase competence in the rights and responsibility domain of adolescent HL [29]. Furthermore, to decrease negative views of doctors, local physicians from the health center could be featured as guest speakers during the required health classes. This would give them an opportunity to provide an overview of the medical services offered through the local clinic and answer student questions. Various studies suggest that increased familiarity with doctors can enhance the patient–provider relationship and provide more positive healthcare experiences for youth [59–61]. Within the context of Massey et al.'s [29] model, interventions to increase positive patient–providers relationships are key in increasing adolescent HL.

Finally, a large number of the youth stated that they turn to a parent or family member for health information. Although this certainly speaks to the importance of immediate and extended family in the Mexican culture [62], it is imperative to consider the quality of the advice they are receiving, especially as appropriate information seeking and the ability to critically examine health information is an essential skill of adolescent HL [29]. Due to the increased availability of the internet, it is critical that youth and parents know about legitimate online resources and websites that can provide accurate information [63]. Furthermore, much like other school subjects such as math [64], if health information from presentations at school are not reinforced at home, it is unlikely to be retained by youth. Incorporating parents into the health education of adolescents is considered a promising model [65], so parents should be offered resources from school or provided the opportunity to attend health classes similar to what their children are receiving. This would reinforce student learning, help parents become more health literate, and increase both students' and parents' ability to critically analyze the health information they are receiving [29]. Further, incorporating family and the community was deemed important in feedback from our Mexican research psychologists who communicated that school is not valued by many students who often did not attend class or who dropped out altogether.

Returning to the Massey's [29] model for adolescent HL, our results suggest that this could act as a cross-cultural framework to explore the critical HL/health behaviors of adolescents, as the original thematic analysis fit within Massey's framework. There were, however, clear cultural differences of themes within each dominion between our focus groups in Mexico and the American context of Massey et al.'s model [29]. For example, while the youth in this study spoke about issues with the token system, adolescents in Massey's interviews spoke to the difficulty in getting an appointment time, which show the differences between the Mexican and American healthcare system. Perhaps the biggest cultural difference was the inclusion of parents, and especially extended family, in healthcare. While the American youth in Massey's interviews viewed parents as gatekeepers to their healthcare and did not bring up the role of extended family [29], the Mexican youth spoke specifically to the role of the extended family in their healthcare. Therefore, the model could serve by including a domain on the role of social networks (including peers, family, extended family, and extended kin), especially in international research on adolescent HL, to gain a fuller understanding of the role family has in adolescent healthcare. Overall, though the domains fit well with the themes that emerged, they also provided a framework to organize the aspects of critical HL of adolescents in rural Mexico, which can guide future policies and interventions.

5. Conclusions

A qualitative methodological approach has inherent limitations, including the limited transferability of findings and inability to provide statistical support for policy and practice recommendations. Furthermore, this particular study had a relatively small sample size and rather large focus groups. Although smaller groups may have afforded more opportunity for individual discussion and cohesion, Toner [66] discusses the importance of conducting groups of various sizes with understudied populations. In doing so, this project provides key insights about the critical HL skills of an understudied sample of youth in Central Mexico, and expands on and supports a current framework for studying adolescent HL [29]. Moreover, this study offers a unique understanding into the negative perceptions of a local healthcare clinic among youth—perceptions that are influential in shaping their future utilization of proper medical services and trust in health professionals. Whether these negative perceptions and experiences are common among youth throughout Mexico and are part of a broader systemwide challenge relating to the implementation of Seguro Popular is a question that deserves further attention. Future research should use mixed-methods approaches to gather information on the lived experiences of Mexican youth as they interact with their local health clinics, as well as provide quantitative data with larger and representative samples to support qualitative findings. Such studies should incorporate culturally appropriate assessments of functional and critical HL so meaningful interventions aimed at reducing health disparities can be created. In brief, the positive changes taking place within the Mexican healthcare system should be accompanied by efforts to ensure at-risk populations—such as youth—are able to access, understand, and use the new resources at their disposal.

Author Contributions: Conceptualization, S.H.; Formal analysis, H.R. and L.B.; Methodology, H.R.; Project Administration, S.H. and H.R.; Supervision, S.H.; Writing—Original Draft, L.B.; Writing—Review & Editing, S.H., H.R., and L.B.

Funding: This research received no external funding

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

- Where do you go for health information?
    - [For those who said "parents"] If you had a question about your health that your parents could not answer, where would you go for health information?
- Has it ever been difficult to obtain health information? If so, what types of challenges or barriers have you faced when seeking health information?
- In going to the doctor and speaking with them about your health, what challenges have you faced?
- In your opinion, when is it important to see a doctor or health professional?
- How easy or difficult is it to understand the health information you are given when you go to the doctor?
- How much do family members (e.g., mom, dad, siblings, aunt/uncle) assist you in setting up appointments, visiting with the doctor, and understanding the health instructions you are given?
- Have you ever been taught how to take care of your own health needs? If so, who taught you and what did you speak about?
- If you were to have a health problem that required medical attention would you know how to independently obtain the medical services you need without assistance from family or friends?
- What are your overall impressions of health services and healthcare in your community?

References

1. Sorensen, K.; Van den Broucke, S.; Fullam, J.; Doyle, G.; Pelikan, J.; Slonska, Z.; Brand, H.; Consortium Health Literacy Project European. Health literacy and public health: A systematic review and integration of definitions and models. *BMC Public Health* **2012**, *12*, 80. [CrossRef]
2. Berkman, N.D.; Dewalt, D.A.; Pignone, M.P.; Sheridan, S.L.; Lohr, K.N.; Lux, L.; Sutton, S.F.; Swinson, T.; Bonito, A.J. Literacy and health outcomes. *Evid. Rep. Technol. Assess* **2004**, *19*, 1228–1239.
3. Heinrich, C. Health literacy: The sixth vital sign. *J. Am. Acad. Nurse Pract.* **2012**, *24*, 218–223. [CrossRef] [PubMed]

4. Berkman, N.D.; Sheridan, S.L.; Donahue, K.E.; Halpern, D.J.; Viera, A.; Crotty, K.; Holland, A.; Brasure, M.; Lohr, K.N.; Harden, E.; et al. Health literacy interventions and outcomes: An updated systematic review. *Evid. Rep. Technol. Assess* **2011**, *199*, 1–941.
5. Batterham, R.W.; Hawkins, M.; Collins, P.A.; Buchbinder, R.; Osborne, R.H. Health literacy: Applying current concepts to improve health services and reduce health inequalities. *Public Health* **2016**, *132*, 3–12. [CrossRef] [PubMed]
6. Santos, M.G.; Gorukanti, A.L.; Jurkunas, L.M.; Handley, M.A. The Health Literacy of U.S. Immigrant Adolescents: A Neglected Research Priority in a Changing World. *Int. J. Environ. Res. Public Health* **2018**, *15*, 2108. [CrossRef] [PubMed]
7. Global Health Equity. Available online: https://www.amsa.org/about/mission-aspirations/global-health-equity/ (accessed on 21 December 2018).
8. Garni, A. Mechanisms of migration: Poverty and social instability in the postwar expansion of Central American migration to the United States. *J. Immigr. Refug. Stud.* **2010**, *8*, 316–338. [CrossRef]
9. National Action Plan to Improve Health Literacy. Available online: https://health.gov/communication/initiatives/health-literacy-action-plan.aspb (accessed on 21 December 2018).
10. Fact File on Health Inequities. Available online: http://www.who.int/sdhconference/background/news/facts/en/ (accessed on 21 December 2018).
11. Knaul, F.M.; Gonzalez-Pier, E.; Gomez-Dantes, O.; Garcia-Junco, D.; Arreola-Ornelas, H.; Barraza-Llorens, M.; Sandoval, R.; Caballero, F.; Hernandez-Avila, M.; Juan, M.; et al. The quest for universal health coverage: Achieving social protection for all in Mexico. *Lancet* **2012**, *380*, 1259–1279. [CrossRef]
12. Galarraga, O.; Sosa-Rubi, S.G.; Salinas-Rodriguez, A.; Sesma-Vazquez, S. Health insurance for the poor: Impact on catastrophic and out-of-pocket health expenditures in Mexico. *Eur. J. Health Econ.* **2010**, *11*, 437–447. [CrossRef] [PubMed]
13. OECD. *OECD Reviews of Health Systems: Mexico 2016*; OECD Publishing: Paris, France, 2016.
14. Alliance for Health Policy and Systems Research. *Primary Health Care Systems: Case Study from Mexico*; World Health Organization: Geneva, Switzerland, 2017.
15. Gutiérrez, J.P.; García-Saisó, S.; Dolci, G.F.; Ávila, M.H. Effective access to health care in Mexico. *BMC Health Serv. Res.* **2014**, *14*, 186. [CrossRef] [PubMed]
16. Urquieta-Salomon, J.E.; Villarreal, H.J. Evolution of health coverage in Mexico: Evidence of progress and challenges in the Mexican health system. *Health Policy Plan.* **2016**, *31*, 28–36. [CrossRef] [PubMed]
17. Dewalt, D.A.; Berkman, N.D.; Sheridan, S.; Lohr, K.N.; Pignone, M.P. Literacy and health outcomes: A systematic review of the literature. *J. Gen. Intern Med.* **2004**, *19*, 1228–1239. [CrossRef] [PubMed]
18. Gazmararian, J.A.; Baker, D.W.; Williams, M.V.; Parker, R.M.; Scott, T.L.; Green, D.C.; Fehrenbach, S.N.; Ren, J.; Koplan, J.P. Health literacy among Medicare enrollees in a managed care organization. *JAMA* **1999**, *281*, 545–551. [CrossRef] [PubMed]
19. Levitt, L. Why health insurance literacy matters. *JAMA* **2015**, *313*, 555–556. [CrossRef] [PubMed]
20. Verastegui, E.L.; De La Garza, J.E.; Allende-Perez, S.R. Health literacy in a comprehensive cancer center in Mexico. *J. Clin. Oncol.* **2018**, *36*. [CrossRef]
21. Sansom-Daly, U.M.; Lin, M.; Robertson, E.G.; Wakefield, C.E.; McGill, B.C.; Girgis, A.; Cohn, R.J. Health Literacy in Adolescents and Young Adults: An Updated Review. *J. Adolesc. Young Adult Oncol.* **2016**, *5*, 106–118. [CrossRef] [PubMed]
22. Manganello, J.A. Health literacy and adolescents: A framework and agenda for future research. *Health Educ. Res.* **2008**, *23*, 840–847. [CrossRef]
23. Chari, R.; Warsh, J.; Ketterer, T.; Hossain, J.; Sharif, I. Association between health literacy and child and adolescent obesity. *Patient Educ. Couns.* **2014**, *94*, 61–66. [CrossRef] [PubMed]
24. Swenson, R.R.; Rizzo, C.J.; Brown, L.K.; Vanable, P.A.; Carey, M.P.; Valois, R.F.; DiClemente, R.J.; Romer, D. HIV knowledge and its contribution to sexual health behaviors of low-income African American adolescents. *J. Natl. Med. Assoc.* **2010**, *102*, 1173–1182. [CrossRef]
25. Hoffman, S.; Marsiglia, F.F.; Nevarez, L.; Porta, M. Health literacy among youth in Guatemala City. *Soc. Work Public Health* **2018**, *23*, 30–36. [CrossRef] [PubMed]
26. Fleary, S.A.; Joseph, P.; Pappagianopoulos, J.E. Adolescent health literacy and health behaviors: A systematic review. *J. Adolesc.* **2018**, *62*, 116–127. [CrossRef] [PubMed]

27. Chang, L.C. Health literacy, self-reported status and health promoting behaviours for adolescents in Taiwan. *J. Clin. Nurs.* **2011**, *20*, 190–196. [CrossRef]
28. Koh, H.K.; Baur, C.; Brach, C.; Harris, L.M.; Rowden, J.N. Toward a systems approach to health literacy research. *J. Health Commun.* **2013**, *18*, 1–5. [CrossRef] [PubMed]
29. Massey, P.M.; Prelip, M.; Calimlim, B.M.; Quiter, E.S.; Glik, D.C. Contextualizing an expanded definition of health literacy among adolescents in the health care setting. *Health Educ. Res.* **2012**, *27*, 961–974. [CrossRef] [PubMed]
30. Nutbeam, D. The evolving concept of health literacy. *Soc. Sci. Med.* **2008**, *67*, 2072–2078. [CrossRef] [PubMed]
31. Chinn, D. Critical health literacy: A review and critical analysis. *Soc. Sci. Med.* **2011**, *73*, 60–67. [CrossRef]
32. Sykes, S.; Wills, J.; Rowlands, G.; Popple, K. Understanding critical health literacy: A concept analysis. *BMC Public Health* **2013**, *13*, 150. [CrossRef]
33. Guzys, D.; Kenny, A.; Dickson-Swift, V.; Threlkeld, G. A critical review of population health literacy assessment. *BMC Public Health* **2015**, *15*, 215. [CrossRef]
34. QSR International Pty Ltd. *NVIVO, Version 10*; QSR International Pty Ltd.: Melbourne, Australia, 2012.
35. Braun, V.; Clark, V. Using thematic analysis in psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101. [CrossRef]
36. Krueger, R.A.; Casey, M.A. *Focus Groups: A Practical Guide for Applied Research*, 5th ed.; Sage: Thousand Oaks, CA, USA, 2014.
37. Crabtree, B.F.; Miller, W.L. *Doing Qualitative Research*, 2nd ed.; Sage: Thousand Oaks, CA, USA, 1999.
38. Padgett, D.K. *Qualitative Methods in Social Work Research*; Sage: Thousand Oaks, CA, USA, 2008.
39. Linton, K.; Rueda, H.A. Dating and sexuality among minority adolescents with disabilities: An application of sociocultural theory. *J. Hum. Behav. Soc. Environ.* **2015**, *25*, 77–89. [CrossRef]
40. Smith, K.B.; Humphreys, J.S.; Wilson, M.G. Addressing the health disadvantage of rural populations: How does epidemiological evidence inform rural health policies and research? *Aust. J. Rural Health* **2008**, *16*, 56–66. [CrossRef] [PubMed]
41. Essig, S.; Steiner, C.; Kuehni, C.E.; Weber, H.; Kiss, A. Improving Communication in Adolescent Cancer Care: A Multiperspective Study. *Pediatr. Blood Cancer* **2016**, *63*, 1423–1430. [CrossRef]
42. Siembida, E.J.; Bellizzi, K.M. The Doctor-Patient Relationship in the Adolescent Cancer Setting: A Developmentally Focused Literature Review. *J. Adolesc. Young Adult Oncol.* **2015**, *4*, 108–117. [CrossRef] [PubMed]
43. Chelvakumar, G.; Ford, N.; Kapa, H.M.; Lange, H.L.H.; McRee, A.L.; Bonny, A.E. Healthcare Barriers and Utilization Among Adolescents and Young Adults Accessing Services for Homeless and Runaway Youth. *J. Community Health* **2017**, *42*, 437–443. [CrossRef]
44. Henderson, E.M.; Keogh, E.; Eccleston, C. Why go online when you have pain? A qualitative analysis of teenagers' use of the Internet for pain management advice. *Child Care Health Dev.* **2014**, *40*, 572–579. [CrossRef] [PubMed]
45. Hoopes, A.J.; Benson, S.K.; Howard, H.B.; Morrison, D.M.; Ko, L.K.; Shafii, T. Adolescent Perspectives on Patient-Provider Sexual Health Communication: A Qualitative Study. *J. Prim. Care Community Health* **2017**, *8*, 332–337. [CrossRef] [PubMed]
46. Harrison, M.E.; Clarkin, C.; Rohde, K.; Worth, K.; Fleming, N. Treat Me But Don't Judge Me: A Qualitative Examination of Health Care Experiences of Pregnant and Parenting Youth. *J. Pediatr. Adolesc. Gynecol.* **2017**, *30*, 209–214. [CrossRef] [PubMed]
47. Katz, I.T.; Bogart, L.M.; Fu, C.M.; Liu, Y.; Cox, J.E.; Samuels, R.C.; Chase, T.; Schubert, P.; Schuster, M.A. Barriers to HPV immunization among blacks and latinos: A qualitative analysis of caregivers, adolescents, and providers. *BMC Public Health* **2016**, *16*, 874. [CrossRef]
48. Raymond-Flesch, M.; Siemons, R.; Pourat, N.; Jacobs, K.; Brindis, C.D. "There is no help out there and if there is, it's really hard to find": A qualitative study of the health concerns and health care access of Latino "DREAMers". *J. Adolesc. Health* **2014**, *55*, 323–328. [CrossRef] [PubMed]
49. Lightfoot, A.F.; Thatcher, K.; Siman, F.M.; Eng, E.; Merino, Y.; Thomas, T.; Coyne-Beasley, T.; Chapman, M.V. "What I wish my doctor knew about my life": Using photovoice with immigrant Latino adolescents to explore barrier to healthcare. *Qual. Soc. Work* **2017**, 1–21. [CrossRef]
50. Ruelas, E. Health Care Quality Improvement in Mexico: Challenges, Opportunities, and Progress. In *Baylor University Medical Center Proceedings*; Taylor & Francis: Boca Raton, FL, USA, 2002; Volume 15, pp. 319–322.

51. Vroom, V.H. *Work and Motivation*; Wiley: New York, NY, USA, 1964.
52. Chisolm, D.J.; Manganello, J.A.; Kelleher, K.J.; Marshal, M.P. Health literacy, alcohol expectancies, and alcohol use behaviors in teens. *Patient Educ. Couns.* **2014**, *97*, 291–296. [CrossRef] [PubMed]
53. Conway, T.; Willcocks, S. The role of expectations in the perception of health care quality: Developing a conceptual model. *Int. J. Health Care Qual. Assur. Inc. Leadersh Health Serv.* **1997**, *10*, 131–140. [CrossRef] [PubMed]
54. Jang, S.H.; Kim, R.H.; Lee, C.W. Effect of u-healthcare service quality on usage intention in a healthcare service. *Technol. Forecast. Soc. Chang.* **2016**, *113*, 396–403. [CrossRef]
55. Lim, P.C.; Tang, N.K. A study of patients' expectations and satisfaction in Singapore hospitals. *Int. J. Health Care Qual. Assur. Inc. Leadersh Health Serv.* **2000**, *13*, 290–299. [PubMed]
56. Naidu, A. Factors affecting patient satisfaction and healthcare quality. *Int. J. Health Care Qual. Assur.* **2009**, *22*, 366–381. [CrossRef] [PubMed]
57. Frenk, J.; Gomez-Dantes, O.; Knaul, F.M. The democratization of health in Mexico: Financial innovations for universal coverage. *Bull. World Health Organ.* **2009**, *87*, 542–548. [PubMed]
58. Health Education. Available online: http://www.nasbe.org/healthy_schools/hs/bytopics.php?topicid=1100&catExpand=acdnbtm_catA (accessed on 21 December 2018).
59. Sleath, B.; Carpenter, D.M.; Coyne, I.; Davis, S.A.; Hayes Watson, C.; Loughlin, C.E.; Garcia, N.; Reuland, D.S.; Tudor, G.E. Provider use of a participatory decision-making style with youth and caregivers and satisfaction with pediatric asthma visits. *Patient Relat. Outcome Meas.* **2018**, *9*, 147–154. [CrossRef] [PubMed]
60. Sing, M. Communication as a bridge to build a sound doctor-patient/parent relationship. *Indian J. Pediatr.* **2016**, *83*, 33–37. [CrossRef] [PubMed]
61. Tates, K.; Meeuwesen, L. Doctor-parent-child communication. A (re)view of the literature. *Soc. Sci. Med.* **2001**, *52*, 839–851. [CrossRef]
62. Coltrane, S.; Gutierrez, E.; Parke, R.D. Stepfathers in Cultural Context: Mexican American Families in the United States. In *The International Handbook of Stepfamilies: Policy and Practice in Legal, Research, and Clinical Environments*; Pryor, J., Ed.; John Wiley & Sons: Hoboken, NJ, USA, 2008; pp. 100–121.
63. Jain, A.V.; Bickham, D. Adolescent health literacy and the Internet: Challenges and opportunities. *Curr. Opin. Pediatr.* **2014**, *26*, 435–439. [CrossRef] [PubMed]
64. Berkowski, T.; Schaeffer, M.W.; Maloney, E.A.; Peterson, L.; Gregor, C.; Levine, S.C.; Beilock, S.L. Math at home adds up to achievement in school. *Science* **2015**, *350*, 196–198.
65. Griebler, U.; Rojatz, D.; Simovska, V.; Forster, R. Effects of student participation in school health promotion: A systematic review. *Health Promot. Int.* **2017**, *32*, 195–206. [CrossRef] [PubMed]
66. Toner, J. Small is not too small: Reflections concerning the validity of very small focus groups (VSFGs). *Qual. Soc. Work* **2009**, *8*, 179–192. [CrossRef]

International Journal of
Environmental Research and Public Health

MDPI

Article

Pathologic Use of Video Games and Motivation: Can the Gaming Motivation Scale (GAMS) Predict Depression and Trait Anxiety?

Sara Peracchia [1], Fabio Presaghi [2] and Giuseppe Curcio [3,*]

1. Dipartimento di Medicina Clinica, Sanità Pubblica, Scienze della Vita e dell'Ambiente, Università degli Studi dell'Aquila, 67100 L'Aquila, Italy; sara.peracchia@gmail.com
2. Dipartimento di Psicologia dei Processi di Sviluppo e Socializzazione, Sapienza Università di Roma, 00185 Roma, Italy; fabio.presaghi@uniroma1.it
3. Dipartimento di Scienze Cliniche Applicate e Biotecnologiche, Università degli Studi dell'Aquila, 67100 L'Aquila, Italy

* Correspondence: giuseppe.curcio@univaq.it; Tel.: +39-0862-433407

Received: 29 January 2019; Accepted: 17 March 2019; Published: 20 March 2019

Abstract: Videogaming is an increasingly prevalent activity among adolescents worldwide. The present study aimed at adapting the Gaming Motivation Scale (GAMS) to the Italian context, assessing its psychometric properties and verifying its sensitivity to predict depression and anxiety levels. From a sample of 1899 participants, a group of 388 adolescents who participated in the survey was divided into two subgroups of Heavy (HG, N = 188) and Light Gamers (LG, N = 200). A sub-sample of N = 172 adolescents also filled-in CESD and STAI to assess, respectively, depression and trait anxiety. Internal consistency and factorial structure of the Italian version of GAMS (GAMS-it) have been evaluated. Moreover, a latent regression structural equation model by predicting the CES-D and STAI scores with the GAMS-it factors has been carried out. GAMS-it has adequate validity and reliability levels, showing a very similar factorial structure to the original version. Therefore, this scale can be used to evaluate gaming motivation, which is useful for gaming motivation screening. Finally, it has been found that lower gaming motivation can be related to high level of depression and anxiety. The present findings provide a coherent picture, supporting the reliability and validity of the GAMS-it, that appears potentially useful in predicting anxiety and depression levels in a population of adolescents.

Keywords: motivation; videogaming; adolescents; adolescents; psychopathology

1. Introduction

In the last decade, a growing amount of literature has been proposed on the new phenomenon of excessive use of video games (VG), and one of the focal points of scientific debate is to clarify if videogaming can have the potential to modify users' thoughts, feelings and behaviour. Looking at the results of previous investigations, conflicting opinions about the effects of exposure to VG did emerge: Some studies show the positive effects on some specific skills, as for example visual performance [1], visual acuity [2], probabilistic learning [3], stimulus-response mapping [4], encoding speed [5], and cognitive functioning [6]. On the other side, also some negative effects have been documented; as increasing aggression [7]; emergence of attention problems [8] or hyperactivity and mood troubles as depression and anxiety [9]; reduction of empathy [10]; impairment of social behaviour [11]; reduction of sleep time, quality and efficiency [12]; and possible addiction [13].

With respect to the issue of "addiction" to VG, researchers and clinicians have found considerable difficulties to assess, identify and define this phenomenon probably because VG does not involve a chemical substance and because problems induced by heavy use of these technological devices tend to be seen as "benign" since videogaming is less likely to pose social threats through illegal activities, as compared to drug addiction conditions [14]. However, research indicated that those who report excessive videogaming tend to show some addiction-like symptoms, including impairment in normal, social and occupational or educational functioning, tolerance, withdrawal and relapse [15], which may be considered "pathological" enough to require clinical attention and intervention. Only recently, the American Psychological Association included Internet Gaming Disorder (IGD) in the appendices of the DSM-5 [16], without specifically mentioning addiction to videogaming. Moreover, more recently, World Health Organization included the Gaming Disorder in the ICD-11 [17]. All these observations highlight the great relevance of these issues under both a scientific and health point of view.

Noteworthy, most studies payed attention only to the adult population, while this phenomenon is ever more relevant in adolescence. Converging evidences from different sources, in fact, indicate that adolescents' use of VG is sharply increasing in recent times and that gambling is becoming a common activity among young people, particularly among boys [18], which is usually associated also to behavioral problems as for example, substance abuse [19]. It needs, in fact, to take into consideration that adolescence is a period of great risk for developing addictive behaviors: Following Johnston and colleagues [20], we know that before the age of 18 years, 60% of individuals start to use drugs and 80% to drink alcohol. A similar trend has also reported for cigarettes smoking (e.g., Reference [14]). Therefore, we cannot exclude the possibility that a long and continuous exposure to VG can become a real and structured addiction.

In the last decade, also in Italy, an increasing amount of young people has been found excessively involved in videogaming, an activity that might increase the possibility to develop addiction and/or gambling. To date, in Italy there are only a few instruments able to measure adolescent motivation to gaming, and a rapid and easy screening scale would be very useful. Such a kind of instrument could help in assessing motivation and involvement in videogaming in order to identify individuals "at risk", and this could, in turn, help to prevent addiction and gambling [14].

In line with Self-Determination Theory (STD; [21]) it has recently developed the Gaming Motivation Scale (GAMS; [22]), designed to assess different aspects of gaming motivation: intrinsic motivation, integrated regulation, identified regulation, introjected regulation, external regulation, and amotivation. Motivation is one of principal components related to addiction e.g., Reference [23] and STD offers a multidimensional conceptualization of motivation that allows the assessment of level and type of motivation that can be applied to several domains including videogaming. Generally speaking, gaming motivation is a complex issue that includes different components: To the ones listed before and identified in the GAMS adaptation study [22] such as intrinsic and extrinsic motives for acting, we could include some others. For example, a sub-theory of SDT, called cognitive evaluation theory (CET [21]), is specifically focused on contextual factors that play a key role in intrinsic motivation. Such a theory hypothesizes that events and conditions able to boost an individual's sense of autonomy and competence can sustain intrinsic motivation, whereas other factors that reduce these aspects tend to weaken this specific type of motivation; see Reference [24]. However, also other factors are associated with intrinsic motivation, such as presence (intended as the sense to be within the game world), intuitive control abilities (crucial to assess need satisfaction in game play), enhancement of individual well-being, and relatedness (when one feels connected with others). On the other hand, an increase of extrinsic motivation with its pivotal aspects of rewards, pressures and evaluations usually diminishes the weight of an intrinsic motive [25]. This complex interplay between different motivations and the several factors affecting them, can reasonably explain different aspects of gaming such as the need for satisfaction in games and short-term well-being; the appeal of violent game content; as well as motivational sources of post-play aggression; the antecedents and consequences of disturbed patterns of game engagement; or the causes and effects of games' aspects as immersion, flow and presence, for a comprehensive review see Reference [25].

A companion issue to gaming is the psychological profile of gamers. As mentioned before, a relationship does exist between gaming abuse and some psychopathological traits such as anxiety and depression [9]. Some recent studies highlighted a reciprocal relationship between pathological gaming and indices of mental health troubles [26], indicating a kind of comorbidity of these traits with maladaptive gaming attitudes. Also, IGD was found to be strongly correlated with such psychopathological traits [27] while other authors hypothesized gaming as a possible maladaptive coping strategy to deal with negative affective disturbances [28].

The aim of the present study is to (a) adapt the Gaming Motivation Scale (GAMS; [22]) to the Italian adolescents' context, (b) examine its psychometric properties in a sample of young Italian students, and (c) verify if GAMS-it factors might predict level of depression and trait anxiety. With respect to the last aim, we hypothesize that high gaming motivation positively correlates with gaming behavior and with psychopathological factors of anxiety and depression.

2. Materials and Methods

2.1. Participants

Different schools of Center-South Italy (I.I.S. "A. Bafile", L'Aquila; I.S.I.S. "E. Mattei" Cerveteri [RM]; Liceo Classico "J. da Todi" Todi [PG]; Licei "T. Campanella", Belvedere M.mo [CS]) have been invited to participate in the study, and a total of 1899 adolescents agreed to participate in the survey (970 males and 929 females; Mage = 15.13; SDage = 1.34 years; range 14–19 years). Together with the number of hours played, the whole sample was also asked to indicate the preferred type of VG. Participants chose different genres of VG (adventure, action, quiz, strategy, arcade, fighting game, role playing game, simulation, sports, educational, and first-person shooter). The different types of video games were divided into two main groups defined as "action video games" (AVG; games mainly based on action, incorporating adventure, action, strategy, fighting game, role playing game, simulation, sports, and first person shooter genres) and "no-action video games" (N-AVG; namely quiz, arcade, and educational kind of video games). Based on these data, it emerged that 40.65% of the total sample preferred AVG, compared with 4.58% that chose N-AVG.

Based on the amount of hours of videogaming, we selected 388 participants by the original sample who composed two subgroups of interest: 188 Heavy Gamers (HG), who reported to play for more than 4 h per week (145 males and 37 females; M-age = 15.12; SD-age = 1.42 years) and 200 Light Gamers (LG), who affirmed to play less than 1 hour per week (43 males and 163 females; M-age = 15.08 years; SD-age= 1.11 years). The analysis conducted on two subgroups highlights that for 70.59%, HG choose AVG, while LG prefer play more with N-AVG (61.34%). This sample of participants was considered for the confirmatory analysis of GAMS.

Finally, among the 388 participants, a sub-sample of 172 participants (defined as Heavy Gamers N = 95, and Light Gamers N = 77) accepted to fill-in the CESD (Center for Epidemiologic Studies Depression Scale) and the STAI (State-Trait Anxiety Inventory) questionnaires.

2.2. Instruments

In the present study, we administered to participants some different scales: (1) the Gaming Motivation Scale (GAMS) developed by Lafreniére and coworkers [22], (2) the Assessment of Internet and Computer Game Addiction Scale (AICA-S; [29]), (3) the Center for Epidemiologic Studies Depression Scale (CES-D; [30]), and 4) the State-Trait Anxiety Inventory (STAI; [31]).

The GAMS was generated on the basis of Deci & Ryan's conceptual definitions of motivations [21]. The purpose of this scale was to investigate the reasons at the base of playing VG; participants were asked to respond to 18 items rated on a 7-point Likert scale (from 1 "do not agree at all", to 7 "very strongly agree") and bearing in mind a same basic question, i.e., "Why do you play to video games?". The scale allows to assess different type of motivations: intrinsic motivation (desire to perform an activity for itself), integrated regulation (first aspect of extrinsic motivation, that refers to engaging in an activity out of choice), identified regulation (second aspect of extrinsic motivation, when people engage in a behavior based on its perceived meaning or its relation to personal goals), introjected regulation (third aspect of extrinsic motivation, that refers to the regulation of behavior through internal pressures such as anxiety and guilt, implying partial internalization), external regulation (fourth aspect of extrinsic motivation, that refers to behavior regulated through external means such as rewards), and amotivation (similar to learned helplessness, that refers to the relative absence of motivation either intrinsic or extrinsic) [22]. GAMS was translated into Italian by two experienced researchers; the translation was then evaluated by two independent experts in experimental psychology. Finally, one translator (an English native speaker) back-translated the questionnaire from Italian to English. After this procedure, the Italian version of the scale (GAMS-it) was obtained. For the general structure of the scale, we maintained the original one and thus the order of the items is the same as in the validation study (see Reference [22]).

The Assessment of Internet and Computer Game Addiction Scale (AICA-S; [29]) is a self-report scale for the assessment of potentially pathological computer and game behavior. Fifteen items are relevant for clinical classification of computer game use behaviour (e.g., craving, tolerance, loss of control, unsuccessful attempts to cut back, and withdrawal). Previous studies on its psychometric properties yielded satisfying results concerning item characteristics, reliability and validity [32]. For the present study, AICA-S was used for determining the amount of hours of gaming and to identify the participants defined as “Heavy Gamers” (HG) and “Light Gamers” (LG). In particular, we took into consideration the answer to the fourth question of this scale (“How you long are you playing computer games?”, with five response categories (1 = “less than one hour per week”, 2 = “1–2 h”, 3 = “2–4 h”, 4 = “4–6 h” and 5 = “more than 6 h to week”): participant who responded “1” were included in LG group, instead the participant that responded “4” and “5” were included in HG group. With respect to the meaning of question 1, “computer games” were intended as all the types of games both those on online platforms and on home console. The Italian version of AICA-S has been developed by following the same procedure followed with GAMS [33].

The 20-item Center for Epidemiologic Studies Depression (CES-D; [30]) is frequently used to estimate the prevalence of depressive symptomatology in the general population. Respondents rated the frequency with which they have experienced particular depressive symptoms during the past week. Answers to each item range from 0 (less than 1 day) to 3 (5–7 days) and are summed to compute a total score. It measures a single depression factor, ranging from 0 to 60; scores of 16 or above are considered potentially pathological.

The State-Trait Anxiety Inventory (STAI; [31]) is a commonly used measure of trait and/or state anxiety all over the world. It can be used in both clinical settings and research. The most popular version (Form Y) has 20 items for assessing trait anxiety and 20 for state anxiety: In the present study, we used the trait version. All items are rated on a 4-point scale (e.g., from “Almost Never” to “Almost Always”). Higher scores indicate greater anxiety level.

2.3. Procedure

Participants were recruited in their school: Only those who were authorized by parents and teachers participated in the study; no incentive was given for participation. Participants completed the “paper and pencil” questionnaire in their class during lessons’ breaks under the supervision of at least one researcher, who was also available if participants asked for clarifications.

The study is divided into two sections: The first one, preparatory to the latter, aimed at verifying the factorial structure of the GAMS-it and compared it with the original version; the second section was aimed at investigating its discriminant validity by correlating the six factors with psychopathological factors like those assessed by CES-D (depressive factors) and STAI (trait anxiety) and verifying if GAMS-it factors can predict the level of depression and trait anxiety.

The whole study protocol was conducted in accordance with the Declaration of Helsinki and approved by the Internal Review Board of the University of L’Aquila (#16/2016).

3. Results

3.1. Analysis Strategy

From a preliminary check for the normality of distribution, it resulted that all the 18 GAMS-it items reported a significant deviation from normality (Shapiro-Wilk test ranged from w = 0.45, $p < 0.01$, to w = 0.86, $p < 0.01$), so Robust Maximum Likelihood estimation method with robust standard error and robust fit statistics was considered for conducting the CFA (i.e., References [34,35]). The goodness of fit of the model was consequently evaluated by means of the Satorra-Bentler (SB) scaled chi-square statistic [34], as well as other well-known fit indices such as: Root Mean Square of Approximation (RMSEA; [36]), Comparative Fit Index (CFI; [37]) and Non-Normed Fit Index (NNFI; [38]). Following guidelines by Hu and coworkers (1992) [39], the model is considered to hold approximatively in the population if the RMSEA value is below 0.08 (the closer to 0.00 the better), and if both CFI and NNFI are above 0.95, it is indicative of a reasonable goodness of fit. The same fit indexes will be used to evaluate the fit for the latent factor regression model.

Given that high- and low-frequency gamers (namely, HG and LG) may differ in terms of motivation and regulatory strategies, it is consequently crucial to test whether the same factor structure (in terms of latent means and covariance structure) is invariant across the two sub-groups. Multi-sample CFA was then performed in order to investigate how well the factor model emerging from the previous analysis could be generalized across high- vs. low-frequency gamers.

A procedure for testing factorial invariance was followed (e.g., Reference [40]). The procedure consisted of a series of hierarchical statistical invariance tests (configural, metric, scalar, unique variance, latent variance and latent means), starting with the omnibus test of the equality of covariance matrices across groups. The scaled difference chi-square statistics, $\Delta SB\chi^2$, [35] were used for comparing the fit between two nested models, i.e., configural and metric invariant and to determine if the more restricted model has or not a non-worsen fit than the less restricted model. The null hypothesis of the statistical test was accepted when the estimated probability of the test was greater than 0.01.

3.2. Results of CFA of GAMS-it

On the selected sample of participants who comprised of low (LG, N = 200) and high (HG N = 188) frequency gamers, a Confirmatory Factor Analysis has been performed. In general, the Six-factor structure reported satisfactory fit statistics (SB-χ^2(120) = 256.35, $p < 0.01$) with an RMSEA of 0.070 (90% C.I: 0.058–0.082), a NNFI of 0.969, and CFI = 0.961. As shown in Table 1, all GAMS-it items reported a significant factor loading on the expected factor; moreover, as described in Table 2, all factors correlated significantly, and the expected direction and reliability was satisfactory (ranging from 0.76 to 0.93). It is worth noting, the correlation between Integrated regulation and Identified Regulation is extremely high ($\theta = 0.945$) and it is of the same entity of that reported by Lafreniére and colleagues [22]. We tested the hypothesis that the correlation between the two factors is equal to 1, meaning that the two factors are the same thing. The difference in model fit was significant (ΔSB-χ^2(1) = 12.23, $p < 0.01$) indicating that the model with the covariance fixed to 1 has a significantly worse fit. In conclusion, even if the correlation is high, the two constructs can be seen as different.

Table 1. Standardized factor loadings with Robust SE for the six factors of GAMS-it (N = 388).

GAMS Factors and Related Items	Factor Loading	SE
Intrinsic motivation		
GAMS.1: Because it is stimulating to play	0.914 **	0.013
GAMS.2: For the pleasure of trying/experiencing new game options (e.g., classes, characters, teams, races, equipment)	0.857 **	0.022
GAMS.3: For the feeling of efficacy I experience when I play	0.881 **	0.015
Integrated regulation		
GAMS.4: Because it is an extension of me	0.896 **	0.020
GAMS.5: Because it is an integral part of my life	0.909 **	0.017
GAMS.6: Because it is aligned with my personal values	0.895 **	0.017
Identified Regulation		
GAMS.7: Because it is a good way to develop important aspects of myself	0.812 **	0.026
GAMS.8: Because it is a good way to develop social and intellectual abilities that are useful to me	0.726 **	0.036
GAMS.9: Because it has personal significance to me	0.844 **	0.022
Introjected Regulation		
GAMS.10: Because I feel that I must play regularly	0.894 **	0.020
GAMS.11: Because I must play to feel good about myself	0.837 **	0.028
GAMS.12: Because otherwise I would feel bad about myself	0.705 **	0.035
External Regulation		
GAMS.13: To acquire powerful and rare items (e.g., armors, weapons) and virtual currency (e.g., gold pieces, gems) or to unlock hidden/restricted elements of the game (e.g., new characters, equipment, maps)	0.850 **	0.025
GAMS.14: For the prestige of being a good player	0.907 **	0.016
GAMS.15: To gain in-game awards and trophies or character/avatar's levels and experiences points	0.933 **	0.014
Amotivation		
GAMS.16: It is not clear anymore; I sometimes ask myself if it is good for me	0.924 **	0.032
GAMS.17: I used to have good reasons, but now I am asking myself if I should continue	0.761 **	0.037
GAMS.18: Honestly, I don't know; I have the impression that I'm wasting my time	0.523 **	0.047

** $p < 0.01$; N = 388.

Table 2. Descriptive Statistics, Cronbach' alphas and Correlations among the six GAMS-it factors (listwise missing deletion, N = 388).

	M	SD	α	Int-Mot	Int-Reg	Id-Reg	Introj-R	Ext-Reg	Amot
Int-Mot	3.18	2.03	0.91	1	0.823 **	0.867 **	0.780 **	0.853 **	0.268 **
Int-Reg	2.23	1.72	0.93		1	0.945 **	0.843 **	0.634 **	0.196 **
Id-Reg	2.22	1.56	0.84			1	0.816 **	0.710 **	0.205 **
Introj-R	1.90	1.45	0.85				1	0.706 **	0.248 **
Ext-Reg	2.86	2.14	0.92					1	0.254 **
Amot	2.51	1.70	0.76						1

Note: Int-Mot = Intrinsic motivation; Int-Reg = Integrated regulation; Id-Reg = Identified Regulation; Introj-R = Introjected Regulation; Ext-Reg = External Regulation; Amot = Amotivation [** $p < 0.01$].

3.3. Results of Multi-Sample CFA

Although the configural model had a significant SB-χ^2, the values of RMSEA were acceptable while those of NNFI and CFI not (SB-$\chi^2(240)$ = 743.33, $p < 0.001$, RMSEA = 0.05, NNFI = 0.86, CFI = 0.87). Under the assumption of strict metric invariance, the fit (SBχ^2) did not worsen significantly

(SB-χ^2(252) = 779.91, $p < 0.001$, RMSEA = 0.05, NNFI = 0.88, CFI = 0.88; $\Delta SB\chi^2$ (12) = 10.40, $p = 0.58$) so the same six factor model with the same loadings in the two groups is tenable. However when we add to the metric invariance also the invariance of intercepts, the model fits much worse (SB-χ^2(264) = 896.72, $p < 0.001$, RMSEA = 0.05, NNFI = 0.84, CFI = 0.84; $\Delta SB\chi^2$ (12) = 33.0, $p = 0.0009$). Given these results, we decided to do not proceed with the remaining hierarchical tests of measurement invariance, as the fit would surely worsen. In conclusion, the metric invariance for the two sub-groups was reached.

3.4. Differences between Heavy Gamers and Light Gamers

Finally, we compared the Heavy (HG) and Light Gamers (LG) groups with respect to the six GAMS-it factors and to the CESD and STAI scores (Table 3). HG scored significantly higher than LG on Intrinsic Motivation, Integrated Regulation, Identified Regulation, Introjected Regulation, and on External Regulation, but not on Amotivation. Concerning psycho-pathological measures, we found that HG reported significantly lower scores than LG on STAI but not on CESD (Table 3). Nonetheless, it should be stressed that such a difference could be a direct consequence of unbalanced gender composition between HG and LG.

Table 3. Descriptive statistics (M, SD) comparing Heavy Gamers (HG) and Light Gamers (LG) on each of the six GAMS-it factors and for CESD and STAI scores.

	HG			LG			F (1, 386)	*p*
	N	M	SD	N	M	SD		
Int-Mot	188	4.8	1.5	200	1.6	1.0	603.9	<0.001
Int-Reg	188	3.4	1.8	200	1.1	0.5	288.2	<0.001
Id-Reg	188	3.3	1.6	200	1.3	0.6	277.7	<0.001
Introj-R	188	2.8	1.6	200	1.1	0.4	202.0	<0.001
Ext-Reg	188	4.4	1.9	200	1.4	0.9	421.3	<0.001
Amot	188	2.6	1.6	200	2.4	1.7	1.2	0.283
CES-D	95	15.9	10.5	77	17.4	10.2	0.9 [a]	0.349
STAI	95	39.8	10.4	77	43.6	11.2	5.3 [a]	0.023

Note: Int-Mot = Intrinsic motivation; Int-Reg = Integrated regulation; Id-Reg = Identified Regulation; Introj-R = Introjected Regulation; Ext-Reg = External Regulation; Amot = Amotivation; CES-D = Center for Epidemiologic Studies Depression; STAI = State-Trait Anxiety Inventory. [a] Data for this combination of variables was available only on 172 participants. So d.f. for this ANOVA was (1, 170).

3.5. Predicting Depression and Trait Anxiety with GAMS-it Factors

As stated before, the main objective of the study was to investigate discriminant validity aspects of GAMS by correlating the six factors of GAMS with psychopathological factors like those assessed by CES-D (depression level) and by STAI (trait anxiety). We performed a latent regression structural equation model by predicting the CES-D and STAI scores with the six latent GAMS-it factors and covarying for the effect of sex (Male vs. Female) and type of gamer (LG vs. HG). We preferred this statistical approach because it is a common model for two outcome variables, also giving the possibility of obtaining latent correlations. As for CFA and for MG-CFA, we used Robust estimation methods.

Results of the fitted Structural Equation Model (as depicted in Figure 1) showed the following fit indices: SB-χ^2(180) = 484.94, $p < 0.01$, with a Robust RMSEA of 0.112 (90% C.I: 0.100–0.124), a NNFI of 0.857, and CFI = 0.888. Table 4 shows the estimated latent regression parameters of the six GAMS-it factors on both CES-D and STAI scores. Regarding CES-D scores, we found that only one of the six GAMS-it factors, Amotivation, significantly and positively predicts (b = 0.209, se = 0.077, $p = 0.007$) depression level. Females (b = 0.400, s.e. = 0.092, $p = 0.000$) as well as HG (b = 0.343, s.e. = 0.091, $p < 0.001$) reported a positive and significant effect. While concerning STAI scores, we found that not only Amotivation has a positive and significant effect (b = 0.381, se = 0.073, $p < 0.001$), but also Intrinsic motivation has a negative and significant effect (b = −0.640, se = 0.231, $p = 0.006$) on the trait anxiety measure. Also, in this case, females reported significantly higher levels of anxiety with respect to males (b = 0.208, s.e. = 0.091, p = 0.022); however, HG did not report anxiety scores significantly higher than that of LG (b = 0.119, s.e. = 0.091, p = 0.189).

Finally, it should be stated that the two clinical measures, CES-D and STAI, reported a significant and positive latent correlation (r = 0.717, se = 0.047, p < 0.001).

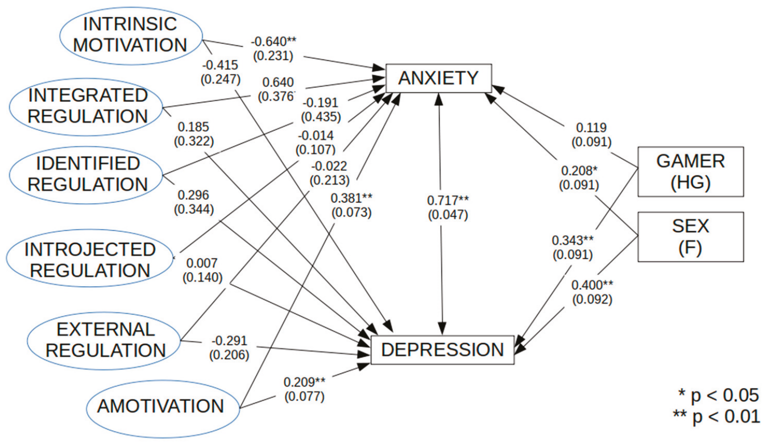


Figure 1. Path Model of the effects (completely standardized regression coefficients) of GAMS latent factors on Depression (CES-D scores) and Anxiety (STAI scores).

Table 4. Effects (Completely standardized regression coefficients) of GAMS latent factors on the CES-D and STAI scores (N = 172).

	CES-D Score			STAI Score		
	b	se	p	b	se	p
Int-Mot	−0.415	0.247	0.093	−0.640	0.231	0.006
Int-Reg	0.185	0.322	0.566	0.619	0.376	0.100
Id-Reg	0.296	0.344	0.389	−0.191	0.435	0.660
Introj-R	0.007	0.140	0.961	−0.014	0.107	0.893
Ext-Reg	−0.291	0.206	0.157	−0.022	0.213	0.919
Amot	0.209	0.077	0.007	0.381	0.073	0.000
Sex (F)	0.400	0.092	0.000	0.208	0.091	0.022
Gamer (HG)	0.343	0.091	0.000	0.119	0.091	0.189

Note: Int-Mot = Intrinsic motivation; Int-Reg = Integrated regulation; Id-Reg = Identified Regulation; Introj-R = Introjected Regulation; Ext-Reg = External Regulation; Amot = Amotivation.

4. Discussion

In the present study, we firstly adapted the Gaming Motivation Scale (GAMS; [22]) to the Italian adolescents' context, examining its psychometric properties in a sample of young Italian students and then assessed the ability of this instrument to predict the presence of possible psychopathological aspects in the young population.

To these aims, after having verified the factorial structure of the GAMS-it and compared it with the original version, we run a Confirmatory Factor Analysis on a subsample of 388 participants with low (LG) and high (HG) frequency of gaming. Results support the six-factor structure of GAMS-it reporting satisfactory statistics fit. As shown in Table 1, all GAMS-it items reported a significant factor loading on the expected factor and all factors correlated significantly and in the expected direction (Table 2); in addition, the reliability also resulted in being satisfactory. In addition, high correlation between Integrated Regulation and Identified Regulation (θ = 0.945) did emerge, as also reported in the original publication [22]. Moreover, the multi-sample CFA highlighted the equality between the various distributions of the correlation between the indicators and their dimensions, so the same six factor model with the same loadings in the two groups is tenable. The results confirmed the hypothesis of an invariant structure for the two groups. As a whole, such results showed adequate levels of validity and reliability of GAMS-it, as in the original validation study [22].

The comparison between two subgroups with respect to the six factors of GAMS-it and the result obtained from CESD and STAI showed a significantly higher score on Intrinsic motivation, Integrated regulation, Identified Regulation, Introjected Regulation, and External Regulation factor in HG, but not Amotivation. Moreover, HG showed a significantly lower score in the STAI, but not in the CESD (Table 3), indicating that a great amount of exposure to video games does not cause an increase of anxiety levels as usually believed.

Another interesting result is the one emerging from Latent Regression Structural Equation Model correlating CESD and STAI scores with latent six factor of GAMS-it, showing satisfactory fit (SB-$\chi^2(144) = 248.22$, $p < 0.01$; Table 3). Regarding CES-D scores, only the GAMS-it' Amotivation factor predicts significantly and positively the depression level (b = 0.247, se = 0.081, $p < 0.001$). This indicates that Amotivation levels can predict the depression level in a directly proportional way. Regarding STAI score, it emerged that not only Amotivation has a positive and significant effect (b = 0.390, se = 0.074, $p < 0.001$), but also that Intrinsic motivation has a negative and significant effect (b = −0.688, se = 0.238, $p = 0.004$). This indicates that Amotivation levels predict different levels of anxiety in a directly proportionate way; conversely, Intrinsic motivation levels predict indirectly the proportional anxiety levels. Finally, the two clinical measures, CES-D and STAI, reported a significant and positive latent correlation (r = 0.717, se = 0.047, $p < 0.001$). The present results only partially confirm the hypothesis of the study about a general correlation between motivation and anxiety/depression.

All these results allow to make some interesting conclusions. From SEM results, three conclusions can be drawn. As a first, we can confirm that the original GAMS factor structure can be generalized to a different culture (i.e., the Italian one). This means that factors underlying motivations to play video-games are the same in the two cultures. Therefore, this scale can be used to evaluate gaming motivation, providing additional statistic support to GAMS and introducing in the Italian context a new scale to be used as a potential screening instrument of gaming motivation. Nonetheless, further research is mandatory in order to establish and clarify psychometric property of the GAMS-it. Secondarily, multi-sample SEM confirmed that the GAMS factor structure is invariant in Heavy and Light gamers; this does mean that motivations to play video-games in both populations did not differ in terms of types of basic drive and that the differences between HG and LG are limited to a quantitative difference, along the same factor structure: Heavy gamers show higher average values with respect to Light gamers on all GAMS dimensions, but one (i.e., Amotivation). Finally, consequent to this, the Amotivation factor is the only one that significantly predicts psychopathological traits like depression and anxiety. So the present study, in conclusion, shows (and adds to the literature) that motivation of both Heavy and Light gamers is the same, that it differs somewhat on average between the two groups, and that these differences may be used to discriminate between them; of greater importance, our study shows that Amotivation is not critical for differentiating HG from LG, but is crucial as a precursor of depression and anxiety, in the sense that high levels of Amotivation predict high levels of depression and anxiety scores, namely, lower gaming motivation can be related to high level of depression and anxiety. These data are in line with current literature on depression and anxiety symptoms in adolescence, where anhedonia, apathy and loss of interest for all kind of activity can predict psychopathological outcomes [41]. An opposite effect has instead been reported for Intrinsic Motivation factor (limitedly to STAI scale), showing that high levels of Intrinsic Motivation can predict a reduced level of anxiety.

The present data contribute to explain the not ever consistent literature on the effects of exposure to videogames on psychopathological components as depression and anxiety. For example, the relationship between videogaming and psychopathological symptoms is still unclear; is psychopathology to induce the approach to VG (and thus possible excessive use and addiction) or does the exposure to VG tend to unmask a latent condition of depression and anxiety? In our opinion, the results obtained in the present study can help to clarify the problem. Indeed, if a lack of gaming motivation predicts high levels of depression and anxiety, we could suppose that persons with depression and anxiety will not be interested in videogaming, as they are unmotivated. Therefore, a state of anxiety and depression cannot be considered a factor able to encourage the use of video games. This assertion is also confirmed by negative relationship between high Intrinsic Motivation to video games use and anxiety. Healthy participants seek wellness in video games, because technology increases well-being-inducing vigor and resistance in players [42].

Furthermore, research based on Self Determination Theory (SDT) revealed that self-determined forms of motivation induce adaptive consequences as pleasure, persistence and wellness [43]. This statement could also be extended to pathological gambling: Thus, people generally anxious and/or depressed would be not encouraged to play videogames or gambling because of their reduced levels of motivation. In this line, anxiogenic and depressive symptoms of patients with Internet Gaming Disorder (IGD) may emerge after exposure to video games. However, the multifactorial aspects of videogaming do not allow us to consider these definitive conclusions, and more research is needed to clarify the effective relationship among the different factors involved.

As a limitation of the study, the sample size should be highlighted. Further validation work should test the GAMS-it on larger samples in order to obtain a greater statistical validity. For example, a possible future direction could be a full cross-sectional sample, including also a group of "medium level" players. Moreover, another limitation arises from the limited availability of participants to complete psychopathological questionnaires. Also, the two subgroups (HG and LG) considered were unbalanced with respect to gender composition; this did not allow us to investigate the issue of gender weight, although it should be borne in mind that the differences between LG and HG with respect to psychopathological traits could arise by sample unbalancing. Finally, all data were collected using self-reports, which could lead to common problems of counterfeiting. Further studies are needed not only to replicate our results but also for testing the GAMS-it with multiple reports (i.e., relatives, friends) or with behavioral and objective measures.

Future prospects may be the application of GAMS-it on participants with IGD diagnosis, or to investigate the relationship between GAMS-it subscales, gender, and the preference for different kind of video games. Also, the correlation with other factors as aggression, impulsivity or addiction could be studied. Another issue to be studied is the correlation between videogames motivation and specific VG characteristics, namely graphic characteristics, immersivity, game option and history, and their potential link with mental and behavioral health [44]. Finally, it would be interesting to evaluate gaming motivation in different age ranges, in order to have an overview of gaming motivation at every time of life.

5. Conclusions

In conclusion, the present findings provide a coherent picture supporting the reliability and validity of the GAMS-it. Specifically, the results supported the internal consistency of the six subscales; the six-factor structure of the GAMS.it; interesting correlation with CES-D; and STAI scores.

The findings from the present study suggest that the GAMS-it is a valid assessment of gaming motivation in the Italian background and it significantly enables the prediction of anxiety and depression levels.

Author Contributions: Conceptualization, G.C. and S.P.; methodology, G.C. and F.P.; validation, G.C., S.P. and F.P.; formal analysis, F.P.; writing—original draft preparation, G.C., S.P. and F.P.; writing—review and editing, G.C. and S.P.; supervision, G.C.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Föcker, J.; Cole, D.; Beer, A.L.; Bavelier, D. Neural bases of enhanced attentional control: Lessons from action video game players. *Brain Behav.* **2018**, *8*, e01019. [CrossRef] [PubMed]
2. Berard, A.V.; Cain, M.S.; Watanabe, T.; Sasaki, Y. Frequent video game players resist perceptual interference. *PLoS ONE* **2015**, *10*, e0120011. [CrossRef]
3. Schenk, S.; Lech, R.K.; Suchan, B. Games people play: How video games improve probabilistic learning. *Behav. Brain Res.* **2017**, *335*, 208–214. [CrossRef] [PubMed]
4. Mack, D.J.; Wiesmann, H.; Ilg, U.J. Video game players show higher performance but no difference in speed of attention shifts. *Acta Psychol.* **2016**, *169*, 11–19. [CrossRef] [PubMed]
5. Wilms, I.L.; Petersen, A.; Vangkilde, S. Intensive video gaming improves encoding speed to visual short-term memory in young male adults. *Acta Psychol.* **2013**, *142*, 108–118. [CrossRef] [PubMed]
6. Stanmore, E.; Stubbs, B.; Vancampfort, D.; de Bruin, E.D.; Firth, J. The effect of active video games on cognitive functioning in clinical and non-clinical populations: A meta-analysis of randomized controlled trials. *Neurosci. Biobehav. Rev.* **2017**, *78*, 34–43. [CrossRef] [PubMed]
7. Prescott, A.T.; Sargent, J.D.; Hull, J.G. Metaanalysis of the relationship between violent video game play and physical aggression over time. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 9882–9888. [CrossRef] [PubMed]
8. Swing, E.L.; Gentile, D.A.; Anderson, C.A.; Walsh, D.A. Television and video game exposure and the development of attention problems. *Pediatrics* **2010**, *126*, 214–221. [CrossRef]
9. González-Bueso, V.; Santamaría, J.J.; Fernández, D.; Merino, L.; Montero, E.; Ribas, J. Association between Internet Gaming Disorder or Pathological Video-Game Use and Comorbid Psychopathology: A Comprehensive Review. *Int. J. Environ. Res. Public Health* **2018**, *15*, 668. [CrossRef]
10. Stockdale, L.; Morrison, R.G.; Palumbo, R.; Garbarino, J.; Silton, R.L. Cool, callous and in control: Superior inhibitory control in frequent players of video games with violent content. *Soc. Cogn. Affect. Neurosci.* **2017**, *12*, 1869–1880. [CrossRef]
11. Anderson, C.A.; Bushman, B.J.; Bartholow, B.D.; Cantor, J.; Christakis, D.; Coyne, S.M.; Donnerstein, E.; Brockmyer, J.F.; Gentile, D.A.; Green, C.S.; et al. Screen Violence and Youth Behavior. *Pediatrics* **2017**, *140* (Suppl. 2), S142–S147. [CrossRef]
12. Poulain, T.; Vogel, M.; Buzek, T.; Genuneit, J.; Hiemisch, A.; Kiess, W. Reciprocal Longitudinal Associations Between Adolescents' Media Consumption and Sleep. *Behav. Sleep Med.* **2018**, 1–15. [CrossRef]
13. Molde, H.; Holmøy, B.; Merkesdal, A.G.; Torsheim, T.; Mentzoni, R.A.; Hanns, D.; Sagoe, D.; Pallesen, S. Are Video Games a Gateway to Gambling? A Longitudinal Study Based on a Representative Norwegian Sample. *J. Gambl. Stud.* **2018**. [CrossRef]
14. Colasante, E.; Gori, M.; Bastiani, L.; Scalese, M.; Siciliano, V.; Molinaro, S. Italian adolescent gambling behaviour: Psychometric evaluation of the South Oaks Gambling Screen: Revised for Adolescents (SOGS-RA) among a sample of italian students. *J. Gambl. Stud.* **2013**, *30*, 789–801. [CrossRef]
15. Beranuy, M.; Carbonell, X.; Griffiths, M.D. A qualitative analysis of online gaming addicts in treatment. *Int. J. Ment. Health Addict.* **2013**, *11*, 149–161. [CrossRef]
16. American Psychiatric Association (APA). *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed.; American Psychiatric Association: Arlington, VA, USA, 2013; ISBN 0890425558.
17. World Health Organization (WHO). *International Statistical Classification of Diseases and Related Health Problems*; 11th Revision; WHO: Geneva, Switzerland, 2018. Available online: https://icd.who.int/browse11/l-m/en (accessed on 21 January 2019).
18. Donati, M.A.; Sottili, E.; Morsanyi, K.; Primi, C. Time Perspectives and Gambling in Adolescent Boys: Differential Effects of Present- and Future-Orientation. *J. Gambl. Stud.* **2018**. [CrossRef]
19. Buja, A.; Lion, C.; Scioni, M.; Vian, P.; Genetti, B.; Vittadello, F.; Sperotto, M.; Simeoni, E.; Baldo, V. SOGS-RA gambling scores and substance use in adolescents. *J. Behav. Addict.* **2017**, *6*, 425–433. [CrossRef]
20. Johnston, L.D.; O'Malley, P.M.; Bachman, J.G.; Schulenberg, J.E. *Monitoring the Future National Survey Results on Drug Use, 1975–2004*; National Institute on Drug Abuse: Bethesda, MD, USA, 2005.

21. Deci, E.L.; Ryan, R.M. The "what" and "why" of goal pursuit: Human needs and the self-determination of behavior. *Psychol. Inq.* **2000**, *11*, 227–268. [CrossRef]
22. Lafreniére, M.A.; Verner-Filion, J.; Vallerand, R.J. Development and validation of the Gaming Motivation Scale (GAMS). *Personal. Individ. Differ.* **2012**, *53*, 827–831. [CrossRef]
23. Meyer, P.J.; King, C.P.; Ferrario, C.R. Motivational Processes Underlying Substance Abuse Disorder. In *Behavioral Neuroscience of Motivation. Current Topics in Behavioral Neurosciences*; Simpson, E., Balsam, P., Eds.; Springer: Cham, Switzerland, 2015; Volume 27.
24. Ryan, R.M.; Rigby, C.S.; Przybylski, A.K. The motivational pull of video games: A self-determination theory approach. *Motiv. Emot.* **2006**, *30*, 347–364. [CrossRef]
25. Przybylski, A.K.; Rigby, C.S.; Ryan, R.M. A motivational model of video game engagement. *Rev. Gen. Psychol.* **2010**, *14*, 154–166. [CrossRef]
26. Krossbakken, E.; Pallesen, S.; Mentzoni, R.A.; King, D.L.; Molde, H.; Finserås, T.R.; Torsheim, T. A Cross-Lagged Study of Developmental Trajectories of Video Game Engagement, Addiction, and Mental Health. *Front. Psychol.* **2018**, *9*, 2239. [CrossRef]
27. Torres-Rodríguez, A.; Griffiths, M.D.; Carbonell, X.; Oberst, U. Internet gaming disorder in adolescence: Psychological characteristics of a clinical sample. *J. Behav. Addict.* **2018**, *7*, 707–718. [CrossRef]
28. Bonnaire, C.; Baptista, D. Internet gaming disorder in male and female young adults: The role of alexithymia, depression, anxiety and gaming type. *Psychiatry Res.* **2018**, *272*, 521–530. [CrossRef]
29. Wölfing, K.; Müller, K.W.; Beutel, M. Reliability and Validation of the Assessment of Internet and Computer Game Addiction Scale (AICA-S). *Psychother. Psychosom. Med. Psychol.* **2011**, *61*, 216–224.
30. Radloff, L.S. The CES-D scale: A self report depression scale for research in the general population. *Appl. Psychol. Meas.* **1977**, *1*, 385–401. [CrossRef]
31. Spielberger, C.D.; Gorsuch, R.L.; Lushene, R.; Vagg, P.R.; Jacobs, G.A. *Manual for the State-Trait Anxiety Inventory*; Consulting Psychologists Press: Palo Alto, CA, USA, 1983.
32. Müller, K.W.; Koch, A.; Dickenhorst, U.; Beutel, M.E.; Duven, E.; Wölfling, K. Addressing the Question of Disorder-Specific Risk Factors of Internet Addiction: A Comparison of Personality Traits in Patients with Addictive Behaviors and Comorbid Internet Addiction. *BioMed Res. Int.* **2013**, *7*, 546342. [CrossRef]
33. Peracchia, S.; Presaghi, F.; Curcio, G. Adattamento italiano della Assessment of Internet and Computer game Addiction scale (AICAs-Ita). *Rass. Psicol.* **2018**, *XXXV*, 5–17.
34. Flora, D.; Curran, P. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychol. Methods* **2004**, *9*, 466–491. [CrossRef]
35. Satorra, A.; Bentler, P.M. Scaling corrections for chi-square statistics in covariance structure analysis. *ASA Proc. Bus. Econ. Sect.* **1988**, 308–313.
36. Steiger, J.H. Structural model evaluation and modification: An interval estimation approach. *Multivar. Behav. Res.* **1990**, *25*, 173–180. [CrossRef]
37. Bentler, P.M. Comparative fit indices in structural models. *Psychol. Bull.* **1990**, *107*, 238–246. [CrossRef]
38. Bentler, P.M.; Bonett, D.G. Significance tests and goodness of fit in the analysis of covariance structures. *Psychol. Bull.* **1980**, *88*, 588–606. [CrossRef]
39. Hu, L.-T.; Bentler, P.M.; Kano, Y. Can test statistics in covariance structure analysis be trusted? *Psychol. Bull.* **1992**, *112*, 351–362. [CrossRef]
40. Millsap, R.E.; Yun-Tein, J. Assessing factorial invariance in ordered-categorical measures. *Multivar. Behav. Res.* **2004**, *39*, 479–515. [CrossRef]
41. McDowell, C.P.; MacDonncha, C.; Herring, M.P. Associations of physical activity with anxiety and depression symptoms and status among adolescents. *J. Adolesc.* **2017**, *55*, 1–4. [CrossRef]
42. Riva, G.; Gaggioli, A.; Villani, D.; Cipresso, P.; Repetto, C.; Serino, S.; Triberti, S.; Brivio, E.; Galimberti, C.; Graffigna, G. Positive Technology for Healthy Living and Active Ageing. In *Active Ageing and Healthy Living: A Human Centered Approach in Research and Innovation as Source of Quality of Life*; Riva, G., Marsan, P.A., Grassi, C., Eds.; IOS Press: Amsterdam, The Netherlands, 2014; pp. 44–56. [CrossRef]

43. Vallerand, R.J.; Pelletier, L.G.; Blais, M.R.; Brière, N.M.; Senécal, C.B.; Vallières, E.F. The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educ. Psychol. Meas.* **1992**, *52*, 1003–1017. [CrossRef]
44. Colder Carras, M.; Kalbarczyk, A.; Wells, K.; Banks, J.; Kowert, R.; Gillespie, C.; Latkin, C. Connection, meaning, and distraction: A qualitative study of video game play and mental health recovery in veterans treated for mental and/or behavioral health problems. *Soc. Sci. Med.* **2018**, *216*, 124–132. [CrossRef]

International Journal of
Environmental Research and Public Health

MDPI

Article

Measuring the Prevalence of Adverse Childhood Experiences by Survey Research Methods

Anikó Ujhelyi Nagy [1,*], Ildikó Kuritár Szabó [1], Endre Hann [2] and Karolina Kósa [1]

1 Institute of Behavioral Sciences, Faculty of Public Health, University of Debrecen, Móricz Zsigmond krt. 22., H-4032 Debrecen, Hajdú-Bihar, Hungary; szabo.ildiko@sph.unideb.hu (I.K.S.); kosa.karolina@sph.unideb.hu (K.K.)

2 Median Opinion & Market Research Ltd., Fürj utca 2., H-1124 Budapest, Hungary; hann@median.hu

* Correspondence: nagy.aniko@sph.unideb.hu; Tel.: +36-305355520

Received: 30 December 2018; Accepted: 20 March 2019; Published: 22 March 2019

Abstract: *Background*: Child maltreatment has been firmly established as a fundamental risk factor for adult health. However, its quantification poses many questions methodologically, psychologically, and culturally alike. We carried out the first nationally representative survey research in Hungary and in Central–Eastern Europe to assess the prevalence of adverse childhood experiences (ACEs) among adults. *Methods*: Data were collected by an opinion research company using a screening tool of the Adverse Childhood Experiences study. *Results*: 25% (n = 293) of adults reported any childhood adversity; 5% (n = 59) of them had four or more ACEs. The most prevalent forms of child maltreatment were emotional (5%, n = 59) and physical abuse (5%, n = 59), sexual abuse (1%, n = 12) being the least prevalent. The most frequent dysfunctional household condition was parental divorce or separation (13%, n = 153), followed by household substance abuse (11%, n = 129). *Conclusions*: Nationally representative surveys on ACEs found a range of overall prevalence of various forms of child maltreatment between 14.1 and 35.2% into which our results fall. Nevertheless, our survey most likely underestimates the prevalence of child maltreatment in Hungary, reflecting the impact of a host of factors influencing awareness. Survey research methods are appropriate to obtain nationally representative data on child maltreatment that not only contribute to designing interventions but can also be used to monitor the effectiveness of interventions to improve child and adult health in the long run.

Keywords: adverse childhood experiences (ACEs); Hungarian representative adult sample; opinion poll; ACE Score Calculator

1. Introduction

Exposure to various forms of adversity early in life has been shown to lead to an increased risk of a broad range of developmental difficulties, principally cognitive, emotional, and behavioral impairments during childhood that are mediated by compromised neurodevelopment affecting various parts of the brain [1–7]. The consequences of childhood maltreatment can last well into adulthood or even throughout life, impacting adult physical health, mental pathology, and quality of life [8–11]. Numerous studies have shown that adverse childhood experiences (ACEs, including forms of child maltreatment and household dysfunctions) are major risk factors for acute and chronic somatic and mental diseases such as anxiety or post-traumatic disorders mediated by risk behaviors such as smoking, alcohol and drug abuse, suicide attempts, aggressive behaviors, risky sexual behaviors, and low mental resilience [12–20]. Previous studies provided strong evidence that ACEs tend to co-occur in which intergenerational transmission of adversity might be a contributing factor [21–24].

Prevention of these early adversities is much more effective than treatment of their consequences with their enormous burden in health and social care, as well as in the education system [25,26]. National policies and evidence-based prevention programs (at local and societal levels) based on early recognition of ACEs may contribute to preventing a wide range of health-harming behaviors, somatic and mental disorders, and early death [15,27,28]. All such policies, programs, and interventions should be based on an in-depth knowledge of the population pattern of ACEs. However, collecting relevant information has been hindered either by lack of awareness about the issue and/or by a lack of relatively simple and cost-effective methods of collecting information in various population groups.

Tested Methodologies for Studying Childhood Adversity

The causal relationship between childhood adversity and its adult health consequences, including mental and somatic health impairments, have been established by prospective longitudinal cohort studies such as the Lehigh Longitudinal Study of the US established in 1976 [29], the Christchurch Health and Development Study established in 1977 [30], and the Adverse Childhood Experiences (ACE) Study in 1995 [15]. The majority of research collected information on childhood adversity either from the primary caretaker of the child in cases of prospective studies or from adult self-reports in terms of their childhood in retrospective or cross-sectional studies. Retrospective assessment of ACEs based on self-report was shown to be reliable and valid for research purposes [31–33]. Retrospective recall of ACEs can be considered valid if these experiences are operationalized unequivocally, making interpretation and judgment of the questions unnecessary [31,32]. Data can be collected in various ways such as by questionnaire during personal interview [30,34,35]; mailing the questionnaire to respondents by post or by email [15]; or by telephone interviews [36,37].

In order to make an evidence-based statement about the pattern of childhood adversity in any given population, survey research should be designed producing reliable population estimates from samples that represent the entire population of interest. A practical handbook on measuring and monitoring national prevalence of child maltreatment published by the World Health Organization promotes system-wide monitoring of child maltreatment in European countries and globally with the emphasis on estimating population-wide prevalence rates based on representative survey samples [38]. However, many studies reporting child maltreatment rely on clinical and other nonrepresentative samples drawn from various public services such as education, health care, social services, or family and child protective services that make the generalization of findings difficult (Figure 1) [39].

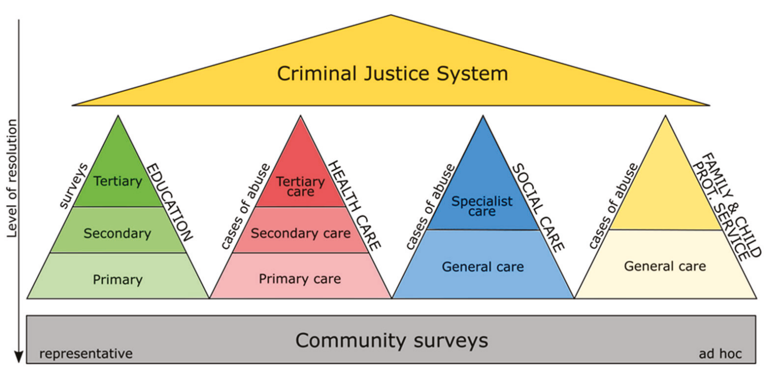


Figure 1. Sources of data for child maltreatment.

So far, only one research study has been published in the literature that assessed childhood adversity in a nationally representative sample in which fieldwork was carried out by a government-financed agency. Namely, the second wave (2004–2005) of the National Epidemiological Survey on Alcohol and Related Conditions (NESARC) in the US collected, among others, data on adverse childhood events by face-to-face interviews conducted by trained lay interviewers of the US Census Bureau in a nationally representative adult sample of 34,653 persons from the United States. Based on these data, the prevalence of emotional abuse (4.8%) and of emotional neglect (6.2%) were estimated [35].

The ACE study was initiated by a health insurance organization among a subset of its clients, and its ongoing surveillance is limited to the participants of the original study [40].

The Behavioral Risk Factor Surveillance System (BRFSS) of the Centers for Disease Control and Prevention (CDC) in the US has been collecting data about adverse childhood experiences since 2009 by the request of individual states of which 32 requested such data collection [41].

Survey research, that is, data collection from a carefully selected nationally representative sample, requires human and financial resources that are beyond the reach of academic institutes, or even governmental agencies in most countries—save for the US. On the other hand, polling companies have vast survey research experience gathering information on a wide range of topics. This experience was taken advantage of in two European studies that used survey research methods to study the epidemiology of ACEs in nationally representative samples. One of the studies was carried out on a representative sample of 2504 German participants between 14 and 92 years by face-to-face interviews on childhood abuse and neglect, as well as current anxiety and depression. Data collection was carried out by an independent institute for opinion and social research [42]. The other study was done in a sample of 3885 adults representative of England in which information on childhood experiences and adult mental well-being was collected during personal visits by a professional survey company directed by researchers [43].

Encouraged by these antecedents, our aim was to obtain data on the prevalence of adverse childhood experiences in the adult Hungarian population in line with the recommendations of the World Health Organization using opinion research methodology.

2. Materials and Methods

2.1. Sampling and Data Collection

A market research company (Median Opinion and Market Research Institute) was contracted based on its outstanding performance predicting election results in Hungary. A multistage stratified cluster sampling was carried out using the most recent census list (2011) of the Central Statistical Office of Hungary. Based on detailed maps of the country, 120 sampling units were selected by a computer program of Median Opinion and Market Research Institute. Sampling units represented the entire territory of the country according to EUROSTAT NUTS II levels and according to the distribution of the resident population in terms of metropolitan, urban, and rural areas. One starting address was randomly drawn in each selected sampling unit from which nine other households were accessed by random walking (10 households per sampling unit). One respondent 18 years or older was interviewed in each household by using the Kish selection grid [44]. Each selected person was contacted in person at least three times if the first attempt was not successful. In case of refusal, the interviewer had to select another respondent in another household based on a preset algorithm.

General questions were asked face to face, whereas the questions relating to childhood adversity were filled by the respondents themselves. The interviewer handed over the paper-based questionnaire to the respondent in person. The respondents were allowed to submit their responses in a sealed envelope upon request; 10 % of the completed interviews were validated by face-to-face or telephone re-interview.

Altogether, 1200 persons aged 18 years or older were interviewed out of 1608 who were attempted to be interviewed (74.6% response rate). All interviewees received a written statement about data collection being voluntary and conforming to the requirements of the national data protection act; none of them received incentive in any form. Data collection was carried out by trained interviewers in person in March 2016.

Median Opinion and Market Research Institute is one of leading research companies in Hungary, conducting high-quality market, opinion, and social research. The institute follows the professional and ethical guidelines specified in the ESOMAR Code of Conduct [45]. During the present research, informed consent was provided and the appropriate ethical standards (according to the World Medical Association Declaration of Helsinki) were followed. The protocol of research was approved by the Medical Research Council of the University of Debrecen (4499-2015).

2.2. Measures

2.2.1. Sociodemographic Variables

Age, gender, marital status (unmarried, married, registered partnership, divorced, widow), type of the settlement of permanent residence (capital, city, village), education (less than primary, primary, vocational, high school diploma, college/university), type of work (manual vs nonmanual), employment (nine categories), and current household income (four quartiles) were registered. Sociodemographic categories were identical to those used by the Central Statistical Office of Hungary.

2.2.2. Adverse Childhood Experiences

Adverse childhood experiences were assessed by the ACE Score Calculator, a validated screening instrument used to estimate the prevalence of ACEs [46]. This tool of 10 items, developed by the researchers of the ACE study based on the original ACE Questionnaire, is appropriate for screening purposes and allows individuals to calculate their own scores based on the original scoring criteria of the ACE Study [15]. A short form of eight items of the original ACE Questionnaire was also used in the Health Behavior in School-Aged Children (HBSC) Study and proved to be reliable [38].

The ACE Score Calculator helps assess exposure to 10 types of ACEs including 5 types of abuse (emotional, physical, and sexual), neglect (physical and emotional), and 5 types of dysfunctional family environment (mentally ill or substance-abusing member of household, physical violence in the household, parental separation/divorce, incarcerated family member(s) prior to age 18). The ACE Score is calculated by summing up all 10 ACE variables and serves as a measure of overall ACE exposure ranging from 0 (meaning no exposure to the 10 categories of ACEs) to 10 (meaning exposure to all 10 categories). Responses were categorized by type of ACE and were dichotomized into no history or any history of adversity prior to age 18. Responses were analyzed also by number of adverse experiences (none, 1, 2–3, 4 or more) prior to age 18.

Dube et al. (2014) found good to excellent reliability in the reports of ACEs during adulthood. The test–retest reliability in the responses to questions about ACEs and the resulting ACE score was found to be good and moderate to substantial. These findings confirm that retrospective responses to the forms of childhood maltreatment and household dysfunction are generally stable over time [31]. Wingenfeld et al. (2010) investigated the psychometric characteristics of the ACE Score Calculator and revealed that it is a reliable, valid, and economic screen for the retrospective assessment of ACEs [47].

The English version of the ACE Score Calculator was translated to Hungarian by the authors, and cross-cultural adaptation was carried out through an iterative forward–backward translation compared by an independent third person. The preambles, item contents, and response options for items can be found in the Appendix A (Table A1).

2.3. Data Analysis

In order to obtain estimates of adverse childhood experiences in the adult noninstitutionalized population of Hungary, statistical weights were applied to ensure that estimates reflect the general adult Hungarian population gender, age group, education, settlement type, and region. The sample defined as survey sample was analyzed using Stata/IC 13.1. Single-stage design was used stratifying the sample based on the sampling units, that is, regions of the country. The Taylor method was used to estimate sampling errors; primary sampling units were sampled without replacement [48].

Analysis of variance was computed to examine the prevalence of child maltreatment and household dysfunctions by total ACE core and by type of ACE stratified by gender. The sociodemographic characteristics of respondents reporting no history (ACE score = 0) or any history of adversity (ACE score >0) were described defining the ACE score as a discrete interval variable. The number of categories of demographic variables such as education and employment type was combined to reduce the number of categories and to simplify interpretation. Considering the weighted estimates, all prevalence data were rounded.

Logistic regression was carried out by backward stepwise regression to identify the independent variables of childhood adversity. The ACE score was defined as the binary outcome variable as described above (no childhood adversity vs any history of adversity). One binary ACE score was also created considering only the five types of childhood maltreatment, and another by including only the five types of family dysfunction. Age (in years), sex (female, male), place of residence (capital, city, village), education (higher education vs less), type of work (nonmanual vs manual), and marital status (single, divorced, married, cohabiting, widowed) were tested as categorical explanatory variables. In terms of marital status, two models were tested. The first model compared those in an ongoing relationship (married or cohabiting) to all other marital categories (single, divorced, widowed) including all respondents. The second model compared those in an ongoing relationship (married or cohabiting) to those who ended their relationship (divorced) including only those with a (supposedly) living present or past partner. Post-test analysis was carried out by the adjusted Wald test.

3. Results

3.1. Sample

Altogether, 1200 persons aged 18 years or older were interviewed, representing 0.012% of the Hungarian adult noninstitutionalized population according to the census in 2011. Respondents ranged in age from 18 years to 112 years and the mean age for the sample was 53.2 (SD = 16.5) years; 37.65% of the respondents were men. The sample was weighted to represent the Hungarian adult population by gender, age group, education, settlement type, and region. Of all the persons who had been approached, 74.6% were willing to fill out the questionnaire. Twenty-six individuals filled out the screening instrument incompletely: 17 did not complete the full questionnaire, 9 respondents answered all but one question. They were dropped from the analysis, leaving a total sample size of 1174 corresponding to a completion rate of 97.8%.

3.2. The prevalence of Adverse Childhood Experiences in Hungary

The distribution of sociodemographic characteristics overall and by reporting an ACE (no ACE vs at least one ACE) is provided in Table 1. The highest prevalence of any adversity, 28% (n = 82) was found in the youngest age group (18–29 years) that was declining and was half of that among the 50–59-year-olds, but somewhat increased in the oldest age group (60+, 23%, n = 67). Experience of childhood adversity was more than twice as high in cities compared to the capital or to villages. Interestingly, the ACE score was by far the lowest among the least-educated group and highest among those with high school qualification. Income was mildly significantly related to the experience of adversity: those in the lowest income quartile had the highest proportion of any adversity. One-quarter more of those who suffered any ACE had been unmarried or divorced compared to persons who did not report any ACE.

Of the adult Hungarian population, 25% (n = 293) reported having experienced some kind of childhood adversity before the age of 18 years; 5% (n = 59) of the respondents had four or more ACEs. There were no significant gender differences regarding the co-occurrence of ACEs (Pearson's chi-squared test, p = 0.29) (Table 2). Considering only those between the ages of 18 and 80 years, the prevalence of any abuse did not change (25%).

Table 1. Distribution of sociodemographic characteristics by reporting/not reporting an ACE.

Sociodemographic Variable	Did not Report an ACE (*N* = 904) *N* (%)	Reported an ACE (*N* = 292) *N* (%)	Total (*N* = 1196) *N* (%)	*p*-Value [a]
Age group				
18–29	127 (14)	82 (28)	209 (17)	
30–39	163 (18)	55 (19)	218 (18)	
40–49	154 (17)	50 (17)	204 (17)	0.0003 *
50–59	181 (20)	38 (13)	219 (18)	
60+	280 (31)	67 (23)	347 (30)	
Type of the settlement of permanent residence				
Capital	154 (17)	58 (20)	212 (18)	
City	470 (52)	162 (55)	632 (53)	0.04 *
Village	280 (31)	72 (25)	352 (29)	
Education				
Less than 8 grades	18 (2)	9 (3)	27 (2)	
Primary school	181 (20)	44 (15)	225 (19)	
Vocational training	262 (29)	96 (33)	358 (30)	0.265
High school	271 (30)	105 (36)	376 (31)	
College graduate	172 (19)	38 (13)	210 (18)	
Type of work				
Manual	515 (57)	181 (62)	696 (58)	0.647
Nonmanual	389 (43)	111 (38)	500 (42)	
Income quartiles				
Low	199 (22)	85 (29)	284 (24)	
Lower middle	244 (27)	70 (24)	314 (26)	0.035 *
Upper middle	181 (20)	67 (23)	248 (21)	
Upper	280 (31)	70 (24)	350 (29)	
Number of children				
0	660 (73)	184 (63)	844 (71)	
1	118 (13)	55 (19)	173 (14)	
2	81 (9)	38 (13)	119 (10)	0.161
3	36 (4)	12 (4)	48 (4)	
>4	9 (1)	3 (1)	12 (1)	

[a] Indicates the application of the Pearson's chi-squared test, * $p < 0.05$.

Table 2. Distribution of ACE Score overall and by gender.

ACE Score	Women (N = 732) N (%)	Men (N = 442) N (%)	Total (N = 1174) N (%)	*p*-Value [a]
0	542 (74)	336 (76)	881 (75)	
1	88 (12)	53 (12)	141 (12)	
2	44 (6)	18 (4)	59 (5)	$p = 0.29$
3	22 (3)	13 (3)	34 (3)	
≥4	36 (5)	22 (5)	59 (5)	

[a] Indicates the application of the Pearson's chi-squared test, $p = 0.29$.

The most prevalent form of self-reported child maltreatment was emotional abuse (5%, n = 59), and physical abuse (5%, n = 59) in this nationally representative sample. The least prevalent pattern was sexual abuse (1%, n = 12). The most frequent dysfunctional household condition was parental divorce or separation (13%, n = 153), followed by household substance abuse (11%, n = 129). The least prevalent household dysfunction was having an incarcerated household member (4%, n = 47).

Among women, emotional abuse and physical abuse were more prevalent (7% (n = 51) for emotional abuse and 6% (n = 44) for physical abuse) than among men (4% (n = 18) for physical abuse and 3% (n = 13) for emotional abuse). Male participants witnessed household physical violence more often (violence against their mother or stepmother) according to their self-report. Gender differences in emotional abuse have been shown to be significant (Pearson's chi-squared test, $p = 0.03$) (Table 3).

Table 3. Reported prevalence of ACEs in the sample and by gender.

Adverse Childhood Experiences (ACEs)	Women (N = 732) N (%)	Men (N = 442) N (%)	Total (N = 1174) N (%)	*p*-Value [a]
Maltreatment				
Emotional abuse	51 (7)	13 (3)	59 (5)	0.03 *
Physical abuse	44 (6)	18 (4)	59 (5)	0.09
Sexual abuse	15 (2)	0 (0)	12 (1)	
Emotional neglect	29 (4)	18 (4)	47 (4)	0.54
Physical neglect	22 (3)	9 (2)	35 (3)	0.68
Family dysfunction				
Parental separation/ divorce	102 (14)	53 (12)	153 (13)	0.45
Household physical violence	29 (4)	27 (6)	59 (5)	0.12
Household substance abuse	95 (13)	44 (10)	129 (11)	0.16
Household mental illness	44 (6)	18 (4)	59 (5)	0.27
Incarcerated household member	29 (4)	18 (4)	47 (4)	0.69

[a] Indicates the application of the Pearson's chi-squared test, * $p < 0.05$.

Respondents having experienced four or more ACEs were younger (57% (n = 29) were 18–39 years old), more of them lived outside the capital (75% (n = 38) lived in cities or villages), belonged to the lowest income category (42%, n = 21), were married (49%, n = 24), and had no children (43%, n = 21) (Table 4).

Table 4. Distribution of sociodemographic characteristics by ACE Score.

Sociodemographic Variable	ACE Score = 0 (N = 904) N (%)	ACE Score = 1 (N = 153) N (%)	ACE Score = 2–3 (N = 89) N (%)	ACE Score $\geq$ 4 (N = 50) N (%)
Age group				
18–29	127 (14)	44 (29)	20 (23)	14 (27)
30–39	163 (18)	20 (13)	20 (23)	15 (30)
40–49	154 (17)	29 (18)	15 (16)	6 (13)
50–59	181 (20)	20 (13)	13 (15)	6 (13)
60+	280 (31)	40 (26)	21 (24)	9 (17)
Type of the settlement of permanent residence				
Capital	154 (17)	35 (23)	11 (12)	12 (25)
City	470 (52)	83 (53)	60 (67)	19 (38)
Village	280 (31)	35 (23)	18 (21)	19 (37)
Education				
Less than 8 grades	18 (2)	2 (1)	3 (4)	2 (5)
Primary school	181 (20)	23 (15)	15 (16)	6 (12)
Vocational training	262 (29)	45 (30)	31 (35)	18 (36)
High school	271 (30)	63 (41)	30 (34)	14 (28)
College graduate	172 (19)	20 (13)	10 (11)	10 (19)
Type of work				
Manual	515 (57)	86 (56)	65 (73)	30 (59)
Nonmanual	389 (43)	67 (44)	24 (27)	20 (41)
Income quartiles				
Low	199 (22)	35 (23)	25 (28)	21 (42)
Lower middle	244 (27)	35 (23)	22 (25)	2 (23)
Upper middle	181 (20)	40 (26)	24 (27)	5 (11)
Upper	280 (31)	43 (28)	18 (20)	12 (24)
Marital status				
Married	389 (43)	43 (28)	31 (35)	24 (49)
Registered partnership	63 (7)	9 (6)	10 (11)	8 (17)
Divorced	145 (16)	25 (23)	18 (21)	5 (11)
Widow	118 (13)	14 (21)	10 (9)	7 (4)
Unmarried	190 (21)	29 (45)	23 (20)	17 (9)
Number of children				
0	660 (73)	112 (73)	54 (61)	21 (43)
1	118 (13)	21 (14)	26 (29)	10 (19)
2	81 (9)	15 (10)	7 (8)	11 (22)
3	36 (4)	5 (3)	2 (2)	5 (11)
>4	9 (1)	0 (0)	0 (0)	3 (5)

3.3. Modeling the Determinants of Adverse Childhood Experiences in Hungary

Independent determinants of adverse childhood experiences were analyzed by logistic regression defining childhood abuse as a binary variable (experienced vs not experienced). Any childhood maltreatment (5 types of abuse or neglect of any sort), family dysfunction (5 types), and both combined, that is, any adverse experience (10 types), as described in Methods, as outcome variables were modeled. Marital status was defined in three different ways. In Model 1, all respondents were divided into two categories: those in an actual relationship (married or cohabiting) and those not currently in a relationship (single, divorced, widowed) (Table 5). One quartile increase of income decreased the odds of reporting any childhood maltreatment by 24% in Model 1, but income did not remain a significant determinant of reporting childhood adversity in either of the models.

In Model 2, only persons with a living present or past partner were included. Those in actual relationships (married or cohabiting) were compared to those who were divorced (Table 6). The latter produced a better model in which living in a relationship decreased the odds of reporting maltreatment by 35%. In the same model, one year increase in age decreased the odds of reporting any childhood maltreatment by 2.3%, and living in a city compared to a village increased the odds of reporting any maltreatment by 76%. Age, type of permanent residence, and marital status were found to be independent determinants of family dysfunction or any adverse childhood experience in the best adjusted logistic regression model (Model 2).

In Model 3, persons with relationship experience were included. Those in actual relationships (married or cohabiting) were compared to those who were divorced or widowed (Table 6). Model 2 and 3 both were statistically significant, showing that currently living in a relationship decreased the odds of reporting any childhood adversity by at least 40% compared to those who had relationship experience but did not currently live in one.

Other variables such as education, gender, number of children, type of work (manual, nonmanual) were not found to be significant determinants of reporting any childhood maltreatment, any family dysfunction, or a combination of both (data not shown).

Table 5. Independent determinants of experiencing any of the five types of maltreatment or five types of dysfunctions in the family (Model 1).

Sociodemographic Variables	**Any Childhood Maltreatment**			**Family Dysfunction**			**Maltreatment & Family Dysfunction Together**		
	OR	***p*-Value**	**Model**	**OR**	***p*-Value**	**Model**	**OR**	***p*-Value**	**Model**
Model 1: All respondents									
Age (year)	1.000	0.925	F = 2.19 $p = 0.068$	0.979	0.000	F = 7.53 $p < 0.001$	0.983	0.002	F = 5.99 $p < 0.001$
City (ref: village)	1.376	0.226		1.765	0.007		1.546	0.026	
Living in relationship -married or cohabiting (ref: not in relationship)	0.827	0.409		0.726	0.079		0.698	0.038	
Income (quartiles)	0.759	0.015		0.874	0.092		0.894	0.141	

Table 6. Independent determinants of experiencing any of the five types of maltreatment or five types of dysfunctions in the family (Model 2 and Model 3).

	Any Childhood Maltreatment			**Family Dysfunction**			**Maltreatment & Family Dysfunction**		
Sociodemographic Variables	**OR**	***p*-Value**	**Model**	**OR**	***p*-Value**	**Model**	**OR**	***p*-Value**	**Model**
Model 2: Only respondents in current relationship or divorced									
Age (year)	0.981	0.104	F = 1.88 $p = 0.132$	0.968	0.000	F = 8.55 $p < 0.001$	0.973	0.001	F = 7.32 $p < 0.001$
City (ref: village)	1.464	0.259		2.032	0.006		1.761	0.020	
Living in relationship-married or cohabiting (ref: divorced)	0.652	0.139		0.607	0.025		0.571	0.009	

Table 6. *Cont.*

	Any Childhood Maltreatment			Family Dysfunction			Maltreatment & Family Dysfunction		
Sociodemographic Variables	**OR**	***p*-Value**	**Model**	**OR**	***p*-Value**	**Model**	**OR**	***p*-Value**	**Model**
Model 3: Only respondents in current relationship or divorced/widowed									
Age (year)	0.984	0.103	F = 1.73 $p = 0.158$	0.97	0.000	F = 10.03 $p < 0.001$	0.974	0.000	F = 8.13 $p < 0.001$
City (ref: village)	1.469	0.193		1.981	0.002		1.706	0.011	
Living in relationship-married or cohabiting (ref: divorced)	0.701	0.17		0.58	0.006		0.565	0.003	

4. Discussion

Our research produced the first national and the third European representative survey on adverse childhood experiences in Hungary according to which 25% (n = 293) of the Hungarian general population reported experiencing any childhood adversity before the age of 18 years with no gender difference; 5% (n = 59) of the respondents had four or more ACEs. The most prevalent form of child maltreatment was emotional (5%) and physical abuse (5%, n = 59); sexual abuse (1%, n = 12) was least prevalent. Parental divorce or separation (13%, n = 153), followed by household substance abuse (11%, n = 129) were the most frequent dysfunctional household conditions. The higher prevalence of ACEs among the youngest age group of adults may indicate an increasing awareness due to a more open public attitude and changing public opinion.

Our study is the first survey on adverse childhood experiences in a nationally representative adult sample of the Hungarian population; the first survey in any Central–Eastern European country; and the third such survey in developed countries that used a marketing research company for data collection. The European Commission has already established the feasibility of using marketing research/opinion polling agencies in health research: two reports were published on various aspects of the mental health of the population of EU member states in 2006 and 2010 by contracting companies to carry out representative surveys with multistage probability sampling and face-to-face interviews [49,50].

In order to interpret our results, data from the other two representative national surveys cited in the Introduction were considered. As it is shown in Table 7, the occurrence of the most frequent ACEs substantially varies in these countries, with Hungary having the lowest and Germany the highest prevalence.

Table 7. Prevalence of some ACEs in representative national samples.

	Germany	England	Hungary
Year of survey	2010	2013	2016
Tool	CTQ (28 items)	ACE screening tool (11 items)	ACE screening tool (10 items)
Age group	14–92	18–69	18–112
Sample size	2504	3885	1174
Sample as proportion of the population in the year of survey (%)	0.003	0.006	0.012
Response rate	56%	-	74.6%
Any adversity	68.2	46.4%	25%
Physical abuse (%)	12.0	14.3	5
Emotional abuse (%)	15.0	17.3	5
Sexual abuse (%)	12.6	6.2	1
	Schilling et al. [51] Hauser et al. [42]	Hughes et al. [43]	present study

The NESARC and BRFSS surveys have unique features (study design and implementation by public agencies funded by the federal government of the US) based on phone interviews, repeated measures that may not be easily copied by other countries.

The strengths of our study include the use of an international standardized screening tool (ACE Score Calculator) in a nationally representative adult sample. Sampling and data collection were carried out by an experienced opinion poll company that used refined and tested sampling methods and had trained interviewers with experience in face-to-face data collection. This not only increased the reliability of data but was also cost-effective.

However, the study has limitations as well. The cross-sectional design and the retrospective nature of data collection limits the scope of interpretation; low awareness of the topic in the country probably increases recall bias, especially among older persons. However, since the prevalence of childhood maltreatment did not change significantly when those above the age of 80 years were removed from the analysis, and since the items of the ACE questionnaire are quite specific, not requiring interpretation, recall bias likely did not influence our results. The interference of dissociative defense mechanisms with recall cannot be excluded, but this bias cannot be avoided by any questionnaires. The conspicuously low frequency of childhood adversity among those with the lowest education merits further investigation.

In order to further probe the comparability of our data, the literature was searched for meta-analyses on the prevalence of child maltreatment and dysfunctional households reported by adults (Table 8). According to Stoltenborgh et al., global estimates of the prevalence in self-report studies were 22.6% for physical abuse, 36.3% for emotional abuse, 12.7% for sexual abuse (7.6% among boys and 18.0% among girls), 16.3% for physical neglect, and 18.4% for emotional neglect. These authors opined that the prevalence of child maltreatment seems to be largely similar across the globe. However, this statement is based mostly on research in western countries, mainly in North America and Europe [52].

Some meta-analyses were identified which focused on the prevalence of child maltreatment and dysfunctional households reported by children (Table 9).

The WHO Regional Office for Europe used 105 prevalence estimates from 50 community surveys to estimate the prevalence of sexual abuse as 9.6% (13.4% in girls and 5.7% in boys), physical abuse 22.9%, and emotional abuse 29.1% with no gender difference in the two latter types of abuse. The few studies that focused on neglect found high prevalence: 16.3% for physical and 18.4% for emotional neglect. As Table 8 shows, there are no differences between global and European prevalence estimates considering the majority of forms of maltreatment—the only exception being female sexual abuse with slightly lower prevalence in Europe [53]. The European report opined that prevalence estimates of child maltreatment would be higher in Eastern Europe. However, Gilbert et al. (2009) reported prevalences with a much greater variability in high-income countries: 3.7–16.3% of children experienced parental violence per year, 10.3% suffered from emotional abuse, and 1.4–15.7% suffered from neglect [27].

Table 8. A comparison of prevalence rates (estimates and measured data) of child maltreatment across the globe reported by adults.

Country	Source/ Sample Characteristics	Tools	Prevalence of Child Maltreatment					Prevalence of Household Dysfunctions				
			Physical Abuse	Emotional Abuse	Sexual Abuse (Women, Men)	Emotional Neglect	Physical Neglect	Mental Illness	Substance Abuse	Domestic Violence	Parental Divorce/ Separation	Incarceration
Worldwide	Stoltenborgh et al. [52], WHO [53] meta-analysis, 244 publication, 577 estimates	various questionnaires	22.6	36.3	18, 7.6	18.4	16.3	-	-	-	-	-
Africa			22.8	46.7	20.2, 19.3	-	-	-	-	-	-	-
Asia			16.7	41.6	11.3, 4.1	30.1	-	-	-	-	-	-
Australia			14.3	11.3	21.5, 7.5	40.0	-	-	-	-	-	-
North America			24.0	36.5	20.1, 8.0	-	6.5	-	-	-	-	-
South America			54.8	-	13.4, 13.8	14.5	19.2	-	-	-	-	-
Europe			22.9	29.2	13.5, 5.6	18.4	16.3	-	-	-	-	-
EU	WHO [53] meta-analysis, 50 publication, 105 estimates	various questionnaires	22.9	29.2	13.5, 5.6	-	-	-	-	-	-	-
USA (2016)	Taillieu et al. [35], representative sample, N = 34,653	ACE questionnaire five-point ordinal scale	-	4.8	-	6.2	-	-	-	-	-	-
USA (2004)	Dong et al. [21], N = 18,175	ACE questionnaire	26.4	10.2	21	14.8	9.9	20.3	28.8	24.1	13	6
UK (2016)	Hughes, Lowey, Quigg, & Bellis [43]; nationally representative household survey; N = 3885	ACE questionnaire	14.3	17.3	6.2	-	-	12.1	6.5	12.1	22.6	4.1
Germany (2016)	Schilling et al. [51], representative sample N = 2504	German version of the Childhood Trauma Questionnaire (28 items)	2.8	1.6	1.9	6.6	10.8	-	-	-	-	-
Hungary (2017)	Present study representative sample N = 1174	ACE Score Calculator (10-item screening tool)	5	5	1	3	3	6	12	4	13	3

Table 9. A comparison of prevalence rates (estimates and measured data) of child maltreatment across the globe reported by children.

Country	Source/Sample Characteristics	Tools	Prevalence of Child Maltreatment					Prevalence of Household Dysfunctions				
			Physical Abuse	Emotional Abuse	Sexual Abuse	Emotional Neglect	Physical Neglect	Mental Illness	Substance Abuse	Domestic Violence	Parental Divorce/ Separation	Incarceration
USA (2016)	Turney & Wildeman [54] nationally representative survey N = 95,677 children placed in and adopted from foster care	ACE questionnaire	-	-	-	-	-	8.5	10.5	-	19.9	6.9
Switzerland (1996)	Halpérin et al. [55] nationally representative high-school children	self-constructed questionnaire	20	-	-	-	-	-	-	-	-	-
Romania (2000)	Browne [56] self-report survey children aged 13–14 N = 1295	n.a.	24.0	21.0	9.0	46.0	44.0	-	-	-	-	-
Latvia Lithuania Macedonia Moldova	Sebre et al. [57], UNICEF [58] multicountry survey; children aged 10–14 Latvia (N = 297) Lithuania (N = 300) Macedonia (N = 302) Moldova (N = 246)	various questionnaires	19.0 26.0 12.0 30.0	29.0 33.0 13.0 32.0	- - - -	- - - -	- - - -	- - - -	- - - -	- - - -	- - - -	- - - -

Our population survey measured a considerably lower prevalence of childhood adversity compared to population surveys in Germany or England. This, on one hand, probably reflects underestimation, supported by other data such as the homicide rate under 15 years of age in Hungary that was as high as 0.89 per 100,000 children or the fact that Hungary ranked 23rd out of 27 developed countries based on deaths due to abuse and/or neglect per 100,000 children under the age of 15 [53,59,60]; or that satisfaction with life among young teenagers was the second lowest in Hungary out of 21 developed countries in 2013 [61].

On the other hand, the widely different methods and measurements in various samples (community, clinical, and chance samples) selected by a wide variety of methods severely restricts the comparability of surveys carried out in different countries.

Third, the strong influence of culture, traditions, and religion on the treatment of children including what counts and what does not as maltreatment [62], as well as the possibility of false-negative statements due to psychological motives, must also be taken into account when comparing data on child maltreatment in various countries [32]. The ACE study was seminal in drawing attention to childhood adversity in the US and other developed countries [63], but this topic only recently has commanded attention in Hungary, reflected by the fact that no community-based data collection on childhood adversity had been carried out in the country.

Taking all these points together, the statement of Stoltenborgh and his coauthors (2015) that the prevalence of child maltreatment seems to be largely similar across the globe must be called into question [52]. Moreover, the opposite seems to be likely, which points to the importance of population- or community-based prevalence estimates measured by consistent methodology in each country.

5. Conclusions

Our survey provides a population-based set of reference data upon which a strategy to address childhood adversity should be built and to which future data can be compared. Considering the fact that 1) the design and implementation of national surveys is beyond the resources of Hungarian academic institutes, 2) to our knowledge, no similar survey is being designed or planned by national institutions of public health or child protection, and 3) clinical samples have been known to overestimate the population prevalence of ACEs [39], marketing research methods provide a viable and cost-effective alternative to collect data on this important topic.

Even underestimated population-based data on childhood adversity are better than estimates based on clinical or chance samples or no data at all. Our survey provides the first data on ACEs in Central and Eastern Europe with the aim of advocating for the monitoring of ACEs in the future for which the use of marketing research methods seems to be appropriate. The European report on preventing child maltreatment states that community surveys using international standardized tools should be conducted regularly in order to identify the changes in prevalence rates and the potential risks and to have the opportunity to evaluate the implemented prevention programs [53]. However, until international standardized methods of measuring childhood adversity are developed, countries should aim at quantifying this important public health problem in a scientifically acceptable way for which less or more complex methods are available [64–66], and keep monitoring its tendency in time. If there is an issue in which national surveillance is more important than international comparability, it is childhood adversity, especially considering its long-term impact on the population's well-being.

Author Contributions: Conceptualization, A.U.N., I.K.S., and K.K.; Data curation, E.H.; Formal analysis, A.U.N. and K.K.; Funding acquisition, K.K.; Investigation, E.H.; Methodology, A.U.N., I.K.S., E.H., and K.K.; Project administration, K.K.; Resources, E.H.; Supervision, K.K.; Writing—original draft, A.U.N., I.K.S., and K.K.; Writing—review & editing, A.U.N., I.K.S., and K.K.

Funding: Karolina Kósa has been supported by the grant GINOP-2.3.2-15 while preparing the paper.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A

Table A1. The ACE Score Calculator—preambles, item contents and response options.

Item.	Preamble and Content	ACE Category
	During your first 18 years of life:	
1 [a]	Did a parent or other adult in the household often or very often . . . Swear at you, insult you, put you down, or humiliate you? or Act in a way that made you afraid that you might be physically hurt?	Emotional abuse
2 [a]	Did a parent or other adult in the household often or very often . . . Push, grab, slap, or throw something at you? or Ever hit you so hard that you had marks or were injured?	Physical abuse
3 [a]	Did an adult person at least 5 years older than you ever . . . Touch or fondle you or have you touch their body in a sexual way? or Attempt or actually have oral, anal, or vaginal intercourse with you?	Sexual abuse
4 [a]	Did you often or very often feel that . . . No one in your family loved you or thought you were important or special? or Your family didn't look out for each other, feel close to each other, or support each other?	Emotional neglect
5 [a]	Did you often or very often feel that . . . You didn't have enough to eat, had to wear dirty clothes, and had no one to protect you? or Your parents were too drunk or high to take care of you or take you to the doctor if you needed it?	Physical neglect
6 [a]	Were your parents ever separated or divorced?	Parental separation/divorce
7 [a]	Was your mother or stepmother: Often or very often pushed, grabbed, slapped, or had something thrown at her? or Sometimes, often, or very often kicked, bitten, hit with a fist, or hit with something hard? or Ever repeatedly hit for at least a few minutes or threatened with a gun or knife?	Household physical violence
8 [a]	Did you live with anyone who was a problem drinker or alcoholic or who used street drugs?	Household substance abuse
9 [a]	Was a household member depressed or mentally ill, or did a household member attempt suicide?	Household mental illness
10 [a]	Did a household member go to prison?	Incarcerated household member

[a] Dichotomous scales–yes/no.

References

1. Bick, J.; Nelson, C.A. Early Adverse Experiences and the Developing Brain. *Neuropsychopharmacology* **2016**, *41*, 177–196. [CrossRef] [PubMed]
2. Holt, S.; Buckley, H.; Whelan, S. The impact of exposure to domestic violence on children and young people: A review of the literature. *Child Abuse Negl.* **2008**, *32*, 797–810. [CrossRef]
3. Glaser, D. Child abuse and neglect and the brain—A review. *J. Child Psychol. Psychiatry* **2000**, *41*, 97–116. [CrossRef] [PubMed]
4. Teicher, M.H.; Andersen, S.L.; Polcari, A.; Anderson, C.M.; Navalta, C.P.; Kim, D.M. The neurobiological consequences of early stress and childhood maltreatment. *Neurosci. Biobehav. Rew.* **2003**, *27*, 33–44. [CrossRef]
5. Twardosz, S.; Lutzker, J.R. Child maltreatment and the developing brain: A review of neuroscience perspectives. *Aggress. Violent Behav.* **2010**, *15*, 59–68. [CrossRef]
6. Riem, M.M.E.; Alink, L.R.A.; Out, D.; Van Ijzendoorn, M.H.; Bakermans-Kranenburg, M.J. Beating the brain about abuse: Empirical and meta-analytic studies of the association between maltreatment and hippocampal volume across childhood and adolescence. *Dev. Psychopathol.* **2015**, *27*, 507–520. [CrossRef] [PubMed]
7. Van Voorhees, E.; Scarpa, A. The effects of child maltreatment on the hypothalamic-pituitary-adrenal axis. *Trauma Violence Abuse* **2004**, *5*, 333–352. [CrossRef] [PubMed]

8. Ehlert, U. Enduring psychobiological effects of childhood adversity. *Psychoneuroendocrinology* **2013**, *38*, 1850–1857. [CrossRef]
9. Wegman, H.L.; Stetler, C. A meta-analytic review of the effects of childhood abuse on medical outcomes in adulthood. *Psychosom. Med.* **2009**, *71*, 805–812. [CrossRef]
10. Fergusson, D.M.; McLeod, G.F.H.; Horwood, L.J. Childhood sexual abuse and adult developmental outcomes: Findings from a 30-year longitudinal study in New Zealand. *Child Abuse Negl.* **2013**, *37*, 664–674. [CrossRef]
11. Weber, S.; Jud, A.; Landolt, M.A. Quality of life in maltreated children and adult survivors of child maltreatment: A systematic review. *Qual. Life Res.* **2016**, *25*, 237–255. [CrossRef] [PubMed]
12. Dube, S.R.; Anda, R.F.; Felitti, V.J.; Chapman, D.P.; Williamson, D.F.; Giles, W.H. Childhood abuse, household dysfunction, and the risk of attempted suicide throughout the life span: Findings from the Adverse Childhood Experiences Study. *JAMA* **2001**, *286*, 3089–3096. [CrossRef] [PubMed]
13. Dube, S.R.; Anda, R.F.; Felitti, V.J.; Edwards, V.J.; Croft, J.B. Adverse childhood experiences and personal alcohol abuse as an adult. *Addict. Behav.* **2002**, *27*, 713–725. [CrossRef]
14. Dube, S.R.; Felitti, V.J.; Dong, M.; Chapman, D.P.; Giles, W.H.; Anda, R.F. Childhood abuse, neglect, and household dysfunction and the risk of illicit drug use: The adverse childhood experiences study. *Pediatrics* **2003**, *111*, 564–572. [CrossRef]
15. Felitti, V.J.; Anda, R.F.; Nordenberg, D.; Williamson, D.F.; Spitz, A.M.; Edwards, V.; Koss, M.P.; Marks, J.S. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults—The adverse childhood experiences (ACE) study. *Am. J. Prev. Med.* **1998**, *14*, 245–258. [CrossRef]
16. Maniglio, R. The impact of child sexual abuse on health: A systematic review of reviews. *Clin. Psychol. Rev.* **2009**, *29*, 647–657. [CrossRef] [PubMed]
17. Norman, R.E.; Byambaa, M.; De, R.; Butchart, A.; Scott, J.; Vos, T. The Long-Term Health Consequences of Child Physical Abuse, Emotional Abuse, and Neglect: A Systematic Review and Meta-Analysis. *PLoS Med.* **2012**, *9*, e1001349. [CrossRef] [PubMed]
18. Cleare, S.; Wetherall, K.; Clark, A.; Ryan, C.; Kirtley, O.J.; Smith, M.; O'Connor, R.C. Adverse Childhood Experiences and Hospital-Treated Self-Harm. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1235. [CrossRef] [PubMed]
19. DeLisi, M.; Alcala, J.; Kusow, A.; Hochstetler, A.; Heirigs, M.H.; Caudill, J.W.; Trulson, C.R.; Baglivio, M.T. Adverse Childhood Experiences, Commitment Offense, and Race/Ethnicity: Are the Effects Crime-, Race-, and Ethnicity-Specific? *Int. J. Environ. Res. Public Health* **2017**, *14*, 331. [CrossRef]
20. Ogai, Y.; Senoo, E.; Gardner, F.C.; Haraguchi, A.; Saito, T.; Morita, N.; Ikeda, K. Association between Experience of Child Abuse and Severity of Drug Addiction Measured by the Addiction Severity Index among Japanese Drug-Dependent Patients. *Int. J. Environ. Res. Public Health* **2015**, *12*, 2781–2792. [CrossRef]
21. Dong, M.; Anda, R.F.; Felitti, V.J.; Dube, S.R.; Williamson, D.F.; Thompson, T.J.; Loo, C.M.; Giles, W.H. The interrelatedness of multiple forms of childhood abuse, neglect, and household dysfunction. *Child Abuse Negl.* **2004**, *28*, 771–784. [CrossRef]
22. Berlin, L.J.; Appleyard, K.; Dodge, K.A. Intergenerational continuity in child maltreatment: Mediating mechanisms and implications for prevention. *Child Dev.* **2011**, *82*, 162–176. [CrossRef] [PubMed]
23. Schofield, T.J.; Lee, R.D.; Merrick, M.T. Safe, stable, nurturing relationships as a moderator of intergenerational continuity of child maltreatment: A meta-analysis. *J. Adolesc. Health* **2013**, *53* (Suppl. 4), S32–S38. [CrossRef] [PubMed]
24. Castro, M.; Alcántara-López, M.; Martínez, A.; Fernández, V.; Sánchez-Meca, J.; López-Soler, C. Mother's IPV, Child Maltreatment Type and the Presence of PTSD in Children and Adolescents. *Int. J. Environ. Res. Public Health* **2017**, *14*, 1077. [CrossRef]
25. Fang, X.; Brown, D.S.; Florence, C.S.; Mercy, J.A. The economic burden of child maltreatment in the United States and implications for prevention. *Child Abuse Negl.* **2012**, *36*, 156–165. [CrossRef] [PubMed]
26. Felitti, V.J.; Anda, R.F. The relationship of adverse childhood experiences to adult health, well-being, social function, and healthcare. In *The Impact of Early Life Trauma on Health and Disease: The Hidden Epidemic*; Lanius, R.A., Vermetten, E., Pain, C., Eds.; Cambridge University Press: Cambridge, UK, 2010; pp. 77–88.
27. Gilbert, R.; Widom, C.S.; Browne, K.; Fergusson, D.; Webb, E.; Janson, S. Burden and consequences of child maltreatment in high-income countries. *Lancet* **2009**, *373*, 68–81. [CrossRef]

28. WHO. *Investing in Children: The European Child Maltreatment Prevention Action Plan 2015–2020*; World Health Organization Regional Office for Europe: Copenhagen, Denmark, 2014; Available online: http://www.euro.who.int/__data/assets/pdf_file/0009/253728/64wd13e_InvestChildMaltreat_140439.pdf (accessed on 29 October 2018).
29. Sousa, C.; Herrenkohl, T.I.; Moylan, C.A.; Tajima, E.A.; Klika, J.B.; Herrenkohl, R.C.; Russo, M.J. Longitudinal study on the effects of child abuse and children's exposure to domestic violence, parent-child attachments, and antisocial behavior in adolescence. *J. Interpers. Violence* **2011**, *26*, 111–136. [CrossRef] [PubMed]
30. Fergusson, D.M.; Boden, J.M.; Horwood, L.J. From evidence to policy: Findings from the Christchurch Health and Development Study. *Aust. N. Z. J. Criminol.* **2015**, *48*, 386–408. [CrossRef]
31. Dube, S.R.; Williamson, D.F.; Thompson, T.; Felitti, V.J.; Anda, R.F. Assessing the reliability of retrospective reports of adverse childhood experiences among adult HMO members attending a primary care clinic. *Child Abuse Negl.* **2004**, *28*, 729–737. [CrossRef] [PubMed]
32. Hardt, J.; Rutter, M. Validity of adult retrospective reports of adverse childhood experiences: Review of the evidence. *J. Child Psychol. Psychiatry* **2004**, *45*, 260–273. [CrossRef] [PubMed]
33. Hardt, J.; Sidor, A.; Bracko, M.; Egle, U.T. Reliability of retrospective assessments of childhood experiences in Germany. *J. Nerv. Ment. Dis.* **2006**, *194*, 676–683. [CrossRef] [PubMed]
34. Hussey, J.M.; Chang, J.J.; Kotch, J.B. Child maltreatment in the United States: Prevalence, risk factors, and adolescent health consequences. *Pediatrics* **2006**, *118*, 933–942. [CrossRef]
35. Taillieu, T.L.; Brownridge, D.A.; Sareen, J.; Afifi, T.O. Childhood emotional maltreatment and mental disorders: Results from a nationally representative adult sample from the United States. *Child Abuse Negl.* **2016**, *59*, 1–12. [CrossRef] [PubMed]
36. Daro, D. Public Opinion and Behaviors Regarding Child Abuse Prevention: 1999 Survey. 1999. Available online: https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=182137 (accessed on 29 October 2018).
37. Larrabee-Warner, H.M.; Nielsen, E. Gallup Poll: Child Abuse Study, 1995. 2005. Available online: https://www.ndacan.acf.hhs.gov/datasets/pdfs_user_guides/Dataset106UsersGuideCodebook.pdf (accessed on 29 October 2018).
38. WHO. *Measuring and Monitoring National Prevalence of Child Maltreatment: A Practical Handbook*; World Health Organization Regional Office for Europe: Copenhagen, Denmark, 2016. Available online: http://www.euro.who.int/__data/assets/pdf_file/0003/317505/Measuring-monitoring-national-prevalence-child-maltreatment-practical-handbook.pdf?ua=1 (accessed on 29 October 2018).
39. Baker, A.J.L. Adult recall of childhood psychological maltreatment: Definitional strategies and challenges. *Child Youth Serv. Rev.* **2009**, *31*, 703–714. [CrossRef]
40. CDC. About the CDC-Kaiser ACE Study. Available online: https://www.cdc.gov/violenceprevention/acestudy/about.html (accessed on 29 October 2018).
41. CDC. About Behavioral Risk Factor Surveillance System ACE Data. 2015. Available online: https://www.cdc.gov/violenceprevention/acestudy/ace_brfss.html (accessed on 29 October 2018).
42. Hauser, W.; Schmutzer, G.; Brahler, E.; Glaesmer, H. Maltreatment in Childhood and Adolescence Results From a Survey of a Representative Sample of the German Population. *Dtsch. Arztebl. Int.* **2011**, *108*, 287–294. [CrossRef] [PubMed]
43. Hughes, K.; Lowey, H.; Quigg, Z.; Bellis, M.A. Relationships between adverse childhood experiences and adult mental well-being: Results from an English national household survey. *BMC Public Health* **2016**, *16*, 222. [CrossRef] [PubMed]
44. Kish, L. *Survey Sampling*; John Wiley & Sons: New York, NY, USA, 1965.
45. ICC/ESOMAR International Code on Market, Opinion and Social Research and Data Analysis 2016. Available online: https://www.esomar.org/uploads/public/knowledge-and-standards/codes-and-guidelines/ICCESOMAR_Code_English.pdf (accessed on 1 March 2019).
46. Anda, R.F.; Butchart, A.; Felitti, V.J.; Brown, D.W. Building a framework for global surveillance of the public health implications of adverse childhood experiences. *Am. J. Prev. Med.* **2010**, *39*, 93–98. [CrossRef] [PubMed]
47. Wingenfeld, K.; Schäfer, I.; Terfehr, K.; Grabski, H.; Driessen, M.; Grabe, H.; Löwe, B.; Spitzer, C. The reliable, valid and economic assessment of early traumatization: First psychometric characteristics of the German version of the Adverse Childhood Experiences Questionnaire (ACE). *Psychotherapie Psychosomatik Medizinische Psychologie* **2010**, *61*, 10–14. [CrossRef]

48. Vonesh, E.F.; Wang, H.; Majumdar, D. Generalized least squares, Taylor series linearization, and Fisher's scoring in multivariate nonlinear regression. *J. Am. Stat. Assoc.* **2001**, *96*, 282–291. [CrossRef]
49. EC. Special EUROBAROMETER 248—Mental Wellbeing. 2006. Available online: http://ec.europa.eu/public_opinion/archives/ebs/ebs_248_en.pdf (accessed on 29 October 2018).
50. EC. Special EUROBAROMETER 345—Mental Health. 2010. Available online: http://ec.europa.eu/health/sites/health/files/mental_health/docs/ebs_345_en.pdf (accessed on 29 October 2018).
51. Schilling, C.; Weidner, K.; Brahler, E.; Glaesmer, H.; Hauser, W.; Pohlmann, K. Patterns of Childhood Abuse and Neglect in a Representative German Population Sample. *PLoS ONE* **2016**, *11*, e0159510. [CrossRef]
52. Stoltenborgh, M.; Bakermans-Kranenburg, M.J.; Alink, L.R.A.; van IJzendoorn, M.H. The Prevalence of Child Maltreatment across the Globe: Review of a Series of Meta-Analyses. *Child Abuse Rev.* **2015**, *24*, 37–50. [CrossRef]
53. WHO. European Report on Preventing Child Maltreatment. 2013. Available online: http://www.euro.who.int/__data/assets/pdf_file/0019/217018/European-Report-on-Preventing-Child-Maltreatment.pdf (accessed on 29 October 2018).
54. Turney, K.; Wildeman, C. Adverse childhood experiences among children placed in and adopted from foster care: Evidence from a nationally representative survey. *Child Abuse Negl.* **2017**, *64*, 117–129. [CrossRef]
55. Halperin, D.S.; Bouvier, P.; Jaffe, P.D.; Mounoud, R.L.; Pawlak, C.H.; Laederach, J.; Wicky, H.R.; Astie, F. Prevalence of child sexual abuse among adolescents in Geneva: Results of a cross sectional survey. *Br. Med. J.* **1996**, *312*, 1326–1329. [CrossRef]
56. Browne, K.D. *National Prevalence Study of Child Abuse and Neglect in Romanian Families*; WHO Regional Office for Europe: Copenhagen, Denmark, 2002.
57. Sebre, S.; Sprugevica, I.; Novotni, A.; Bonevski, D.; Pakalniskiene, V.; Popescu, D.; Turchina, T.; Friedrich, W.; Lewis, O. Cross-cultural comparisons of child-reported emotional and physical abuse: Rates, risk factors and psychosocial symptoms. *Child Abuse Negl.* **2004**, *28*, 113–127. [CrossRef] [PubMed]
58. UNICEF. Child Disciplinary Practices at Home: Evidence from a Range of Low- and Middle-Income Countries. 2010. Available online: https://www.unicef.org/protection/Child_Disciplinary_Practices_at_Home.pdf (accessed on 29 October 2018).
59. Herczog, M. *Gyermekbántalmazás*; Complex Kiadó Kft.: Budapest, Hungary, 2007.
60. UNICEF. *A League Table of Child Maltreatment Deaths in Rich Nations*; UNICEF Innocenti Research Centre: Florence, Italy, 2003; Innocenti Report Card No. 5. Available online: https://www.unicef-irc.org/publications/pdf/repcard5e.pdf (accessed on 29 October 2018).
61. UNICEF. *Child Well-Being in Rich Countries: A Comparative Overview*; Innocenti Report Card 11; UNICEF Office of Research: Florence, Italy, 2013. Available online: https://www.unicef-irc.org/publications/pdf/rc11_eng.pdf (accessed on 29 October 2018).
62. Raman, S.; Hodes, D. Cultural issues in child maltreatment. *J. Paediatr. Child Health* **2012**, *48*, 30–37. [CrossRef] [PubMed]
63. CDC. Adverse Childhood Experiences (ACEs). 2016. Available online: https://www.cdc.gov/violenceprevention/acestudy/ (accessed on 29 October 2018).
64. Cronholm, P.F.; Forke, C.M.; Wade, R.; Bair-Merritt, M.H.; Davis, M.; Harkins-Schwarz, M.; Pachter, L.M.; Fein, J.A. Adverse Childhood Experiences: Expanding the Concept of Adversity. *Am. J. Prev. Med.* **2015**, *49*, 354–361. [CrossRef] [PubMed]
65. Finkelhor, D.; Shattuck, A.; Turner, H.; Hamby, S. A revised inventory of Adverse Childhood Experiences. *Child Abuse Negl.* **2015**, *48*, 13–21. [CrossRef]
66. Mersky, J.P.; Janczewski, C.E.; Topitzes, J. Rethinking the Measurement of Adversity: Moving Toward Second-Generation Research on Adverse Childhood Experiences. *Child Maltreat.* **2017**, *22*, 58–68. [CrossRef]

International Journal of
Environmental Research and Public Health

MDPI

Article

Association of Consecutive Influenza Vaccinations and Pneumonia: A Population-Based Case-Control Study

Liang-Tsai Yeh [1,2,†], Chi-Ho Chan [3,†], Shun-Fa Yang [1,4], Han-Wei Yeh [5], Ying-Tung Yeh [6,7,8], Yu-Hsun Wang [4], Ming-Chih Chou [1], Chao-Bin Yeh [1,9,10] and Ying-Hock Teng [9,10,11,*]

1 Institute of Medicine, Chung Shan Medical University, Taichung 40201, Taiwan; 68990@cch.org.tw (L.-T.Y.); ysf@csmu.edu.tw (S.-F.Y.); cshy236@csh.org.tw (M.-C.C.); sky5ff@gmail.com (C.-B.Y.)
2 Department of Anesthesiology, Changhua Christian Hospital, Changhua 50006, Taiwan
3 Department of Microbiology and Immunology, Chung Shan Medical University, Taichung 40201, Taiwan; chiho@csmu.edu.tw
4 Department of Medical Research, Chung Shan Medical University Hospital, Taichung 40201, Taiwan; cshe731@csh.org.tw
5 School of Medicine, Chang Gung University, Taoyuan City 33302, Taiwan; george66889@gmail.com
6 Graduate Institute of Oral Sciences, Chung Shan Medical University, Taichung 40201, Taiwan; yehtungtung@hotmail.com
7 School of Dentistry, Chung Shan Medical University, Taichung 40201, Taiwan
8 Department of Dentistry, Chung Shan Medical University Hospital, Taichung 40201, Taiwan
9 Department of Emergency Medicine, School of Medicine, Chung Shan Medical University, Taichung 40201, Taiwan
10 Department of Emergency Medicine, Chung Shan Medical University Hospital, Taichung 40201, Taiwan
11 Department of Health Policy and Management, Chung Shan Medical University, Taichung 40201, Taiwan
* Correspondence: cshy392@csh.org.tw
† These authors contributed equally to the work.

Received: 14 February 2019; Accepted: 23 March 2019; Published: 26 March 2019

Abstract: The purpose of this study was to investigate whether individuals receiving influenza vaccines have a lower risk of pneumonia. A nationwide population-based case-control study was conducted using data from the National Health Insurance Research Database in Taiwan. We enrolled 7565 patients each in pneumonia and non-pneumonia groups after diagnosis of patients with chronic pulmonary disease, and these patients were individually age and sex matched in a 1:1 ratio. Using conditional logistic regression analysis, adjusted odds ratios (aORs) were estimated in patients who received influenza vaccination and those who had not previously had pneumonia. Moreover, we also analyzed the interval between vaccination and the onset of pneumonia and the number of vaccinations received by patients. This was compared with patients who never received influenza vaccination. Patients who had received influenza vaccination and had been vaccinated for two consecutive years (aOR = 0.85, confidence interval (CI) = 0.79–0.93 and aOR = 0.75, CI = 0.67–0.85, respectively) showed lower rates of pneumonia occurrence by 15–25%. In conclusion, influenza vaccination significantly reduces the occurrence of pneumonia, especially in individuals who receive vaccination in consecutive years.

Keywords: influenza vaccination; pneumonia; population-based case-control study

1. Introduction

Pneumonia is a common clinical presentation after a respiratory infection. Since 2016, pneumonia has been the third most common cause of death in Taiwan. The majority of pneumonia cases can be classified as either community-acquired, hospital-acquired, or acquired after traveling to foreign countries. Bacterial pneumonia is usually a complication of influenza virus infection [1]. The American Thoracic Society and Infectious Diseases Society of America have classified pneumonia into three types, namely community-acquired pneumonia (CAP), hospital-acquired (or nosocomial) pneumonia, and ventilator-associated pneumonia, according to the epidemiology, pathogenesis, and risk factors for infection in patients with pneumonia [2,3]. The etiology of pneumonia includes bacteria, viruses, fungi, and protozoa. Generally, pathogens that potentially cause pneumonia subsist of "typical" bacteria and "atypical" organisms, including *Mycoplasma pneumoniae*, and respiratory viruses, such as influenza viruses. However, 11–20% of pneumonia cases are polymicrobial, and the etiology usually consists of a combination of typical and atypical pathogens [4]. *Streptococcus pneumoniae* has been reported as the major pathogen in secondary infection after individuals were infected by influenza viruses, and it increases the mortality risk of patients [5–8]. Previous studies such as Tessmer et al. [9] show that prior influenza vaccination is associated with a less severe clinical course and improved long-term survival in patients with CAP, especially during the influenza epidemic season. However, Shinjoh et al. [10] reported that children who were immunized for two consecutive seasons experienced decreased vaccine effectiveness and were more likely to acquire influenza and that this might be associated with immunity against influenza infection in the previous season. Therefore, the aims of this study are also to clarify this.

Among the pathogens, influenza is one of the primary causes of pneumonia and influential respiratory diseases [11,12]. In general, human influenza viruses include A/H1N1, A/H3N2, and B viruses. Because of the antigenic drift of viruses, epidemics of influenza are reported every year in Taiwan. To effectively prevent the disease, large influenza vaccination programs have been held by the health authority of the Taiwanese government, and the trivalent influenza vaccine containing influenza A/H1N1, A/H3N2, and B viruses is employed annually. To increase the influenza vaccination coverage, free-of-charge vaccination programs have been held since 1988 for several groups, including young children under 6 months old, students from primary school to senior high school, and adults with high influenza risk (i.e., people aged over 50 years and patients of any age with a chronic illness like diabetes, chronic hepatitis, cardiovascular disease, chronic pulmonary disease, chronic renal disease, etc.) [13]. In fact, a characteristic of influenza infection in the elderly is a high frequency of pneumonia complications. Therefore, inoculation with the influenza vaccine is critical for both preventing influenza infection and lowering the risk of post-influenza pneumonia development in the elderly [14]. Influenza vaccination has been reported to be associated not only with prevention of influenza epidemics but also reduced risk of several diseases, such as acute kidney injury, diabetes, cardiovascular disease, and respiratory failure in chronic obstructive pulmonary disease, especially in elderly people [15–20]. However, the relationship between pneumonia and influenza vaccination remains unclear. This study investigated whether individuals receiving influenza vaccines have decreased risk of developing pneumonia.

2. Materials and Methods

2.1. Data Source

We used the National Health Insurance Research Database (NHIRD) of Taiwan from 1 January 2009 to 31 December 2013. The NHIRD contains all medical data, including information on disease diagnosis, drug prescriptions, medical operations, and medical expenses of insurance claims data. The National Health Insurance covered more than 99% of Taiwan's population by 2010. The Longitudinal Health Insurance Database 2010 included data from 1 million people randomly sampled from the NHIRD and who were registered by the end of 2010. Diagnostic codes were recorded according to the International Classification of Diseases, Ninth Edition, Clinical Modification (ICD-9-CM). This study was approved by the ethical review board of Chung Shan Medical University Hospital (CSMU No. CS 18096).

2.2. Study Groups

This study used a case-control study design. All subjects had been diagnosed with chronic pulmonary diseases (ICD-9-CM 490–496) from 2010 to 2012. To confirm the accuracy of diagnosis, we enrolled patients with a minimum of two outpatient visits or one admission. The case group was defined as newly diagnosed pneumonia (ICD-9-CM 481, 482, 483, 485, and 486) from emergency or hospitalization after chronic pulmonary disease diagnosis. The first date of pneumonia was the index date. Patients with no pneumonia diagnosis from 2009 to 2013 were selected as the control group. We performed a 1:1 age and sex match to obtain index dates corresponding to the control group.

2.3. Exposure Measurement

Influenza vaccination was defined using an ICD-9-CM (V04.7 or V04.8) diagnosis or a drug code within two years before the index date. Comorbidities considered were hypertension (ICD-9-CM 401–405), diabetes mellitus (ICD-9-CM 250), cerebrovascular disease (ICD-9-CM 430–438), renal disease (ICD-9-CM 582–582.9, 583–583.7, 585, 586, and 588–588.9), liver disease (ICD-9-CM 456.0–456.21, 571.2, 571.4–571.49, 571.5, 571.6, and 572.2–572.8), ischemic heart disease (ICD-9-CM 410–414), dementia (ICD-9-CM 290.0–290.4, 294.1, 331.0–331.2, and 331.82), alcohol-related disorder (ICD-9-CM 291, 303, 305.0, 571.0, 571.1, 571.3, 790.3, and V11.3), and malignancy (ICD-9-CM 140–208). All patients were diagnosed with the disease before the index date.

2.4. Statistical Analysis

A comparison of pneumonia and non-pneumonia groups was performed using a Chi-squared test or independent *t*-test, as was appropriate. The association between influenza vaccination and pneumonia was analyzed using conditional logistical regression analysis. The statistical software used was SPSS V.18.0 (SPSS, Chicago, IL, USA), and the significance was defined as $p < 0.05$. On the use of post hoc power analysis of the logistic regression, the alpha error was 0.05 and the beta error was 0.2. The achieved power of the study was 0.83.

3. Results

In total, 90,491 patients diagnosed with chronic pulmonary disease were initially recruited from 2010 to 2012. Next, 8010 of the 90,491 patients were selected according to a new diagnosis of pneumonia after chronic pulmonary disease diagnosis. Another 56,233 patients who were never diagnosed with pneumonia were also selected. After this, we performed a 1:1 age and sex match, and 7565 patients were enrolled in each of the two groups: pneumonia and non-pneumonia (Figure 1). The distributions of age and sex in the groups were not significant (Table 1).

Table 1. Demographic characteristics of pneumonia and non-pneumonia.

Variable	Pneumonia (N = 7565)		Non-Pneumonia (N = 7565)		p-Value
	n	%	n	%	
Vaccination					0.026 *
No	5006	66.2	4876	64.5	
Yes	2559	33.8	2689	35.5	
Age					1
<40	1739	23.0	1739	23.0	
40–65	1383	18.3	1383	18.3	
≥65	4443	58.7	4443	58.7	
Mean ± SD	59.1 ± 29.1		59.1 ± 29.1		1

Table 1. *Cont.*

Variable	Pneumonia (N = 7565)		Non-Pneumonia (N = 7565)		p-Value
	n	%	n	%	
Gender					1
Female	2991	39.5	2991	39.5	
Male	4574	60.5	4574	60.5	
Hypertension	3694	48.8	3173	41.9	<0.001 **
Diabetes	1673	22.1	1116	14.8	<0.001 **
Cerebrovascular disease	1425	18.8	693	9.2	<0.001 **
Renal disease	591	7.8	301	4.0	<0.001 **
Liver disease	385	5.1	298	3.9	0.001 **
Ischemic heart disease	1435	19.0	1111	14.7	<0.001 **
Dementia	665	8.8	262	3.5	<0.001 **
Alcohol-related disorder	58	0.8	15	0.2	<0.001 **
Malignancy	1009	13.3	406	5.4	<0.001 **

* $p < 0.05$, ** $p < 0.01$.

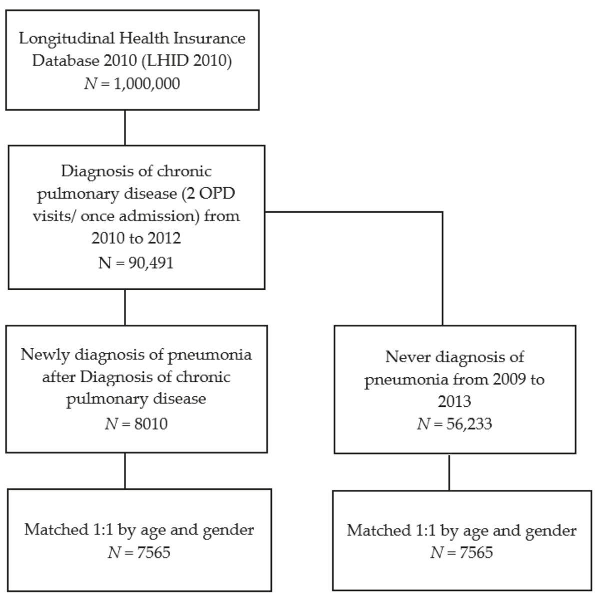


Figure 1. Flow chart for patient selection.

Next, we analyzed the risk of pneumonia and the selected comorbidities. Patients who had received an influenza vaccination showed a 15% decreased risk of pneumonia (adjusted odds ratio (aOR) = 0.85, confidence interval (CI) = 0.79–0.93). Furthermore, patients had an increased risk of pneumonia with following comorbidities: hypertension, diabetes, cerebrovascular disease, renal disease, and liver disease (Table 2).

Table 2. Conditional logistic regression between influenza vaccination and comorbidities and the risk of pneumonia.

Variable	Crude OR	95% CI	Adjusted OR †	95% CI
Vaccination				
No	1		1	
Yes	0.91	** 0.84–0.98	0.85	** 0.79–0.93
Hypertension	1.48	** 1.37–1.60	1.26	** 1.16–1.37
Diabetes	1.73	** 1.58–1.89	1.46	** 1.33–1.62
Cerebrovascular disease	2.48	** 2.24–2.76	2.15	** 1.92–2.41
Renal disease	2.06	** 1.78–2.38	1.64	** 1.41–1.92
Liver disease	1.31	** 1.12–1.54	1.11	0.94–1.32
Ischemic heart disease	1.40	** 1.28–1.53	1.35	** 1.22–1.49
Dementia	2.95	** 2.52–3.45	2.72	** 2.3–3.23
Alcohol-related disorder	4.07	** 2.27–7.31	3.68	** 1.99–6.82
Malignancy	2.88	** 2.54–3.27	3.10	** 2.71–3.55

† Adjusted for hypertension, diabetes, cerebrovascular disease, renal disease, liver disease, ischemic heart disease, dementia, alcohol-related disorder, and malignancy. OR: odds ratio; CI: confidence interval; ** $p < 0.01$.

We also analyzed the interval between the onset of pneumonia and receiving the influenza vaccination, as well as the number of vaccinations. We compared these results with patients who had never received an influenza vaccination. Patients that received the vaccine one year prior to the study showed a 13% reduction in the risk of developing pneumonia (aOR = 0.87, CI = 0.78–0.98). Patients vaccinated for two consecutive years prior to the study showed a 25% decreased risk of developing pneumonia (aOR = 0.75, CI = 0.67–0.85). Furthermore, patients vaccinated for three consecutive years compared with those who had never received influenza vaccination showed a 44% decreased risk of developing pneumonia (aOR = 0.56, CI = 0.45–0.69) (Table 3).

Table 3. Conditional logistic regression of the frequency of receiving an influenza vaccination.

Variable		*N*	No. of Pneumonia	Crude OR	95% C.I.	Adjusted OR †	95% C.I.
Vaccination							
1 yr	2 yr						
No	No	9882	5006	1		1	
No	Yes	1282	665	1.03	0.91–1.17	0.98	0.86–1.12
Yes	No	1817	908	0.96	0.86–1.07	0.87 *	0.78–0.98
Yes	Yes	2149	986	0.79 **	0.71–0.88	0.75 **	0.67–0.85
Vaccination							
Never		9243	4685	1		1	
Three consecutive years		1096	475	0.64 **	0.53–0.77	0.56 **	0.45–0.69

† Adjusted for hypertension, diabetes, cerebrovascular disease, renal disease, liver disease, ischemic heart disease, dementia, alcohol-related disorder, and malignancy. 1 yr: receive influenza vaccine in the first year before the pneumonia diagnosis. 2 yr: receive influenza vaccine in 1 to 2 years before the pneumonia diagnosis. * $p < 0.05$, ** $p < 0.01$.

Finally, we performed a subgroup analysis by age and sex for the vaccination and non-vaccination groups. After we subdivided the patients by age into three subgroups (<40, 40–65, and ≥65), we observed that only those patients aged ≥ 65 had an obvious reduced risk of developing pneumonia (aOR = 0.78, CI = 0.71–0.86). Male patients also showed a lower risk of pneumonia than the female patients. However, vaccination did not reduce the risk of pneumonia in patients aged 40–65 years (Table 4).

Table 4. Subgroup analysis of the conditional logistic regression between vaccination and non-vaccination groups.

Variable	Vaccination		Non-Vaccination		OR	95% CI
	N	No. of Pneumonia Event	*N*	No. of Pneumonia Event		
Age [a]						
<40	513	271	2965	1468	1.17	0.94–1.46
40–65	219	143	2547	1240	1.69 **	1.23–2.33
≥65	4516	2145	4370	2298	0.78 **	0.71–0.86
Gender [b]						
Female	1869	943	4113	2048	0.99	0.87–1.13
Male	3379	1616	5769	2958	0.78 **	0.7–0.87

[a] Adjusted for hypertension, diabetes, cerebrovascular disease, renal disease, liver disease, ischemic heart disease, and malignancy. [b] Adjusted for hypertension, diabetes, cerebrovascular disease, renal disease, liver disease, ischemic heart disease, dementia, alcohol-related disorder, and malignancy. ** $p < 0.01$.

4. Discussion

Patients diagnosed with chronic pulmonary disease were enrolled to analyze the association between influenza vaccination and the prevention of pneumonia. Our study indicated that patients who received vaccination one year prior to the study had a significantly decreased risk of developing pneumonia. Additionally, patients who received vaccination consecutively for two and three years showed a continuous reduction of their risk of pneumonia. Although the patients who received influenza vaccinations only in the second year did not show any decrease in their risk of developing pneumonia, the data was not statistically significant. Repeated influenza vaccination was demonstrated to be effective for preventing severe and fatal influenza infection in elderly individuals [17]. Our study showed that influenza vaccination can lower the risk of pneumonia in individuals aged ≥ 65. A previous study used health administrative databases to investigate the association between influenza vaccination and the all-cause death of elderly people >65 years. They also concluded that influenza vaccination was associated with reductions in the total hospitalizations for pneumonia and influenza and all-cause mortality during the influenza season [21]. Moreover, some studies, such as Li et al., confirmed that previous pneumococcal and influenza vaccination in elderly patients reduced the length of hospital stay and reduced the risk of bacteremia [22], which was similar to our study. However, Demirdogen Cetinoglu et al. showed that influenza vaccination did not affect the clinical outcome of hospitalized adult CAP patients. This result may be related to the low influenza vaccination rate in the elderly [23]. Therefore, further research is needed to analyze this in the future.

Several studies have shown that influenza vaccination not only prevents influenza infection but also reduces the risk of several diseases in certain groups of patients. According to a retrospective cohort study, the administration of the influenza vaccination in elderly patients with diabetes reduced risks of hospitalization, lung failure, and 12-month mortality [15]. Influenza vaccination in elderly individuals reduced the risk of acute kidney injury in a nested-control study. Although the actual mechanism is unclear, influenza vaccination is proposed to be associated with the reduction of the inflammation cascade [19]. A study showed that influenza vaccination reduced dementia risk in patients with chronic kidney disease [24]. Another study also showed that influenza vaccines might prevent cardiovascular disease [16]. A high-dose influenza vaccine (60 μg of hemagglutinin from each of the three viral strains against 15 μg of hemagglutinin from each of the four viral strains) might be more effective in reducing poor clinical outcomes in patients who have a heart failure or myocardial infarction history [20]. In another study, it was found that influenza infection increased the risk of atrial fibrillation, and people who received influenza vaccination showed a lower atrial fibrillation risk [18]. Influenza vaccination could decrease respiratory failure risk in patients with chronic obstructive pulmonary disease [13]. Our study found that influenza vaccination could significantly decrease the risk of pneumonia.

One of the most serious complications of influenza infection is bacterial pneumonia, which increases morbidity and mortality [25]. Evidence shows that older patients, or those with severe illness with pneumonia, had higher 30-day mortality rates [5]. In general, influenza viruses cause only the desquamation of the epithelial cells in the respiratory tract, but this disrupts the outermost part of the mucosal defense and promotes secondary bacterial pneumonia. The most common pathogens associated with secondary infection are *S. pneumoniae*, *Staphylococcus aureus*, and *Haemophilus influenzae* [26]. In Taiwan, to prevent *S. pneumoniae* infection in elderly individuals, the government has provided free vaccination of the 23-valent pneumococcal polysaccharide vaccine for people aged > 75 since 2008 [27]. We also considered that one of the factors for lowering pneumonia risk in our enrolled patients might be *S. pneumoniae* vaccination. However, only 257 (1.7%) patients in our study population had received *S. pneumoniae* vaccine. Therefore, the observation that influenza vaccination could reduce the risk of pneumonia in elderly individuals may not be overestimated.

The prevalent strains of influenza viruses in Taiwan are A/H1N1, A/H3N2, and B viruses. For the prevention of an influenza outbreak, the health authority of the Taiwanese government recommends trivalent influenza vaccine use. However, because of genetic divergence, influenza B viruses are divided into two lineages: Victoria and Yamagata lineages. Simultaneous cocirculation of these lineages has been observed, and a mismatch of the vaccine lineage and circulating strains has been reported [28–30]. This might decrease the effectiveness of influenza vaccines. However, there was no mismatch of the vaccine strains and the circulating strains in our study during the observation period (2010–2013).

Our study has several limitations. First, in Taiwan, influenza vaccines are purchased from different pharmaceutical companies. The manufacturing process of influenza vaccines varies between pharmaceutical companies. For instance, the inactivated influenza vaccine (i.e., Optaflu©) produced by Novartis contains whole virus, whereas the vaccine produced by Sanofi Pasteur (i.e., Fluzone©) contains split virus. We did not examine whether the protective effect of both these influenza vaccines is the same. In addition, the protective effect may be different for different age groups. Second, we were unable to obtain potentially relevant personal behavioral information, such as alcohol consumption, smoking habits, and body mass index. These confounding factors may have affected the outcome. Third, Taiwan's National Health Insurance system involves the Taiwanese population. Our data accurately reflects the situation in Taiwan; however, our results may not be applicable to other regions. Fourth, the National Health Insurance Research Database (NHIRD) consists of claim data. The database used does not contain information about clinical parameters, such as pneumonia infection severity. Therefore, we could not further distinguish the severities of respiratory diseases. Fifth, the laboratory data and microbial culture data that may affect the occurrence of pneumonia infection are not included in the database. Moreover, information about from sputum cultures or viral swabs and the causative agents of pneumonia was not available. Thus, the cause of pneumonia identified could not be divided into viral or bacterial.

5. Conclusions

In conclusion, patients who receive influenza vaccination have a significantly lower risk of developing pneumonia, especially for the elderly (aged $\geq$ 65). The preventative effects against pneumonia depend on consecutive years of vaccination.

Author Contributions: Conceptualization: Y.-H.T., L.-T.Y., C.-H.C., and C.-B.Y. Validation: S.-F.Y. and C.-B.Y. Formal analysis: H.-W.Y., Y.-T.Y., and Y.-H.W. Writing—Original draft preparation: Y.-H.T., C.-H.C., M.-C.C., and C.-B.Y. Writing—Review and editing: Y.-H.T., L.-T.Y., C.-H.C., and C.-B.Y.

Funding: This research received no external funding.

Acknowledgments: This study was partly based on data from the NHIRD provided by the NHI Administration, Ministry of Health and Welfare, and managed by the National Health Research Institutes (registration number: NHIRD-104-148). The interpretation and conclusions contained herein do not represent those of the NHI Administration, Ministry of Health and Welfare, or National Health Research Institutes. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cheng, C.W.; Chien, M.H.; Su, S.C.; Yang, S.F. New markers in pneumonia. *Clin. Chim. Acta Int. J. Clin. Chem.* **2013**, *419*, 19–25.
2. Anand, N.; Kollef, M.H. The alphabet soup of pneumonia: CAP, HAP, HCAP, NHAP, and VAP. *Semin. Respir. Crit. Care Med.* **2009**, *30*, 3–9. [CrossRef] [PubMed]
3. American Thoracic Society, & Infectious Diseases Society of America. Guidelines for the management of adults with hospital-acquired, ventilator-associated, and healthcare-associated pneumonia. *Am. J. Respir. Crit. Care Med.* **2005**, *171*, 388–416. [CrossRef] [PubMed]
4. Jereb, M.; Kotar, T. Usefulness of procalcitonin to differentiate typical from atypical community-acquired pneumonia. *Wien. Klin. Wochenschr.* **2006**, *118*, 170–174. [CrossRef]
5. Chang, G.M.; Tung, Y.C. Factors associated with pneumonia outcomes: A nationwide population-based study over the 1997–2008 period. *J. Gen. Intern. Med.* **2012**, *27*, 527–533. [CrossRef]
6. Chiu, C.Y.; Chen, C.J.; Wong, K.S.; Tsai, M.H.; Chiu, C.H.; Huang, Y.C. Impact of bacterial and viral coinfection on mycoplasmal pneumonia in childhood community-acquired pneumonia. *J. Microbiol. Immunol. Infect. = Wei Mian Yu Gan Ran Za Zhi* **2015**, *48*, 51–56. [CrossRef] [PubMed]
7. Cordoba, E.; Maduro, G.; Huynh, M.; Varma, J.K.; Vora, N.M. Deaths from pneumonia-New York City, 1999–2015. *Open Forum Infect. Dis.* **2018**, *5*, ofy020. [CrossRef] [PubMed]
8. Huang, C.Y.; Chang, L.; Liu, C.C.; Huang, Y.C.; Chang, L.Y.; Huang, Y.C.; Chiu, N.C.; Lin, H.C.; Ho, Y.H.; Chi, H.; et al. Risk factors of progressive community-acquired pneumonia in hospitalized children: A prospective study. *J. Microbiol. Immunol. Infect. = Wei Mian Yu Gan Ran Za Zhi* **2015**, *48*, 36–42. [CrossRef] [PubMed]
9. Tessmer, A.; Welte, T.; Schmidt-Ott, R.; Eberle, S.; Barten, G.; Suttorp, N.; Schaberg, T. Influenza vaccination is associated with reduced severity of community-acquired pneumonia. *Eur. Respir. J.* **2011**, *38*, 147–153. [CrossRef]
10. Shinjoh, M.; Sugaya, N.; Yamaguchi, Y.; Iibuchi, N.; Kamimaki, I.; Goto, A.; Kobayashi, H.; Kobayashi, Y.; Shibata, M.; Tamaoka, S.; et al. Inactivated influenza vaccine effectiveness and an analysis of repeated vaccination for children during the 2016/17 season. *Vaccine* **2018**, *36*, 5510–5518. [CrossRef]
11. Zhang, X.; Zhang, J.; Chen, L.; Feng, L.; Yu, H.; Zhao, G.; Zhang, T. Pneumonia and influenza hospitalizations among children under 5 years of age in Suzhou, China, 2005–2011. *Influenza Other Respir. Viruses* **2017**, *11*, 15–22. [CrossRef] [PubMed]
12. Almond, M.H.; McAuley, D.F.; Wise, M.P.; Griffiths, M.J. Influenza-related pneumonia. *Clin. Med.* **2012**, *12*, 67–70.
13. Huang, H.H.; Chen, S.J.; Chao, T.F.; Liu, C.J.; Chen, T.J.; Chou, P.; Wang, F.D. Influenza vaccination and risk of respiratory failure in patients with chronic obstructive pulmonary disease: A nationwide population-based case-cohort study. *J. Microbiol. Immunol. Infect.* **2019**, *52*, 22–29. [CrossRef] [PubMed]
14. Heo, J.Y.; Song, J.Y.; Noh, J.Y.; Choi, M.J.; Yoon, J.G.; Lee, S.N.; Cheong, H.J.; Kim, W.J. Effects of influenza immunization on pneumonia in the elderly. *Hum. Vaccines Immunother.* **2018**, *14*, 744–749. [CrossRef]
15. Wang, I.K.; Lin, C.L.; Chang, Y.C.; Lin, P.C.; Liang, C.C.; Liu, Y.L.; Chang, C.T.; Yen, T.H.; Huang, C.C.; Sung, F.C. Effectiveness of influenza vaccination in elderly diabetic patients: A retrospective cohort study. *Vaccine* **2013**, *31*, 718–724. [CrossRef] [PubMed]
16. Clar, C.; Oseni, Z.; Flowers, N.; Keshtkar-Jahromi, M.; Rees, K. Influenza vaccines for preventing cardiovascular disease. *Cochrane Database Syst. Rev.* **2015**, CD005050. [CrossRef]
17. Casado, I.; Dominguez, A.; Toledo, D.; Chamorro, J.; Astray, J.; Egurrola, M.; Fernandez-Sierra, M.A.; Martin, V.; Morales-Suarez-Varela, M.; Godoy, P.; et al. Repeated influenza vaccination for preventing severe and fatal influenza infection in older adults: A multicentre case-control study. *Can. Med. Assoc. J.* **2018**, *190*, E3–E12. [CrossRef]
18. Chang, T.Y.; Chao, T.F.; Liu, C.J.; Chen, S.J.; Chung, F.P.; Liao, J.N.; Tuan, T.C.; Chen, T.J.; Chen, S.A. The association between influenza infection, vaccination, and atrial fibrillation: A nationwide case-control study. *Heart Rhythm* **2016**, *13*, 1189–1194. [CrossRef]

19. Shih, C.H.; Lee, Y.J.; Chao, P.W.; Kuo, S.C.; Ou, S.M.; Huang, H.M.; Chen, Y.T. Association between influenza vaccination and the reduced risk of acute kidney injury among older people: A nested case-control study. *Eur. J. Intern. Med.* **2018**, *54*, 65–69. [CrossRef]
20. Vardeny, O.; Udell, J.A.; Joseph, J.; Farkouh, M.E.; Hernandez, A.F.; McGeer, A.J.; Talbot, H.K.; Bhatt, D.L.; Cannon, C.P.; Goodman, S.G.; et al. High-dose influenza vaccine to reduce clinical outcomes in high-risk cardiovascular patients: Rationale and design of the invested trial. *Am. Heart J.* **2018**, *202*, 97–103. [CrossRef]
21. Wong, K.; Campitelli, M.A.; Stukel, T.A.; Kwong, J.C. Estimating influenza vaccine effectiveness in community-dwelling elderly patients using the instrumental variable analysis method. *Arch. Intern. Med.* **2012**, *172*, 484–491. [CrossRef]
22. Li, C.; Gubbins, P.O.; Chen, G.J. Prior pneumococcal and influenza vaccinations and in-hospital outcomes for community-acquired pneumonia in elderly veterans. *J. Hosp. Med.* **2015**, *10*, 287–293. [CrossRef]
23. Demirdogen Cetinoglu, E.; Uzaslan, E.; Sayiner, A.; Cilli, A.; Kilinc, O.; Sakar Coskun, A.; Hazar, A.; Kokturk, N.; Filiz, A.; Polatli, M. Pneumococcal and influenza vaccination status of hospitalized adults with community acquired pneumonia and the effects of vaccination on clinical presentation. *Hum. Vaccines Immunother.* **2017**, *13*, 2072–2077. [CrossRef]
24. Liu, J.C.; Hsu, Y.P.; Kao, P.F.; Hao, W.R.; Liu, S.H.; Lin, C.F.; Sung, L.C.; Wu, S.Y. Influenza vaccination reduces dementia risk in chronic kidney disease patients: A population-based cohort study. *Medicine* **2016**, *95*, e2868. [CrossRef]
25. Cauley, L.S.; Vella, A.T. Why is coinfection with influenza virus and bacteria so difficult to control? *Discov. Med.* **2015**, *19*, 33–40. [PubMed]
26. Joseph, C.; Togawa, Y.; Shindo, N. Bacterial and viral infections associated with influenza. *Influenza Other Respir. Viruses* **2013**, *7* (Suppl. 2), 105–113. [CrossRef]
27. Tsai, Y.H.; Hsieh, M.J.; Chang, C.J.; Wen, Y.W.; Hu, H.C.; Chao, Y.N.; Huang, Y.C.; Yang, C.T.; Huang, C.C. The 23-valent pneumococcal polysaccharide vaccine is effective in elderly adults over 75 years old—Taiwan's PPV vaccination program. *Vaccine* **2015**, *33*, 2897–2902. [CrossRef]
28. Chen, G.W.; Shih, S.R.; Hsiao, M.R.; Chang, S.C.; Lin, S.H.; Sun, C.F.; Tsao, K.C. Multiple genotypes of influenza b viruses cocirculated in taiwan in 2004 and 2005. *J. Clin. Microbiol.* **2007**, *45*, 1515–1522. [CrossRef] [PubMed]
29. Alfelali, M.; Khandaker, G.; Booy, R.; Rashid, H. Mismatching between circulating strains and vaccine strains of influenza: Effect on hajj pilgrims from both hemispheres. *Hum. Vaccines Immunother.* **2016**, *12*, 709–715. [CrossRef]
30. Noh, J.Y.; Choi, W.S.; Song, J.Y.; Lee, H.S.; Lim, S.; Lee, J.; Seo, Y.B.; Lee, J.S.; Wie, S.H.; Jeong, H.W.; et al. Significant circulation of influenza b viruses mismatching the recommended vaccine-lineage in South Korea, 2007–2014. *Vaccine* **2018**, *36*, 5304–5308. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

Increased Prevalence of Psychosocial, Behavioral, and Socio-Environmental Risk Factors among Overweight and Obese Youths in Mexico and the United States

Yvonne N. Flores [1,2], Zuelma A. Contreras [3], Paula Ramírez-Palacios [1,*], Leo S. Morales [4], Todd C. Edwards [4], Katia Gallegos-Carrillo [1], Jorge Salmerón [5], Cathy M. Lang [6], Noémie Sportiche [1] and Donald L. Patrick [4]

1 Unidad de Investigación Epidemiológica y en Servicios de Salud, Delegación Morelos, Instituto Mexicano del Seguro Social, Cuernavaca, Morelos 62000, Mexico; ynflores@ucla.edu (Y.N.F.); kgallegosc13@gmail.com (K.G.-C.); nsportiche@gmail.com (N.S.)
2 UCLA Department of Health Policy and Management, Center for Cancer Prevention and Control Research, Fielding School of Public Health and Jonsson Comprehensive Cancer Center, Los Angeles, CA 90095, USA
3 UCLA Department of Epidemiology, Fielding School of Public Health, Los Angeles, CA 90095, USA; zuelmaarellano@ucla.edu
4 Department of Health Services, School of Public Health, University of Washington, Seattle, WA 98195, USA; lsm2010@uw.edu (L.S.M.); toddce.uw@gmail.com (T.C.E.); donald@uw.edu (D.L.P.)
5 Universidad Nacional Autónoma de México, Centro de Investigación en Políticas, Población y Salud. Ciudad Universitaria, Ciudad de México 04510, Mexico; jorge.salmec@gmail.com
6 UCLA Department of Community Health Sciences, Fielding School of Public Health, Los Angeles, CA 90095, USA; clang@ucla.edu
* Correspondence: paula_rzps@hotmail.com; Tel.: +52-777-100-1364

Received: 29 January 2019; Accepted: 27 April 2019; Published: 30 April 2019

Abstract: The aim of this study was to examine various psychosocial, behavioral, and socio-environmental factors in a multiethnic sample of healthy-weight, overweight, and obese youths in the United States (US) and Mexico and determine differences by sex. We conducted a cross-sectional analysis of 633 youths aged 11–18 years who completed a self-reported questionnaire. Height and weight were measured to determine body mass index (BMI). Overweight and obese youth in both countries were significantly more likely to report a higher body image dissatisfaction (Odds Ratio (OR) = 1.67 and OR= 2.95, respectively), depressive symptoms (OR = 1.08 and OR = 1.12, respectively), perceive themselves as overweight (OR = 2.57) or obese (OR = 5.30), and a lower weight-specific quality of life (OR = 0.97 and OR = 0.95, respectively) than healthy-weight youth. Obese youth have lower healthy lifestyle priorities (OR = 0.75) and are less likely to be physically active (OR = 0.79) and eat breakfast (OR = 0.47) than healthy-weight youth. Additionally, overweight and obese youth are more likely to engage in weight control behaviors (OR = 5.19 and OR = 8.88, respectively) and restrained eating than healthy-weight youth. All the aforementioned results had a p-value of <0.05, which was considered statistically significant. The association between these factors and overweight or obesity remained significant after controlling for age, sex, race/ethnicity, and country. In conclusion, obesity was associated with a range of psychosocial, behavioral, and socio-environmental risk factors in both countries. Our findings support the need for multifactorial approaches when developing interventions to address the growing problem of obesity among youth in the US and Mexico.

Keywords: obesity; quality of life; adolescent; risk factors; psychosocial; socio-environmental; behavior; United States; Mexico; Latinos

1. Introduction

The high prevalence of obesity among youth is one of the most concerning public health issues in both developed and developing countries [1]. Most overweight or obese children live in developing countries, where the rate of increase is over 30% higher than in more developed countries [1]. In the past 40 years, the number of obese children and adolescents (aged five to 19 years) has increased from 11 million in 1975 to 124 million in 2016 [2]. If current trends continue, by 2022, there will be more obese children and adolescents worldwide than moderately or severely underweight children [2]. Childhood and adolescent overweight and obesity are associated with increased risk of subsequent diabetes, stroke, coronary heart disease, hypertension, functional disability, as well as premature adult mortality and morbidity [3,4]. In the United States (US), an estimated 34.5% of adolescents aged 12–19 years were overweight or obese from 2011–2012, and of these, 16.9% were obese [5]. This number rose to 20.6% in 2015–2016 [6]. Disparities in obesity and overweight exist across racial and ethnic groups in the US, with African American and Mexican-American adolescents ranking highest in prevalence [7,8]. During 2011–2012, the prevalence of overweight and obesity was 39.8% and 38.1% among African American and Latino adolescents, respectively, followed by non-Latino white (31.2%) and Asian (24.6%) adolescents [5]. From 2013–2016, the prevalence of obesity among youth of Mexican origin aged 12–19 years was 26.6%, as compared to 17.2% among non-Latino whites [9]. Studies have also shown that US-born Mexicans are significantly more likely to be overweight or obese than Mexican immigrants [10,11]. In Mexico, the prevalence of obesity and overweight among adolescents aged 12–19 years was 36.3% in 2016 [12].

Addressing obesity is complex, due to its multi-causal nature that includes various psychosocial, behavioral, and socio-environmental factors. Previous studies have found an association between a range of psychosocial factors and increased obesity risk, such as body size dissatisfaction and self-perception of overweight, because they may promote unhealthy weight control behaviors [13–15]. Other psychosocial factors that have been examined include exposure to adverse life events and the influence of the family and peer environment, which may be associated with a greater risk of childhood overweight/obesity [16]. Studies have also found that depressive symptoms are a risk factor for obesity because binge eating may be used as a coping mechanism [14,17]. Obese youth report having a lower quality of life (QOL) [18–20], which improves when they lose weight [21].

Unhealthy weight control and dietary restraint behaviors have been found to predict the onset of obesity [14,15,17,22]. Studies also show that prioritizing healthy eating may protect youth from becoming overweight or obese [23,24], whereas prioritizing physical activity appears to be less protective [23]. Various studies have demonstrated a negative association between breakfast consumption and an increase in body mass index (BMI) [14,22,25–27], which could be due to its association with favorable nutrient intake, improved food choices, and higher physical activity levels [26,28]. The protective effect of physical activity has also been observed in both cross-sectional and longitudinal studies [14,22]. However, the relationship between fast food consumption and obesity has not been established conclusively in the literature. Some studies have shown an inverse association between fast food consumption and obesity [14,27], while others report that fast food consumption is a predictor of weight gain [28,29].

Several socio-environmental factors have also been associated with risk of obesity in adolescents. Parents who have unhealthy lifestyles are more likely to have children who become overweight or obese [30–32]. Conversely, positive parental influence regarding healthy diet and frequent physical activity have been associated with reductions in BMI among overweight and obese adolescents [33]. Studies also report that increased availability of healthy food at home encourages healthy eating in adolescents and is protective against overweight and obesity [34,35], while parental obesity is associated with an increased risk of adolescent and ensuing adult obesity [17]. However, Haines et al. found that the availability of healthy food at home and perceived parental obesity did not predict onset of obesity [14]. Parental concern regarding their child's weight has been positively associated with their child being overweight or obese [14,36]. Parental concern may lead to parental pressure to lose weight and encouragement of restrictive feeding practices, which could lead to weight gain [14,36]. However, parents who reported being concerned about their child's weight were more likely to improve the family's diet, limit child screen time, and attempt to increase their child's physical activity levels [36].

Although there is no individual factor that causes obesity, most research to date has lacked an integrated approach to examine the factors that may be contributing to the high rates of overweight and obesity among youth [17]. An exception would be a study by Haines et al., which looked at the effects of personal, behavioral, and socio-environmental factors on risk of overweight in an ethnically diverse population in Minnesota [14]. To the best of our knowledge, the present study is the first to compare the effects of multiple domains on overweight or obesity risk among a bi-national, ethnically diverse sample of youth. The objective of this study was to identify risk and protective factors for overweight or obesity within the following three domains: Psychosocial, behavioral, and socio-environmental, in a sample of African American, Caucasian, and Latino youths in the US, and Mexican youths in Mexico, and determine differences by sex.

2. Research Methods and Procedures

2.1. Study Population and Data Collection Procedures

US participants were recruited from community centers, schools, clinics, and youth programs in Seattle, Washington and Los Angeles, California (n = 452). A convenience sample of youth was also recruited from the main Mexican Institute of Social Security (IMSS, as per its Spanish abbreviation) hospital in Cuernavaca, Morelos (n = 181). Study flyers were posted in various areas of the IMSS clinic, and potential participants were also informed of the study by staff during their visit to the primary care clinics. All individuals who expressed an interest in the study were contacted by a study recruiter who conducted a telephone interview with the primary caregivers of the potential participants to determine eligibility. Participants had to be African American, Caucasian, or Latino, and between the age of 11–18 years. Youths who met study inclusion criteria of age, 5th grade reading ability, and no serious physical or mental illness diagnosis were informed that participation in the study would involve completing a 40-min questionnaire and having their weight, height, and waist circumference measured. All study participants were enrolled between 2006 and 2008, and informed consent was obtained from each participant and a parent or guardian prior to their inclusion in the study. Further details regarding study design, methodology, and baseline participant characteristics are specified elsewhere [19,37,38]. The Institutional Review Boards of the University of Washington, the University of California, Los Angeles, and the Mexican Institute of Social Security approved all study materials including the study questionnaire, protocol, and consent forms (Seattle Children's Hospital IRB approval number: 11916; IMSS IRB approval number: R-2007-1701-13; UCLA IRB approval number: G06-09-094-01).

Study participants completed a self-administered questionnaire that included the 21-item youth quality of life weight-specific measure (YQOL-W), a generic youth quality of life Instrument (YQOL-R), as well as measures of perceived general health, physical function, body shape satisfaction, and symptoms of depression. The Spanish versions of these measures have been used extensively and validated in other research studies [21,38–44]. All study materials were designed to be readable and understandable at a 5th grade level.

2.2. Study Measures

The following study variables are all reported as indices except for physical activity, which was measured with a single item. Each index score was derived by summing the individual item scores and dividing by the number of items, with the exception of specific scales that have been established and validated (e.g., the YQOL-W, CDI-S, and DEBQ) [19,21,39,45]. The Cronbach's alpha value indicates the internal consistency of each index that was created.

2.2.1. Psychosocial Factors

Body image dissatisfaction. The body image satisfaction scale consists of the following three questions that ask youth to rate their satisfaction with their weight, body shape, and muscle size [46]: How satisfied are you with your weight?; How satisfied are you with your body shape?; How satisfied are you with your muscle size? A body image dissatisfaction index was created using these questions, which ranges from 1 to 5, with higher values indicating greater dissatisfaction (Cronbach's $\alpha = 0.84$).

Depression symptoms. The children's depression inventory: short version (CDI-S) was used to assess depressive symptoms [39]. The CDI-S consists of 10 items with a total score that ranges from 0 to 20. Higher scores indicate a greater presence of depressive symptoms [39].

Self-perception regarding weight. This index comprises three items with a total score ranging from 1 to 5 (Cronbach's $\alpha = 0.77$). A higher score indicates a greater likelihood to regard oneself as overweight or obese. One of the three items is the pictorial body image assessment (PBIA), which asks youth: Which figure in A (female figures) or B (male figures) above is closest to your usual weight? [38] The PBIA silhouettes were modified from Stunkard et al. [47] to include larger body shapes. The silhouettes range from underweight (BMI < 19) to highly severe obesity (BMI > 50) [48]. The two other items ask youth to describe their weight (How do you describe your weight?) [49] and if they ever feel fat (Do you ever feel fat?) [50].

Youth weight-specific quality of life. Weight-specific QOL was evaluated using the YQOL-W, a 21-item instrument with three domain scores: Self, social, and environmental (Cronbach's α for selfitems = 0.90; Cronbach's α for social items = 0.90; Cronbach's α for environmental items = 0.90). The validity and reliability of this instrument has been tested in a multicultural sample of overweight and obese youth in the US and Mexico [19,21]. The YQOL-W has good reliability and validity for assessing weight-specific QOL in children and adolescents, including one-week test–retest intra-class correlation coefficients that were 0.73 for social, 0.71 for self, and 0.73 for environment [19]. The total score for this survey ranges from 0–100, and 100 indicates the best QOL.

2.2.2. Behavioral Factors

Healthy lifestyle priorities. Youth were asked to indicate how much they care about (1) eating healthy food and (2) staying fit and exercising [34]. The following two items: How much do you care about eating healthy food? and How much do you care about staying fit and exercising? were used to construct an index that ranges from 1 to 4, with 4 being the highest level of interest in maintaining a healthy lifestyle (Cronbach's $\alpha = 0.59$).

Physically active. Physical activity was assessed with a single item that asks youth to rate their physical activity level compared to others their age. Compared with most boys/girls your age, would you say that you are: (3) More active, (2) less active, (1) about the same, (0) not sure? The response for this item is on a 0–3 point scale, with a higher score indicating that youth consider themselves to be more physically active than their peers [51].

Fast food consumption. An index was constructed using three items: During the past 7 days, how many times did you eat French fries, sweets, chips, or other foods sometimes called "junk food"? How many times in the past 7 days did you eat breakfast, lunch, or dinner from a "fast food" restaurant? and We have "junk food" in my home. A 1 to 5 scale was used, with higher scores suggesting a higher consumption of fast foods (Cronbach's $\alpha = 0.73$) [34].

Eats breakfast. A breakfast index was created using two items to classify respondents as having eaten breakfast or not based on a yes–no response (1 = no, 2 = yes): Did you eat breakfast today? and I ate breakfast at home. (Cronbach's $\alpha = 0.58$) [34,52].

Weight control behaviors. Participants were asked if they engaged in any of the following weight control behaviors in the past 30 days to lose or keep from gaining weight: (1) Exercise, (2) eating less food, fewer calories, or low-fat foods, (3) fasting, (4) taking diet pills, powders, or liquids [49,52]. A weight control index was created using the following four items based on yes–no responses (1 = no, 2 = yes): During the past 30 days, did you exercise to lose weight or keep from gaining weight?; During the past 30 days, did you eat less food, fewer calories, or foods low in fat to lose weight or keep from gaining weight?; During the past 30 days, did you go without eating for 24 h or more (also called fasting) to lose weight or to keep from gaining weight?; and During the past 30 days, did you take any diet pills, powders, or liquids without a doctor's advice to lose weight or to keep from gaining weight? (Cronbach's $\alpha = 0.37$).

Restrained eating behaviors. Ten items from the Dutch eating behavior questionnaire (DEBQ) were used to assess restrained eating behaviors [45]. The index score ranges from 1 to 5, with 5 indicating a higher frequency of restrained eating practices.

2.2.3. Socio-Environmental Factors

Perceived parental concern regarding adolescent weight. Youths' perception of their parents' concern regarding their weight and if they are getting sufficient physical activity was assessed with a two-item index on a 5-point scale, where 5 is the highest level of perceived parental concern. The following two items were used: How concerned are your parents about you becoming overweight? and How concerned are your parents about you not getting enough physical activity? (Cronbach's $\alpha = 0.75$) [34].

Perceived parent body size. Two indices that represent how youth perceive their parents' body size were constructed using images of the PBIA, one for males and the other for females. Participants were asked to select the figure that is closest to the usual adult weight of their mother and father: Which number under the figures in the figure Box A is closest to the usual adult weight of your mother? and Which number under the figures in the figure Box B is closest to the usual adult weight of your father? The PBIA silhouettes were modified from Stunkard et al. [47] to include larger body shapes. The silhouettes range from underweight (BMI <19) to highly severe obesity (BMI >50) [48]. The 13-point response scale for each item depicts a spectrum of silhouettes with 1 representing underweight, and 13 representing extremely obese (Cronbach's $\alpha = 0.47$) [38].

Mother/father healthy values. Perception of parent healthy values was evaluated using two items that ask how concerned your parents are about (1) staying fit and exercising and (2) losing weight or preventing weight gain: How much does your mother/father feel about staying fit and exercising (for herself/himself)? and How much does your mother feel about losing weight or keeping from gaining weight (for herself/himself)? Separate indices were constructed for mother and father health values, with each consisting of these two items (Cronbach's $\alpha = 0.73$, 0.83, respectively). The response scale for these indices ranges from 1 to 4, with 4 as the highest level of concern [34].

Home availability of healthy foods. An index of fruit and vegetable availability in the home was created using two items that range from 1 to 4, with 4 indicating the highest frequency of healthy food availability in the home: Fruits and vegetables are available in my home . . . (1) Never, (2) Sometimes, (3) Usually, (4) Always, and Vegetables are served at dinner in my home . . . (1) Never, (2) Sometimes, (3) Usually, (4) Always. (Cronbach's $\alpha = 0.69$) [34].

2.2.4. Body Mass Index (BMI)

Height, weight, and waist circumference were measured by trained study staff. Participants were weighed to the nearest 0.1 kg while wearing minimal clothing using a calibrated electronic TANITA scale (model BC533; Tokyo, Japan). Height was determined to the nearest 0.1 cm using a conventional stadiometer, with the youth standing barefoot, with their shoulders in a normal position. BMI was determined to categorize participants as healthy-weight, overweight, or obese, based on the World Health Organization (WHO) age- and sex-specific classifications for youth aged 5 to 19 years [53].

2.3. Statistical Analysis

A descriptive analysis of various sociodemographic variables was conducted for the total study population by country of residence and BMI status. Psychosocial, behavioral, and socio-environmental factors were also examined by country of residence and BMI status. Differences between proportions were assessed using chi-square tests of homogeneity, and t-tests were used to calculate differences between means. Test for trend *p*-values were calculated to determine whether there was a linear association between the study variables and BMI status. Odds ratios and 95% confidence intervals for the association between psychosocial, behavioral, and socio-environmental factors and being overweight or obese were calculated using multinomial logistic regression. These results were adjusted for sex, age, race/ethnicity, and country of residence. Standardized odds ratios were determined to facilitate comparisons of the study variables since their score range varied considerably. Standardized odds ratios improve comparison and interpretability of the logistic regression results. Multinomial logistic regression models for males and females were also used to examine any differences by sex, after adjusting for age, race/ethnicity, and country. All *p*-values presented are 2-tailed and a *p*-value of <0.05 was considered statistically significant. All statistical analyses were performed using STATA software, version 12.0 (StataCorp LP, College Station, TX, USA).

3. Results

The sociodemographic characteristics of the study sample are compared by BMI status in Table 1. Of the 633 participants, 54% are 11–14 years of age, 46% are between 15 and 18 years old, 52% are female, 22% are African American, 25% are Caucasian, 24% are US Latinos, and 29% are youth who live in Mexico. Thirty percent of youth have a healthy BMI, 30% are overweight, and 40% are obese. Thirty-seven percent of participants are from Seattle, WA, 35% are from Los Angeles, CA, and 29% are from Cuernavaca, Mexico. Chi-square tests were used to assess differences by weight status for each of the study variables, separately by country. There are no significant differences by country of residence in terms of sociodemographic characteristics for each of the three BMI categories, except for education level among the US participants. (Table 1)

Table 1. Sample characteristics by body mass index (BMI) categories and country (n = 633).

Sociodemographic Variables	Mexico (n = 181)			p-Value [1]	United States (n = 452)			p-Value [1]
	Healthy n = 43	Overweight n = 68	Obese n = 70		Healthy n = 143	Overweight n = 124	Obese n = 185	
	n (%)	n (%)	n (%)		n (%)	n (%)	n (%)	
Age (years)								
11–14	20 (46.5)	43 (63.2)	40 (57.1)	0.222	73 (51.1)	67 (54.0)	98 (53.0)	0.883
15–18	23 (53.5)	25 (36.8)	30 (42.9)		70 (49.0)	57 (46.0)	87 (47.0)	
Gender								
Female	19 (44.2)	35 (51.5)	34 (48.6)	0.756	77 (53.9)	66 (53.2)	96 (51.9)	0.936
Male	24 (55.8)	33 (48.5)	36 (51.4)		66 (46.1)	58 (46.8)	89 (48.1)	
Race/ethnicity								
African American	-	-	-	0.451	46 (32.2)	35 (28.2)	59 (31.9)	0.798
Caucasian	-	-	-		53 (37.1)	42 (33.9)	65 (35.1)	
Latino	43 (100.0)	68 (100.0)	70 (100.0)		44 (30.1)	47 (37.9)	61 (33.0)	
Education Level								
Elementary School (≤6th grade)	5 (11.6)	15 (22.1)	14 (20.0)	0.336	24 (16.8)	16 (12.9)	19 (10.3)	**<0.001**
Middle School (7th–9th grade)	18 (41.9)	32 (47.1)	36 (51.4)		28 (19.6)	43 (34.7)	81 (43.58)	
High School (≥10th grade)	20 (46.5)	20 (29.4)	20 (28.6)		41 (28.7)	35 (28.2)	57 (30.8)	
Missing	-	1 (1.5)	-		50 (35.0)	30 (24.2)	28 (15.1)	
Mother's Education								
Less than High School	26 (60.5)	36 (52.9)	44 (62.9)	0.818	28 (19.6)	17 (13.7)	31 (16.7)	0.886
High School/GED	14 (32.6)	22 (32.4)	17 (24.3)		22 (15.4)	20 (16.1)	37 (20.0)	
Some college	1 (2.3)	2 (2.9)	2 (2.9)		38 (26.6)	42 (33.9)	52 (28.1)	
University or higher	2 (4.7)	7 (10.3)	7 (10.0)		46 (32.2)	39 (31.5)	54 (29.2)	
Don't know	-	1 (1.5)	-		4 (2.8)	2 (1.6)	3 (1.6)	
Missing	-	-	-		5 (3.5)	4 (3.2)	8 (4.3)	
Father's Education								
Less than High School	23 (53.5)	34 (50.0)	38 (54.3)	1.124	26 (18.2)	24 (19.4)	34 (18.4)	0.069
High School/GED	13 (30.2)	13 (30.2)	11 (15.7)		28 (19.6)	15 (12.1)	50 (27.0)	
Some college	2 (4.7)	2 (4.7)	8 (11.4)		26 (18.2)	36 (29.0)	30 (16.2)	
University or higher	3 (7.0)	16 (23.5)	9 (12.9)		40 (28.0)	26 (21.0)	43 (23.2)	
Don't know	-	-	-		14 (9.8)	16 (12.9)	17 (9.2)	
Missing	2 (4.7)	2 (2.9)	4 (5.7)		9 (6.3)	7 (5.7)	11 (6.0)	

Sample sizes may not add up to marginal totals due to missing values. [1] Differences between proportions were performed using chi-square tests of homogeneity by weight status for each of the study variables, separately by country. Statistically significant results are in bold.

Tables 2 and 3 compare the mean scores for various psychosocial, behavioral, and socio-environmental variables, by country of residence and BMI status. Within the domain of psychosocial factors, overweight or obese youth in Mexico and the US are more likely to report being dissatisfied with their body image, to perceive themselves as overweight or obese, and to have lower weight-specific QOL scores than healthy-weight youth. However, overweight or obese youths in Mexico are not more likely to report more depressive symptoms than healthy-weight youths, unlike obese youths in the US, who are more likely to report depressive symptoms than healthy-weight youths (3.2 vs. 2.3, respectively). The presence of depressive symptoms is greater among healthy-weight, overweight, and obese youth in Mexico (3.1, 3.7, 3.9, respectively) than those in the US (2.1, 2.8, 3.2, respectively); and the weight-related QOL reported by overweight or obese youths in Mexico is lower than those in the US (65.1 and 52.9 vs. 78.1 and 67.0, respectively).

In terms of behavioral factors, obese youth in the US have lower healthy lifestyle priorities (3.0 vs. 3.1, respectively), are less physically active (1.6 vs. 1.8, respectively), and are less likely to eat breakfast (1.5, 1.6, respectively) than healthy-weight youth in the US. Overweight and obese participants in both countries are also significantly more likely to engage in weight control behaviors, such as exercise and restrained eating, as compared to healthy-weight youth. The only statistically significant socio-environmental factors reported by obese youths in both countries include being more likely to think that their parents are concerned about their weight and that their parents have a larger body size than healthy-weight youth. However, overweight youths in the US are not more likely to report that their parents are concerned about their weight, as compared to healthy weight youths (Tables 2 and 3). All the aforementioned results had a *p*-value of <0.05, which was considered statistically significant.

Table 2. Comparison of various psychosocial, behavioral, and socio-environmental factors by BMI category among youths in Mexico (n = 181).

	Range *	Healthy Mean ± SD	Overweight Mean ± SD	Obese Mean ± SD	$P_{overweight}$ ǂ	P_{obese} ǂ	P_{trend} ǂ
Psychosocial Factors							
Dissatisfied with Body Image	1 to 5	2.7 ± 1.3	**3.1 ± 1.1**	**3.8 ± 0.9**	**0.034**	**<0.001**	**<0.001**
Depression symptoms (CDI-S)	0 to 20	3.1 ± 3.3	3.7 ± 3.2	3.9 ± 3.1	0.364	0.223	0.100
Self-perception regarding weight	1 to 5	2.2 ± 0.8	**2.8 ± 0.7**	**3.5 ± 0.6**	**<0.001**	**<0.001**	**<0.001**
Perceived Body Shape (PBIA)	1 to 13	3.0 ± 1.5	**4.4 ± 1.3**	**6.3 ± 1.6**	**<0.001**	**<0.001**	**<0.001**
Body Weight Description	1 to 5	3.1 ± 0.9	**3.9 ± 0.6**	**4.3 ± 0.6**	**<0.001**	**<0.001**	**<0.001**
Feeling Fat	1 to 5	2.4 ± 1.2	**3.1 ± 1.1**	**3.7 ± 1.0**	**0.006**	**<0.001**	**<0.001**
Youth weight-related quality of life (YQOL-W)	0 to 100	75.8 ± 28.1	**65.1 ± 26.0**	**52.9 ± 26.4**	**0.045**	**<0.001**	**<0.001**
Behavioral Factors							
Healthy lifestyle priorities	1 to 4	2.8 ± 0.6	2.9 ± 0.7	2.8 ± 0.7	0.797	0.723	0.569
Physically active	0 to 3	1.5 ± 1.0	1.6 ± 1.0	1.3 ± 1.0	0.785	0.255	0.176
Fast food consumption	1 to 5	1.1 ± 0.4	1.1 ± 0.5	1.0 ± 0.4	0.765	0.073	0.023
Eats breakfast	1 to 2	1.7 ± 0.4	1.6 ± 0.4	1.6 ± 0.4	0.416	0.205	0.239
Weight control behaviors	1 to 2	1.2 ± 0.2	**1.4 ± 0.2**	**1.4 ± 0.2**	**0.002**	**<0.001**	**0.001**
Exercises	1 to 2	1.5 ± 0.5	**1.8 ± 0.4**	**1.8 ± 0.4**	**0.005**	**0.003**	**0.006**
Eat less, few calories, low-fat foods	1 to 2	1.3 ± 0.5	**1.5 ± 0.5**	**1.6 ± 0.5**	**0.045**	**0.032**	**0.046**
Fasting	1 to 2	1.1 ± 0.3	1.1 ± 0.3	1.2 ± 0.4	0.415	0.174	0.168
Diet pills, powders, or liquids	1 to 2	1.0 ± 0.0	1.0 ± 0.1	1.1 ± 0.2	0.429	0.112	0.058
Restrained eating behaviors (DEBQ-R)	1 to 5	2.1 ± 0.8	**2.6 ± 0.8**	**2.7 ± 0.7**	**0.001**	**<0.001**	**<0.001**
Socio-environmental factors							
Perceived parental concern regarding weight	1 to 5	2.7 ± 1.2	**3.2 ± 1.1**	**3.6 ± 1.0**	**0.0316**	**<0.001**	**<0.001**
Perceived parent body size (PBIA)	1 to 13	4.3 ± 1.2	**4.9 ± 1.5**	**5.6 ± 1.4**	**0.0292**	**<0.001**	**<0.001**
Mother healthy values	1 to 4	2.8 ± 0.7	2.6 ± 0.8	2.6 ± 0.7	0.1260	0.1383	0.178
Father healthy values	1 to 4	2.1 ± 0.6	2.3 ± 0.8	2.3 ± 0.8	0.1656	0.1333	0.191
Home availability of healthy foods	1 to 4	3.1 ± 0.7	3.0 ± 0.7	2.9 ± 0.7	0.5642	0.1474	0.103

* A higher score indicates a greater frequency or agreement; ¶ Reference category for comparisons between BMI groups; ǂ Differences between means were performed using t-tests; ǂ Cuzick's trend test; statistically significant results are in bold.

Table 3. Comparison of various psychosocial, behavioral, and socio-environmental factors by BMI category among youths in the Unites States (n = 452).

	Range *	Healthy Mean ± SD	Overweight Mean ± SD	Obese Mean ± SD	$P_{overweight}$ ⁱ	P_{obese} ⁱ	P_{trend} ⁱ
Psychosocial Factors							
Dissatisfied with Body Image	1 to 5	2.3 ± 0.9	**2.8 ± 1.0**	**3.2 ± 0.9**	**<0.001**	**<0.001**	**<0.001**
Depression symptoms (CDI-S)	0 to 20	2.1 ± 2.6	2.8 ± 3.5	**3.2 ± 3.6**	0.059	**0.001**	**0.003**
Self-perception regarding weight	1 to 5	1.9 ± 0.5	**2.6 ± 0.7**	**3.1 ± 0.7**	**<0.001**	**<0.001**	**<0.001**
Perceived Body Shape (PBIA)	1 to 13	2.4 ± 1.3	**3.9 ± 1.4**	**5.2 ± 1.6**	**<0.001**	**<0.001**	**<0.001**
Body Weight Description	1 to 5	2.8 ± 0.7	**3.5 ± 0.7**	**4.1 ± 0.8**	**<0.001**	**<0.001**	**<0.001**
Feeling Fat	1 to 5	2.0 ± 1.1	**2.8 ± 1.3**	**3.3 ± 1.2**	**<0.001**	**<0.001**	**<0.001**
Youth weight-related quality of life (YQOL-W)	0 to 100	90.4 ± 14.1	**78.1 ± 23.0**	**67.0 ± 27.5**	**<0.001**	**<0.001**	**<0.001**
Behavioral Factors							
Healthy lifestyle priorities	1 to 4	3.1± 0.7	3.0± 0.6	**3.0 ± 0.7**	0.099	**0.042**	**0.028**
Physically active	0 to 3	1.8 ± 1.1	1.6 ± 1.0	**1.6 ± 1.0**	0.105	**0.038**	0.058
Fast food consumption	1 to 5	1.8 ± 0.9	1.8 ± 1.0	**1.6 ± 0.7**	0.639	**0.011**	**0.041**
Eats breakfast	1 to 2	1.6 ± 0.4	1.6 ± 0.4	**1.5 ± 0.4**	0.395	**0.010**	**0.008**
Weight control behaviors	1 to 2	1.3 ± 0.2	**1.3 ± 0.2**	**1.4 ± 0.2**	**0.021**	**0.001**	**<0.001**
Exercises	1 to 2	1.6 ± 0.5	1.7 ± 0.4	**1.7 ± 0.4**	0.053	**0.018**	**0.020**
Eat less, few calories, low-fat foods	1 to 2	1.4 ± 0.5	**1.5 ± 0.5**	**1.6 ± 0.5**	**0.003**	**<0.001**	**0.001**
Fasting	1 to 2	1.1 ± 0.3	1.1 ± 0.3	1.1 ± 0.3	0.577	0.281	0.224
Diet pills, powders, or liquids	1 to 2	1.0 ± 0.2	1.0 ± 0.2	1.1 ± 0.2	0.726	0.691	0.648
Restrained eating behaviors (DEBQ-R)	1 to 5	2.1 ± 0.8	**2.5 ± 0.9**	**2.7 ± 0.8**	**0.001**	**<0.001**	**<0.001**
Socio-environmental factors							
Perceived parental concern regarding weight	1 to 5	2.6 ± 1.3	2.7 ± 1.2	**3.2 ± 1.2**	0.766	**<0.001**	**<0.001**
Perceived parent body size (PBIA)	1 to 13	3.9 ± 1.5	**4.9 ± 1.6**	**5.1 ± 1.5**	**<0.001**	**<0.001**	**<0.001**
Mother healthy values	1 to 4	3.1 ± 0.9	3.1 ± 0.8	3.1 ± 0.8	0.854	0.694	0.992
Father healthy values	1 to 4	2.8 ± 0.9	2.6 ± 1.0	2.9 ± 1.0	0.154	0.255	0.173
Home availability of healthy foods	1 to 4	3.2 ± 0.7	3.2 ± 0.7	3.2 ± 0.8	0.888	0.533	0.391

* A higher score indicates a greater frequency or agreement; ¶ Reference category for comparisons between BMI groups; ⁱ Differences between means were performed using t-tests; ⁱ Cuzick's trend test; statistically significant results are in bold.

The standardized and adjusted odds ratios for various psychosocial, behavioral, and socio-environmental factors, by BMI status, among youths in Mexico and the US (controlling for age, sex, race/ethnicity, and country of residence) are reported in Table 4. Overweight and obese youth have significantly greater odds of reporting body image dissatisfaction (OR = 1.67, OR = 2.95), having depressive symptoms (OR = 1.08, OR = 1.12), perceiving themselves as overweight or obese, and having a lower weight-specific QOL (OR = 0.97, OR = 0.95), than healthy-weight youth. Obese youth in both countries also have significantly lower odds of having healthy lifestyle priorities (OR = 0.75), being physically active (OR = 0.79), consuming fast food (OR = 0.68), and eating breakfast (OR = 0.47), than healthy-weight youth. Overweight and obese youth are significantly more likely to engage in weight control behaviors (OR = 5.19, OR = 8.88), such as exercise (OR = 1.99, OR = 2.12), as well as eating less, fewer calories, and lower-fat food (OR = 2.15, OR = 2.32) than healthy-weight youth. In addition, overweight and obese youth have significantly greater odds of restrained eating behaviors (OR = 1.86, OR = 2.35) than healthy-weight youth. Both groups are also significantly more likely to perceive their parent as overweight or obese (OR = 1.49, OR = 1.71), and obese youth have significantly greater odds of reporting that their parents are very concerned about their weight (OR = 1.56), compared to healthy-weight youth. The standardized odds ratio results indicate that the following psychosocial, behavioral, and socio-environmental factors are most significantly associated with overweight and obesity: Perceived body shape (OR = 6.31, OR = 25.89), restrained eating behaviors (OR = 1.7, OR = 2.08), and perceived parent body shape (OR = 1.88, OR = 2.34), respectively. The standardized odds ratios of measures with scales that have a wider range, such as the CDI-S (0–20) and the YQOL-W (0–100), show a stronger association with overweight and obesity, than the non-standardized odds ratios. (Table 4) All the aforementioned results had a p-value of <0.05, which was considered statistically significant.

Table 4. Standardized and adjusted odds ratios for psychosocial, behavioral, and socio-environmental factors by BMI status, among youth in Mexico and the US (n = 633).

	S_OR	Overweight ǀ	S_OR ¶	Obese ǀ
		OR (95% CI) ∞		OR (95% CI) ∞
Psychosocial Factors				
Dissatisfied with Body Image	1.74	**1.67 (1.3, 2.1) ***	3.22	**2.95 (2.3, 3.7) ***
Depression symptoms (CDI-S)	1.29	**1.08 (1.0, 1.2) ****	1.44	**1.12 (1.1, 1.2) ***
Self-perception regarding weight	4.75	**6.63 (4.3, 10.3) ***	15.13	**27.03 (16.3, 44.8) ***
Perceived body shape	6.31	**2.57 (2.1, 3.2) ***	25.89	**5.30 (4.1, 6.8) ***
Body weight description	3.47	**4.02 (2.8, 5.7) ***	10.77	**14.30 (9.4, 21.7) ***
Feeling fat	2.41	**1.99 (1.6, 2.5) ***	4.29	**3.12 (2.5, 3.9) ***
Youth weight-related quality of life (YQOL-W)	0.46	**0.97 (0.96, 0.98) ***	0.28	**0.95 (0.94, 0.96) ***
Behavioral Factors				
Healthy lifestyle priorities	0.87	0.82 (0.6, 1.1)	0.82	**0.75 (0.6, 1.0) ****
Physically active	0.88	0.88 (0.7, 1.1)	0.79	**0.79 (0.6, 1.0) ****
Fast food consumption	0.97	0.97 (0.7, 1.3)	0.73	**0.68 (0.5, 0.9) ****
Eats breakfast	0.87	0.70 (0.4, 1.2)	0.73	**0.47 (0.3, 0.8) ****
Weight control behaviors	1.46	**5.19 (2.0, 13.2) ***	1.64	**8.88 (3.7, 21.5) ***
Exercises	1.37	**1.99 (1.3, 3.1) ****	1.41	**2.12 (1.4, 3.2) ***
Eat less, few calories, low-fat foods	1.47	**2.15 (1.4, 3.3) ***	1.52	**2.32 (1.6, 3.4) ***
Fasting	1.00	0.99 (0.5, 2.0)	1.18	1.68 (0.9, 3.1)
Diet pills, powders, or liquids	0.97	0.86 (0.3, 2.6)	1.11	1.64 (0.7, 4.1)
Restrained eating behaviors (DEBQ-R)	1.70	**1.86 (1.4, 2.4) ***	2.08	**2.35 (1.8, 3.0)***
Socio-environmental factors				
Perceived parental concern regarding weight	1.15	1.12 (0.9, 1.3)	1.76	**1.56 (1.3, 1.8) ***
Perceived parent body shape	1.88	**1.49 (1.3, 1.7) ***	2.34	**1.71 (1.5, 2.0) ***
Mother healthy values	0.94	0.92 (0.7, 1.2)	0.97	0.96 (0.7, 1.2)
Father healthy values	0.93	0.92 (0.7, 1.2)	1.21	1.22 (1.0, 1.6)
Home availability of healthy foods	0.94	0.96 (0.7, 1.3)	0.97	0.98 (0.7, 1.3)

¶ Standardized odds ratios; ǀ Healthy is reference category for comparison between BMI groups; ∞ Adjusted for age, gender, race/ethnicity, and country; * p-value ≤ 0.001; ** p-value < 0.05; significant results are in bold.

Table 5 presents the logistic regression results for the psychosocial, behavioral, and socio-environmental factors, stratified by sex. Some important differences are observed by sex. For example, overweight or obese boys are more likely to report dissatisfaction with their body image (OR = 1.81 and OR = 3.21, respectively) than girls (OR = 1.59 and OR = 2.78, respectively). However, the presence of depressive symptoms is significantly greater among overweight and obese females (OR = 1.14 and OR = 1.16, respectively) but not among males. Girls are also more likely to perceive themselves as overweight or obese and "feel fat" (OR = 8.91 and OR = 34.28, respectively) than boys (OR = 7.14 and OR = 32.28, respectively). Obese females are significantly less likely to be physically active (OR = 0.72) and eat breakfast than healthy-weight females (OR = 0.40), but this association was not found to be significant among males. Overweight or obese males are more likely to engage in weight control behaviors (OR = 13.77 and OR = 12.69, respectively) than obese females (OR = 8.02), especially exercise (OR = 2.67 and OR = 2.59, respectively) and eating less/few calories/low-fat foods (OR = 3.40 and OR = 2.93, respectively). However, obese girls are significantly more likely to consume diet pills, powders or liquids (OR = 9.59) than boys (Table 5). All the aforementioned results had a p-value of <0.05, which was considered statistically significant.

Table 5. Association between psychosocial, behavioral, and socio-environmental factors and overweight or obesity, by sex (n = 633).

	Female		Male	
	Overweight [ǂ] OR (95% CI) [∞]	Obese [ǂ] OR (95% CI) [∞]	Overweight [ǂ] OR (95% CI) [∞]	Obese [ǂ] OR (95% CI) [∞]
Psychosocial Factors				
Body image dissatisfaction	**1.59 (1.2, 2.2) ****	**2.78 (2.0, 3.8) ***	**1.81 (1.3, 2.5) ***	**3.21 (2.3, 4.5) ***
Depression symptoms (CDI-S)	**1.14 (1.0, 1.3) ****	**1.16 (1.1, 1.3) ***	1.00 (0.9, 1.1)	1.07 (0.97, 1.2)
Self-perception regarding weight	**8.91 (4.5, 17.5) ***	**34.28 (16.1, 73.1) ***	**7.14 (3.7, 13.9) ***	**32.28 (15.1, 69.0) ***
Perceived body shape	**3.43 (2.4, 4.9) ***	**7.07 (4.7, 10.6) ***	**2.16 (1.7, 2.8) ***	**4.45 (3.2, 6.1) ***
Body weight description	**4.45 (2.7, 7.3) ***	**12.94 (7.3, 23.0) ***	**3.87 (2.3, 6.4) ***	**17.37 (9.2, 32.7) ***
Feeling fat	**2.02 (1.5, 2.7) ***	**3.15 (2.3, 4.3) ***	**2.25 (1.6, 3.1) ***	**3.53 (2.5, 4.9) ***
Youth weight-specific quality of life (YQOL-W)	**0.96 (0.95,0.98) ***	**0.95 (0.93, 0.96) ***	**0.98 (0.96, 0.99) ****	**0.96 (0.94, 0.97) ***
Behavioral Factors				
Healthy lifestyle priorities	**0.59 (0.4, 0.9) ****	0.75 (0.5, 1.1)	1.16 (0.8, 1.8)	0.76 (0.5, 1.1)
Physically active	0.76 (0.6, 1.0)	**0.72 (0.5, 0.9) ****	1.01 (0.7, 1.4)	0.87 (0.7, 1.1)
Fast food consumption	1.12 (0.8,1.7)	0.76 (0.5, 1.1)	0.85 (0.6, 1.3)	**0.63 (0.4, 0.9) ****
Eats breakfast	0.51 (0.2, 1.1)	**0.40 (0.2, 0.8) ****	0.90 (0.4, 2.0)	0.50 (0.2, 1.0)
Weight control behaviors	3.00 (0.9, 10.5)	**8.02 (2.4, 26.8) ***	**13.77 (3.2, 59.2) ***	**12.69 (3.3, 49.2) ***
Exercise	1.64 (0.9, 3.0)	**1.88 (1.1, 3.4) ****	**2.67 (1.4, 5.2) ****	**2.59 (1.4, 4.7) ****
Eat less, few calories, low-fat foods	1.53 (0.9, 2.7)	**2.00 (1.2, 3.4) ****	**3.40 (1.8, 6.5) ***	**2.93 (1.6, 5.4) ***
Fasting	0.93 (0.4, 2.4)	1.77 (0.8, 4.0)	1.24 (0.4, 3.9)	1.67 (0.6, 4.6)
Diet pills, powders, or liquids	4.16 (0.4, 38.6)	**9.59 (1.2, 76.7) ****	0.41 (0.1, 2.2)	0.54 (0.1, 2.0)
Restrained eating behaviors (DEBQ-R)	**1.69 (1.2, 2.4) ****	**2.76 (1.9, 4.0) ***	**2.22 (1.5, 3.3) ***	**2.11 (1.5, 3.0) ***
Socio-environmental Factors				
Perceived parental concern regarding weight	1.03 (0.8, 1.3)	**1.50 (1.2, 1.9) ***	1.21 (0.9, 1.5)	**1.61 (1.3, 2.0) ***
Perceived parent body size	**1.47 (1.2, 1.8) ***	**1.67 (1.4, 2.0) ***	**1.56 (1.3, 1.9) ***	**1.83 (1.5, 2.3) ***
Mother healthy values	0.83 (0.6, 1.2)	0.81 (0.6, 1.2)	1.00 (0.7, 1.4)	1.12 (0.8, 1.6)
Father healthy values	0.78 (0.5, 1.1)	1.33 (0.9, 1.9)	1.05 (0.7, 1.6)	1.06 (0.7, 1.5)
Home availability of healthy foods	0.95 (0.6, 1.4)	0.96 (0.7, 1.4)	0.94 (0.6, 1.4)	0.99 (0.7, 1.5)

[ǂ] Healthy is reference category for comparison between BMI groups; [∞] Adjusted for age, race/ethnicity, and country; * p-value ≤ 0.001; ** p-value < 0.05; statistically significant results are in bold.

4. Discussion

The primary objective of this study was to examine the relevance of various psychosocial, behavioral, and socio-environmental factors among overweight and obese youth in the US and Mexico, and to determine differences by sex. We aimed to address gaps in the current research by studying factors in distinct domains among an ethnically diverse, bi-national sample of youth. Our results support the findings of other studies in the US that have examined similar factors within these three domains [13–36]. However, as far as we know, our study is the first to explore the effects of multiple psychosocial factors, behavioral, and socio-environmental factors on overweight or obesity risk in a diverse sample of youth. By simultaneously examining all of these factors in one sample, we were able to contrast the relevance of different risk factors in a single large group, rather than across various studies, which may be difficult to compare. Additionally, to the best of our knowledge, this is one of the first studies to shed light on the association between psychosocial, behavioral, and socio-environmental factors and the presence of overweight or obesity among youth in Mexico and Latinos living in the US.

In our study, psychosocial factors, such as a higher rate of body image dissatisfaction, depressive symptoms, self-perception of overweight, and a lower weight-related QOL, were most strongly associated with overweight or obesity. These results are consistent with other studies, which found a higher prevalence of these factors among overweight or obese youth, as compared to healthy-weight youth [13–15,17,19,20]. We found that depressive symptoms are significantly associated with overweight or obesity among girls but not boys. Inconsistent gender differences have previously been reported for the relationship between depressive symptoms and obesity [14,17]. These mixed results could be attributable to variations in study design or assessment of depression [54], or due to the characteristics of the study sample. A meta-analysis of 17 studies concluded that depression is positively associated with BMI but only among females [54]. Interestingly, overweight or obese youths in Mexico did not report more depressive symptoms than healthy-weight youths, unlike obese youths in the US, who did report more depressive symptoms than healthy-weight youths. Our findings also indicate a higher prevalence of depressive symptoms among youths in Mexico than in the US.

In terms of self-perception regarding weight, overweight or obese girls were more likely to perceive themselves as overweight or obese than boys. Similar differences have been observed with adolescent girls being more likely to perceive themselves as overweight or obese than boys [13]. A recent study investigated brain activation using functional magnetic resonance imaging during a body perception task in healthy males and females. They found that images of their own bodies were more salient for the female participants and concluded that females may be more vulnerable than males to conditions involving own body perception [55]. Youths in Mexico reported higher scores for all the "self-perception regarding weight categories", than youths in the US. Obese adolescents have been shown to report a lower QOL [20], which was also found in this study, with overweight or obese youth reporting significantly lower weight-related QOL than healthy-weight youth. Additionally, the weight-related QOL reported by overweight or obese youths in Mexico was lower than in the US. Notably, self-reported QOL is lower in Mexico than in the US., regardless of weight status.

The multivariate analyses indicate that obese youth were less likely to have healthy lifestyle priorities, be physically active, or eat breakfast. However, when stratified by sex, only obese females were significantly less likely to engage in physical activity. Obese and overweight youth were twice as likely to report that they exercise for weight control, compared to healthy weight youths. There are contradictory findings regarding the effect of physical activity by gender, with one study showing a protective effect only among boys [14] and another only among girls [22]. By contrast, eating breakfast has shown a consistent protective effect for boys and girls in various studies, across different ethnic groups [14,26,27]. Our results also indicate that obese youth are less likely to consume breakfast, but when stratified by sex, this association only remained significant among obese females.

We found that obese males are less likely to report that they eat fast food, as compared to healthy-weight males. Additionally, overweight or obese youths in Mexico are less likely to eat fast food than their counterparts in the US. Previous studies have reported a negative association between eating fast food and obesity among males [27] and females [14,27]. However, other researchers have found that fast food consumption is associated with increased risk of obesity [28,29]. When relying on self-reported behaviors, there may be a higher likelihood of over reporting of socially desirable behaviors, which could explain the inverse association between fast food consumption and obesity observed in this study. Several weight control behaviors were also significantly associated with overweight and obesity in this study. There was a stronger association between weight control behaviors and BMI among males compared to females. Unhealthy weight control behaviors have been shown to predict weight gain in boys and girls [14,15,17,22]. Restrained eating was also found to be a risk factor for obesity in our study, which has previously been reported in other studies [17].

Socio-environmental factors were found to have the least significant associations with overweight or obesity. In this study, obese youth were more likely to believe that their parents are concerned about their weight, which has been previously reported in the literature [14,36]. Parental obesity has also been examined in various studies because children of obese parents may be at greater risk for obesity due to shared genetic and environmental factors [17,56]. In this study, youth who perceived their parents as heavier were more likely to be overweight or obese. Although parental health values and the availability of healthy foods at home have been reported to be significant in other studies [30,31,34,35], no significant associations were found in this study.

This study has some limitations, including that it is cross-sectional, and thus, no conclusions about the direction of causality can be made and there is a possibility of reporting bias. Participants were recruited by means of convenience sampling and might not be representative of their respective weight groups. Additionally, this is an exploratory study with a limited sample size for the participants in Mexico. Future studies should be conducted with a larger sample size that will allow for a higher significance threshold to be set for individual comparisons to compensate for the number of inferences being made. Other limitations include the specific measures that were collected using a self-reported questionnaire, a lack of validated measures, and the fact that some of the behavioral and socio-environmental indices, e.g., "healthy lifestyle priorities," "physically active," "mother/father healthy values," or "home availability of healthy foods", were created based on a limited number of variables and should be interpreted as preliminary findings. The information provided by the study participants was of a quantitative nature, so we were unable to determine the reason for some of the differences observed by sex or country of origin. Future studies should collect more qualitative data to investigate these differences. A strength of this study is that it explored the issue of overweight and obesity among an ethnically diverse group of youth in the US and Mexico, including African Americans and Latinos, who are disproportionately affected by obesity. Additionally, this study examined a breadth of risk factors that have not been analyzed in a comprehensive and comparative manner. Although some of the indices we created to measure eating behaviors do not have a high reliability score, the associations we observed support the expected relationships, especially when obesity is the main outcome variable. The use of indices in this study to combine various factors also allowed for a robust analysis of complex concepts.

5. Conclusions

The results of this bi-national study highlight some of the differences and similarities in various psychosocial, behavioral, and socio-environmental factors among a multiethnic sample of healthy-weight, overweight, and obese youths. We hope our findings help to demonstrate the importance of considering a wide range of risk and protective factors for obesity among adolescents, when planning future studies and interventions. Additionally, our results support the need for multifactorial approaches when developing interventions to address the growing problem of obesity among youth in the US and Mexico. Intervention programs should use an integrated approach that addresses several of these factors to help to reduce the alarmingly high rates of obesity among youth in the US and Mexico. More research is needed on how these factors may interact with each other to cause obesity, since many are interrelated. Our study paves the way for future studies to focus on adopting a transdisciplinary approach to identify and address important risk factors for obesity among youth.

Author Contributions: Conceptualization, Y.N.F., L.S.M., N.S., and D.L.P.; Data curation, Z.A.C. and P.R.-P.; Formal analysis, Z.A.C. and P.R.-P.; Funding acquisition, Y.N.F., L.S.M., J.S., and D.L.P.; Investigation, Y.N.F., L.S.M., T.C.E., and D.L.P.; Methodology, Y.N.F., Z.A.C., P.R.-P., L.S.M., T.C.E., K.G.-C., N.S., and D.L.P.; Project administration, Y.N.F., L.S.M., T.C.E., and D.L.P.; Resources, Y.N.F., L.S.M., J.S., and D.L.P.; Software, Z.A.C. and P.R.-P.; Supervision, Y.N.F., L.S.M., T.C.E., and D.L.P.; Writing—original draft, Y.N.F., and Z.A.C.; Writing—review and editing, Y.N.F., Z.A.C., P.R.-P., L.S.M., T.C.E., K.G.-C., J.S., C.M.L., and D.L.P.

Funding: This study was supported by a grant from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) awarded to Patrick (grant number R01DK071101-01A2), and an NIDDK Research Supplement to Promote Diversity in Health-Related Research Grant awarded to Flores (grant number: 3R01DK071101-02S1). Contreras received a grant from the Drabkin–Neumann Internship at UCLA to support her work on this study. Flores was also supported by NIH/NCI K07CA197179. Additional funding was provided by the Epidemiological and Health Services Research Unit of the Mexican Institute of Social Security.

Conflicts of Interest: The authors declare no conflict of interest.

Compliance with Ethical Standards: All procedures performed in this study were in accordance with the ethical standards of the institutional research committees and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The Institutional Review Boards of the University of Washington, the University of California, Los Angeles, and the Mexican Institute of Social Security approved all study materials, including the study questionnaire, protocol, and consent forms. Informed consent was obtained from all individual participants included in the study.

References

1. World Health Organization. Facts and figures on childhood obesity. Available online: https://www.who.int/end-childhood-obesity/facts/en/ (accessed on 31 March 2019).

2. NCD Risk Factor Collaboration (NCD-Risc). Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: A pooled analysis of 2416 population-based measurement studies in 128.9 million children, adolescents, and adults. *Lancet* **2017**, *390*, 2627–2642. [CrossRef]
3. Lee, H.; Pantazis, A.; Cheng, P.; Dennisuk, L.; Clarke, P.J.; Lee, J.M. The association between adolescent obesity and disability incidence in young adulthood. *J. Adolesc. Health* **2016**, *59*, 472–478. [CrossRef] [PubMed]
4. Reilly, J.J.; Kelly, J. Long-term impact of overweight and obesity in childhood and adolescence on morbidity and premature mortality in adulthood: Systematic review. *Int. J. Obes.* **2011**, *35*, 891–898. [CrossRef] [PubMed]
5. Ogden, C.L.; Carroll, M.D.; Kit, B.K.; Flegal, K.M. Prevalence of childhood and adult obesity in the United States, 2011-2012. *JAMA* **2014**, *311*, 806–814. [CrossRef] [PubMed]
6. Hales, C.M.; Carroll, M.D.; Fryar, C.D.; Ogden, C.L. Prevalence of Obesity among Adults and Youth: United States, 2015–2016. Available online: https://www.cdc.gov/nchs/data/databriefs/db288.pdf (accessed on 26 April 2019).
7. Schaefer, S.E.; Salazar, M.; Bruhn, C.; Saviano, D.; Boushey, C.; Van Loan, M.D. Influence of race, acculturation, and socioeconomic status on tendency toward overweight in Asian-American and Mexican-American early adolescent females. *J. Immigr. Minor. Health* **2009**, *11*, 188–197. [CrossRef] [PubMed]
8. Skelton, J.A.; Cook, S.R.; Auinger, P.; Klein, J.D.; Barlow, S.E. Prevalence and trends of severe obesity among US children and adolescents. *Acad. Pediatr.* **2009**, *9*, 322–329. [CrossRef]
9. Centers for Disease Control and Prevention. National Center for Health Statistics. National Health and Nutrition Examination Survey. Obesity among children and adolescents aged 2–19 years, by selected characteristics: United States, selected years 1988–1994 through 2013–2016. Available online: https://www.cdc.gov/nchs/data/hus/2017/059.pdf (accessed on 21 December 2018).
10. Morales, L.S.; Flores, Y.N.; Leng, M.; Sportiche, N.; Gallegos-Carrillo, K.; Salmeron, J. Risk factors for cardiovascular disease among Mexican-American adults in the United States and Mexico: A comparative study. *Salud Publica Mex.* **2014**, *56*, 197–205. [CrossRef]
11. Liu, J.H.; Chu, Y.H.; Frongillo, E.A.; Probst, J.C. Generation and acculturation status are associated with dietary intake and body weight in Mexican American adolescents. *J. Nutr.* **2012**, *142*, 298–305. [CrossRef]
12. Encuesta Nacional de Salud y Nutrición de Medio Camino. Instituto Nacional de Salud Pública (MX): Cuernavaca, México, 2016. Available online: http://transparencia.insp.mx/2017/auditorias-insp/12701_Resultados_Encuesta_ENSANUT_MC2016.pdf (accessed on 21 December 2018).
13. Chung, A.E.; Perrin, E.M.; Skinner, A.C. Accuracy of child and adolescent weight perceptions and their relationships to dieting and exercise behaviors: A NHANES study. *Acad. Pediatr.* **2013**, *13*, 371–378. [CrossRef]
14. Haines, J.; Neumark-Sztainer, D.; Wall, M.; Story, M. Personal, behavioral, and environmental risk and protective factors for adolescent overweight. *Obesity* **2007**, *15*, 2748–2760. [CrossRef]
15. Vander Wal, J.S. Unhealthy weight control behaviors among adolescents. *J. Health Psychol.* **2012**, *17*, 110–120. [CrossRef]
16. Stein, D.; Weinberger-Litman, S.L.; Latzer, Y. Psychosocial perspectives and the issue of prevention in childhood obesity. *Front. Public Health* **2014**, *2*, 104. [CrossRef]
17. Stice, E.; Presnell, K.; Shaw, H.; Rohde, P. Psychological and behavioral risk factors for obesity onset in adolescent girls: A prospective study. *J. Consult. Clin. Psychol.* **2005**, *73*, 195–202. [CrossRef]
18. Kolodziejczyk, J.K.; Gutzmer, K.; Wright, S.M.; Arredondo, E.M.; Hill, L.; Patrick, K.; Huang, J.S.; Gottschalk, M.; Norman, G.J. Influence of specific individual and environmental variables on the relationship between body mass index and health-related quality of life in overweight and obese adolescents. *Qual. Life Res.* **2015**, *24*, 251–261. [CrossRef]
19. Morales, L.S.; Edwards, T.C.; Flores, Y.; Barr, L.; Patrick, D.L. Measurement properties of a multicultural weight-specific quality-of-life instrument for children and adolescents. *Qual. Life Res.* **2011**, *20*, 215–224. [CrossRef]
20. Tyler, C.; Johnston, C.A.; Fullerton, G.; Foreyt, J.P. Reduced quality of life in very overweight Mexican American adolescents. *J. Adolesc. Health* **2007**, *40*, 366–368. [CrossRef]
21. Patrick, D.L.; Skalicky, A.M.; Edwards, T.C.; Kuniyuki, A.; Morales, L.S.; Leng, M.; Kirschenbaum, D.S. Weight loss and changes in generic and weight-specific quality of life in obese adolescents. *Qual. Life Res.* **2011**, *20*, 961–968. [CrossRef]

22. Mellin, A.E.; Neumark-Sztainer, D.; Story, M.; Ireland, M.; Resnick, M.D. Unhealthy behaviors and psychosocial difficulties among overweight adolescents: The potential impact of familial factors. *J. Adolesc Health* **2002**, *31*, 145–153. [CrossRef]
23. Craeynest, M.; Crombez, G.; Deforche, B.; Tanghe, A.; De Bourdeaudhuij, I. The role of implicit attitudes towards food and physical activity in the treatment of youth obesity. *Eat. Behav.* **2008**, *9*, 41–51. [CrossRef]
24. Martens, M.K.; van Assema, P.; Brug, J. Why do adolescents eat what they eat? Personal and social environmental predictors of fruit, snack and breakfast consumption among 12-14-year-old Dutch students. *Public Health Nutr.* **2005**, *8*, 1258–1265. [CrossRef]
25. McCrory, M.A. Meal skipping and variables related to energy balance in adults: A brief review, with emphasis on the breakfast meal. *Physiol. Behav.* **2014**, *134*, 51–54. [CrossRef] [PubMed]
26. Timlin, M.T.; Pereira, M.A.; Story, M.; Neumark-Sztainer, D. Breakfast eating and weight change in a 5-year prospective analysis of adolescents: Project EAT (Eating Among Teens). *Pediatrics* **2008**, *121*, e638–e645. [CrossRef]
27. Wojcicki, J.M.; Schwartz, N.; Jimenez-Cruz, A.; Bacardi-Gascon, M.; Heyman, M.B. Acculturation, dietary practices and risk for childhood obesity in an ethnically heterogeneous population of Latino school children in the San Francisco bay area. *J. Immigr. Minor. Health* **2012**, *14*, 533–539. [CrossRef]
28. Niemeier, H.M.; Raynor, H.A.; Lloyd-Richardson, E.E.; Rogers, M.L.; Wing, R.R. Fast food consumption and breakfast skipping: Predictors of weight gain from adolescence to adulthood in a nationally representative sample. *J. Adolesc. Health* **2006**, *39*, 842–849. [CrossRef] [PubMed]
29. Thompson, O.M.; Ballew, C.; Resnicow, K.; Must, A.; Bandini, L.G.; Cyr, H.; Dietz, W.H. Food purchased away from home as a predictor of change in BMI z-score among girls. *Int. J. Obes. Relat. Metab. Disord.* **2004**, *28*, 282–289. [CrossRef]
30. Lazarou, C.; Kalavana, T.; Matalas, A.L. The influence of parents' dietary beliefs and behaviours on children's dietary beliefs and behaviours. The CYKIDS study. *Appetite* **2008**, *51*, 690–696. [CrossRef]
31. Pugliese, J.; Tinsley, B. Parental socialization of child and adolescent physical activity: A meta-analysis. *J. Fam. Psychol.* **2007**, *21*, 331–343. [CrossRef] [PubMed]
32. Cadenas-Sanchez, C.; Henriksson, P.; Henriksson, H.; Delisle Nystrom, C.; Pomeroy, J.; Ruiz, J.R.; Ortega, F.B.; Lof, M. Parental body mass index and its association with body composition, physical fitness and lifestyle factors in their 4-year-old children: Results from the MINISTOP trial. *Eur. J. Clin. Nutr.* **2017**, *71*, 1200–1205. [CrossRef]
33. Zarychta, K.; Mullan, B.; Luszczynska, A. It doesn't matter what they say, it matters how they behave: Parental influences and changes in body mass among overweight and obese adolescents. *Appetite* **2016**, *96*, 47–55. [CrossRef] [PubMed]
34. Campbell, K.J.; Crawford, D.A.; Salmon, J.; Carver, A.; Garnett, S.P.; Baur, L.A. Associations between the home food environment and obesity-promoting eating behaviors in adolescence. *Obesity* **2007**, *15*, 719–730. [CrossRef]
35. Hanson, N.I.; Neumark-Sztainer, D.; Eisenberg, M.E.; Story, M.; Wall, M. Associations between parental report of the home food environment and adolescent intakes of fruits, vegetables and dairy foods. *Public Health Nutr.* **2005**, *8*, 77–85. [CrossRef]
36. Moore, L.C.; Harris, C.V.; Bradlyn, A.S. Exploring the relationship between parental concern and the management of childhood obesity. *Matern Child. Health J.* **2012**, *16*, 902–908. [CrossRef]
37. Flores, Y.N.; Shaibi, G.Q.; Morales, L.S.; Salmeron, J.; Skalicky, A.M.; Edwards, T.C.; Gallegos-Carrillo, K.; Patrick, D.L. Perceived health status and cardiometabolic risk among a sample of youth in Mexico. *Qual. Life Res.* **2015**, *24*, 1887–1897. [CrossRef]
38. Edwards, T.C.; Patrick, D.L.; Skalicky, A.M.; Huang, Y.; Hobby, A.D. Perceived body shape, standardized body-mass index, and weight-specific quality of life of African-American, Caucasian, and Mexican-American adolescents. *Qual. Life Res.* **2012**, *21*, 1101–1107. [CrossRef]
39. Kovacs, M. *Children's Depression Inventory (CDI)*; Multi-Health Systems, Inc.: North Tonawanda, NY, USA, 1992.
40. Swallen, K.C.; Reither, E.N.; Haas, S.A.; Meier, A.M. Overweight, obesity, and health-related quality of life among adolescents: The National Longitudinal Study of Adolescent Health. *Pediatrics* **2005**, *115*, 340–347. [CrossRef] [PubMed]

41. Molina, C.S.; Gómez, J.R.; Pastrana, M.C. Psychometric properties of the Spanish-language child depression inventory with Hispanic children who are secondary victims of domestic violence. *Adolescence* **2012**, *44*, 133–148.
42. Segura Camacho, S.; Posada Gomez, S.; Ospina, M.L.; Ospina Gomez, H. Inventory standardization of children depression scale for adolescents aged 12 and 17 years of age, in the Municipality of Sabaneta, Department Antioquia, Colombia. *Int. J. Psychol. Res.* **2010**, *3*, 63–73. [CrossRef]
43. Salum, G.A.; Patrick, D.L.; Isolan, L.R.; Manfro, G.G.; Fleck, M.P. Youth Quality of Life Instrument-Research version (YQOL-R): Psychometric properties in a community sample. *J. Pediatr.* **2012**, *88*, 443–448. [CrossRef]
44. Patrick, D.L.; Edwards, T.C.; Topolski, T.D. Adolescent quality of life, part II: Initial validation of a new instrument. *J. Adolesc.* **2002**, *25*, 287–300. [CrossRef]
45. van Strien, T.; Frijters, J.E.R.; Bergers, G.P.A.; Defares, P.B. The Dutch Eating Behavior Questionnaire (DEBQ) for Assessment of Restrained, Emotional, and External Eating Behavior. *Int. J. Eat. Disord.* **1986**, *5*, 295–315. [CrossRef]
46. McCabe, M.P.; Ricciardelli, L.A. Sociocultural influences on body image and body changes among adolescent boys and girls. *J. Soc. Psychol.* **2003**, *143*, 5–26. [CrossRef] [PubMed]
47. Stunkard, A.J.; Sorensen, T.; Schulsinger, F. Use of the Danish Adoption Register for the study of obesity and thinness. *Res. Publ. Assoc. Res. Nerv. Ment. Dis.* **1983**, *60*, 115–120.
48. Song, A.Y.; Rubin, J.P.; Thomas, V.; Dudas, J.R.; Marra, K.G.; Fernstrom, M.H. Body image and quality of life in post massive weight loss body contouring patients. *Obesity* **2006**, *14*, 1626–1636. [CrossRef]
49. Washington State Department of Health. Healthy Youth Survey. Available online: http://www.doh.wa.gov/DataandStatisticalReports/HealthBehaviors/HealthyYouthSurvey/PastSurveys (accessed on 21 December 2018).
50. Killen, J.D.; Taylor, C.B.; Hayward, C.; Wilson, D.M.; Haydel, K.F.; Hammer, L.D.; Simmonds, B.; Robinson, T.N.; Litt, I.; Varady, A.; et al. Pursuit of thinness and onset of eating disorder symptoms in a community sample of adolescent girls: A three-year prospective analysis. *Int. J. Eat. Disord.* **1994**, *16*, 227–238. [CrossRef]
51. Centers for Disease Control and Prevention. National Center for Health Statistics. National Health and Nutrition Examination Survey. Available online: https://wwwn.cdc.gov/nchs/nhanes/Default.aspx (accessed on 21 December 2018).
52. Grunbaum, J.A.; Kann, L.; Kinchen, S.A.; Williams, B.; Ross, J.G.; Lowry, R.; Kolbe, L. Youth risk behavior surveillance–United States, 2001. *Mmwr. Surveill Summ.* **2002**, *51*, 1–62. [CrossRef] [PubMed]
53. de Onis, M.; Onyango, A.W.; Borghi, E.; Siyam, A.; Nishida, C.; Siekmann, J. Development of a WHO growth reference for school-aged children and adolescents. *Bull. World Health Organ.* **2007**, *85*, 660–667. [CrossRef] [PubMed]
54. de Wit, L.; Luppino, F.; van Straten, A.; Penninx, B.; Zitman, F.; Cuijpers, P. Depression and obesity: A meta-analysis of community-based studies. *Psychiatry Res.* **2010**, *178*, 230–235. [CrossRef]
55. Burke, S.M.; Majid, D.S.A.; Manzouri, A.H.; Moody, T.; Feusner, J.D.; Savic, I. Sex differences in own and other body perception. *Hum. Brain Mapp.* **2019**, *40*, 474–488. [CrossRef]
56. Whitaker, R.C.; Wright, J.A.; Pepe, M.S.; Seidel, K.D.; Dietz, W.H. Predicting obesity in young adulthood from childhood and parental obesity. *N. Engl. J. Med.* **1997**, *337*, 869–873. [CrossRef] [PubMed]

International Journal of
Environmental Research and Public Health

MDPI

Article

Improving Adolescents' Subjective Well-Being, Trait Emotional Intelligence and Social Anxiety through a Programme Based on the Sport Education Model

Pablo Luna, Jerónimo Guerrero and Javier Cejudo *

Department of Psychology, Faculty of Education of Ciudad Real, Universidad de Castilla-La Mancha, Ronda de Calatrava 3, 13071 Ciudad Real, Spain; pablo.luna@uclm.es (P.L.); jeronimo.guerrero@alu.uclm.es (J.G.)

* Correspondence: manueljavier.cejudo@uclm.es; Tel.: +34-926295300 (ext. 3215)

Received: 4 April 2019; Accepted: 21 May 2019; Published: 23 May 2019

Abstract: This study aimed to evaluate the impact of a physical-sport education pilot programme on adolescents' subjective well-being (health-related quality of life, positive affect and negative affect), trait emotional intelligence and social anxiety. The programme was based on the pedagogical sport education model within a quality physical education framework, and approached from the perspective of social and emotional learning. Participants were 113 compulsory secondary education students aged 12–15 years that were assigned to a control group (n = 44) and an experimental group (n = 69). A quasi-experimental design with repeated pre-test and post-test measures was used. Bonferroni correction was applied for multiple comparisons. The preliminary results obtained in this investigation revealed that the physical-sport education pilot programme promoted significant improvements in a specific indicator of subjective well-being and trait emotional intelligence in the experimental group. These encouraging findings support the pedagogical efficiency of the programme with regard to the programme aim. The findings also highlight the feasibility and appropriateness of the programme in terms of an innovative teaching proposal.

Keywords: physical education; social and emotional learning; sport education model; subjective well-being; trait emotional intelligence; social anxiety

1. Introduction

In the education field, physical education is a subject that can contribute to improving the well-being and health of children and adolescents. The concept of quality physical education, which is understood as an interrelated system of inclusive and active teaching and learning, must be considered a key framework for integral approaches (i.e., education and health) [1]. It can also be seen as a physically active teaching and learning experience that can positively impact students' psychomotor abilities, cognitive comprehension and social and affective aptitudes [2]. Moreover, in our view, quality physical education could be grouped in the social and emotional learning category.

Quality physical education aims to achieve an integral education commitment [3,4] that allows students to be physically literate [5,6]. Physical literacy is the pillar of quality physical education, and can be defined as the motivation and cognitive, physical and affective competence necessary to encourage and preserve an active attitude in life, enabling a positive development of the aptitude to achieve, understand and use decisions about one's health efficiently [7]. Students who are physically literate intrinsically value their own psychomotor capabilities, as well as the contribution of these abilities to well-being and health [6].

The connection between health and physical activity is widely accepted [8,9]. However, the effects of physical activity on health in the educational context should be deepened through experimental studies. Designing teaching and learning processes in this area will support physical, psychological, emotional and social development [10].

The synergy between the practice of physical-sport activity together with physical and psychological health is a gradually growing interest area for education researchers [11–14]. Moreover, different investigations in the education framework of the evolution of quality physical education have emphasised the need for methodological change [4,15]. Several pedagogical models share the same features [16].

This study is based on quality physical education, manifested through a specific sport education model [17]. Sport education is a pedagogical model that uses essential features of sports (seasons, competitions, membership, data register, culminating event and festivity), and aims to achieve the inclusive goal of all students living real and meaningful sport experiences in physical education. In addition, this model aspires to develop competence, enthusiasm and a physical-sport culture in students [18].

The pedagogical potential of the sport education model, if correctly implemented [19], results in benefits at a physical level [20–22]. Similarly, it has been shown to have a positive impact on psychological variables in adolescents. These positive benefits include: basic psychological needs [23]; improvement in competence [24] and the feeling of belonging to a group [25]; decrease in attitudes towards violence and improvements in social responsibility and participants' relationships [25,26]; more self-determined behaviour [27]; improvements in friendship and sport goals [24]; decrease in aggressive behaviour and improvements in friendship relationships [25,28]; positive changes in the perception of the social climate [29]; improvement in social relationships [30]; improvement in trait emotional intelligence and motivational mediators [31]; and improvement in sport culture and enthusiasm. However, no benefits in terms of life satisfaction have been found [32]. By making use of sport, these studies provide evidence of a meaningful and positive impact on the psychological and physical development of the school-age population [11,33].

Health is generally determined by several physical–biological, psychological and social indicators [34]. Therefore, good health is a fundamental dimension in personal and social progress, and an important sphere in quality of life [35]. Approached from the perspective of positive health, a state of wellness encourages individuals to reach complete social and psychological development [36].

The World Health Organization (WHO) aims to promote physical and psychological health [34] that supports a good quality of life [37]. The construct of subjective well-being is among the factors that affect health. Consequently, research on the influence of subjective well-being in different social and educational contexts has received increasing attention in recent decades [38]. Subjective well-being comprises two main factors: a cognitive aspect (satisfaction with one's own life) and an affective aspect (positive and negative affect) [39,40]. The cognitive side of well-being reflects the assessment of how individuals process information in their lives [41]. The affective side of well-being implies a hedonistic individual balance; that is, how often individuals experience positive and negative emotions [39,40].

Recent research has focused on studying the effects of positive psychological variables on personal and social development [42]. These studies have been categorised as positive psychology [43]. Emotional intelligence is a positive variable that currently has broad support because of its close connections to subjective well-being and physical and mental health [38,44,45].

Variables such as social anxiety have a negative effect on subjective well-being [46]. In this sense, social anxiety can be defined as a person's constant fear of one or more social or performance situations, in which they are exposed to unknown persons or the possible scrutiny of other people [47]. Social anxiety has a negative impact on subjective well-being in adolescents because of the anguish individuals may feel [48], which in turn may negatively affect the quality of their interpersonal relationships [49,50].

We consider that education should promote social and emotional learning, which the WHO defines as a heterogeneous set of life skills, as this is a potential factor that supports and encourages mental health [51]. Scholars in favour of this teaching proposal argue that emotional education may also promote public health [52,53], because it's ultimate goal is improvement of the general quality of health and well-being in citizens. In the school context, many researchers claim that a key purpose of education is to improve peoples' lives so that they can reach an optimal degree of personal happiness and well-being in adulthood [54]. This suggests that a healthy pedagogical and psychological school environment may facilitate students' positive adjustment; therefore, such an environment is essential for the development of well-being in children and adolescents [55].

The theoretical and practical justification for this study was rooted in the work of various authors who developed educational interventions based on the sport education model and recommended that further research should evaluate the impact of this model on the promotion of optimal personal and social development [25,26,56,57]. In this sense, Metzler's [16] contributions are very relevant, stating that a sports model teaching program is mainly focused on different domains [16]: affective, cognitive and motor. In line with this statement, our study focuses on the affective domain. We agree with several previous authors [39,40] that subjective well-being is a key variable that influences balanced personal and social development. In addition, the existing relationship between subjective well-being and trait emotional intelligence suggests that it is necessary to further explore this topic. Social anxiety generates inappropriate social relationships in adolescents [46,48]. Given the positive effects of the sport education model on social relationships [30], it is possible that such interventions may reduce social anxiety.

This study aimed to evaluate the impact of a pilot programme based on the sport education model on the three variables: subjective well-being, trait emotional intelligence and social anxiety. The hypotheses focused on the assumptions that the programme will result in improvements in our participants' subjective well-being (Hypothesis 1), trait emotional intelligence (Hypothesis 2) and social anxiety (Hypothesis 3).

2. Methods

2.1. Participants

This study used non-probability incidental or accessibility sampling. The sample comprised 113 students in compulsory secondary education aged 12–15 years (mean age (*M*) = 13.82 years, standard deviation (*SD*) = 0.79 years). The research was conducted in a state school with students from five class groups. The control group comprised 44 students (two class groups) and the experimental group included 69 students (three class groups). The experimental and control group assignment was based on a cluster-randomised controlled trial. The gender distribution was 64 (57%) boys and 49 (43%) girls (Table 1).

The main inclusion criterion was parental consent. The exclusion criteria were: (a) attending less than 80% of the sessions of the intervention programme (less than 13 sessions); (b) students with special educational needs associated with intellectual disability; and (c) students that were removed from school for disciplinary reasons.

Table 1. Sex and age of the sample.

		n	%
Sex	Male	64	57
	Female	49	43
Age	12	42	37
	13	45	40
	14	23	20
	15	3	3

2.2. Procedure

We requested the collaboration of the educational centre (Spain) in this study. The management board of the participating school was contacted to obtain their approval and authorisation for the study. Permission was also obtained from the families of the participating students, and from the teaching staff and school council. This study respected the relevant ethical values and guaranteed participants' confidentiality and anonymity. In addition, this study was developed in accordance with the Declaration of Helsinki regarding human experimentation. The study procedures were conducted in accordance with the Universidad de Castilla-La Mancha code of ethics.

This study used a quasi-experimental design with repeated pre-test and post-test measures and a control group. The study was conducted over three stages. First, before the intervention began, the assessment instruments (pre-test evaluation) were handed out for completion in the first 20 minutes of two sessions, to avoid burdening the students. Next, the programme based on the sport education model was implemented. The programme sessions took place during the second term of the school year. Finally, the assessment instruments were completed a second time (post-test evaluation). The independent variable was the intervention programme, and the dependent variables were subjective well-being, trait emotional intelligence and social anxiety.

2.3. Measures

Four evaluation instruments were used to assess the variables in this study, under the psychometric parameters of reliability and validity.

The Kidscreen-10 Index [58] was used to assess subjective health-related quality of life and well-being. This 10-item scale was designed for children and adolescents aged 8–18 years. Each item has five response options, ranging from 'Never' to 'Always' or 'Not at all' to 'Extremely'. The 10 items cover: affective symptoms of depressed mood; cognitive symptoms of disturbed concentration; psychovegetative aspects of vitality, energy and feeling well; and psychosocial aspects correlated with mental health, such as the ability to experience fun with friends or getting along well with others at school. The adapted version of the questionnaire used in this study has adequate internal consistency reliability (*Cronbach's alpha* = 0.82) and test-retest stability (r = 0.73; *ICC* = 0.72) [59].

The Positive and Negative Affect Schedule [60] was used to assess participants' positive and negative affect. The Spanish version of this scale for children and adolescents was validated by Sandín [61]. This scale comprises 20 items on two dimensions: positive affect and negative affect. Each subscale contains 10 items. The questionnaire is completed by participants based on the way they normally feel and behave. The scale has three response options: 'Never' = 1, 'Sometimes' = 2, and 'Many times' = 3.

We used the Trait Emotional Intelligence Questionnaire Adolescents Short Form (TEIQue-ASF) [62] (adapted into Spanish in its abridged version for teenagers by Ferrando and Serra [63]) to evaluate trait emotional intelligence based on the theoretical model of Petrides and Furnham [64]. The 30 items that make up the TEIQue-ASF are scored on a 7-point Likert scale (1 = 'Completely disagree' to 7 = 'Completely agree'). The general emotional intelligence score of the total scale is obtained by summing the 30 items.

Finally, participants completed the Social Anxiety Scale for Adolescents (SAS-A) [65]. The SAS-A comprises 22 items; 18 items are self-descriptive and four are distracting elements that are not taken into account for the score. The SAS-A contains three subscales: (a) fear of negative evaluation (eight items), (b) anxiety and social avoidance before strangers or new social situations (six items) and (c) anxiety and social avoidance in social situations in general (four items). Responses are on a 5-point Likert-type scale from 1 ('Never') to 5 ('Always'). In addition, a global index of social anxiety (SAS-T) is obtained by summing the scores for the items (excluding neutral items). High scores reflect high levels of social anxiety [65]. The scale was adapted to the Spanish population by Olivares, Ruiz, Hidalgo, García-López, Rosa and Piqueras [66]. Only the SAS-T score was used in this study.

2.4. Intervention Programme

The physical-sport programme was completed following the sport education model structure [17]: (1) *season*: lengthy didactic units; (2) *membership*: development of a team spirit and cooperation; (3) *regular competition*: showing technical–tactical abilities; (4) *data register*: giving evidence of and analysing the process that has been followed; and (5) *festivity*: a festive atmosphere. This highlighted other important education aspects such as: cooperative learning; autonomy and personal initiative; positive interdependence; and self-management of responsibility roles in conflict resolution (i.e., referee and coach). This helped to make the sport experience more real and positive, including how students transferred responsibilities by means of organisation roles (i.e., referee and scorer), team roles (i.e., coach and physical trainer) and how sport content was modified when adapted to the students [17].

Hastie and Casey's guidelines were followed for the design and validation of the programme [19] (p. 423): (a) thoroughly detailed curricular elements; (b) precise certification of the applied model; and (c) an in-depth explanation of the context of the programme. The intervention programme was implemented in the experimental group following sequencing of content and activities in three stages (initial, intermediate and final) over 16 sessions (Table 2).

Table 2. Sequencing of stages and activity sessions in the intervention programme.

Stage	Session	Sport Education Model
Initial (Theoretical Sessions)	1–2	Introduction and presentation of the Sport Education Model with digital and audio-visual support (ICT). Presentation and distribution of learning resources. Division and organisation of classroom groups in teams (assignment of team names with a didactic and cross curricular theme). Distribution and selection of responsibility roles.
	3	Explanation for the self-design of learning resources on digital format (ICT). Selection and assignment of anthems, badges, mascots and t-shirts representing a team.
Intermediate (Practical Sessions)	4–7	Practical implementation of the roles of each member of the teams. Learning of technical-tactical elements and abilities: kicking-off, catching, moving, throwing, defence and attack. Learning game rules.
	8–9	Warming-up, training and friendly matches. Meetings for comprehension and reflection with intervention of the responsibility roles.
	10–14	Regular stage competition (Round Robin).
Final (Practical Sessions)	15–16	Inter-class groups final competitions (final matches with class groups), final event, giving awards and diplomas.

This pilot programme was developed to reflect the teaching hours of the physical education subject, which covers 16 55 minute sessions (2–3 sessions per week for 6 weeks). The total duration was considered sufficient to analyse the possible effects of the programme on the dependent variables, as indicated by previous research [67].

The educational intervention applied to the experimental group consisted of a didactic unit that used an alternative sport, called ringo [68,69]. Ringo is an alternative, modified and reduced sport of divided court and net. It is played with a hoop (ringo) and a volleyball net. The objective is to score when the ringo falls on the opposite court (Figure 1). In this pilot programme, the application of an alternative sport that was novel and unknown to students meant that everyone started with the same theoretical and practical sports knowledge, and there were few initial differences in their levels of technical–tactical sports skill. Alternative sports are characterised by being motivating, cooperative, socialising and adapted to participants' characteristics. The selection and organisation of teams (five teams per classroom) was developed by drawing lots. In addition, different responsibility roles were assigned to participating students: player; referee; coach–captain; physical trainer; person responsible for statistics and reports; and member of the discipline and organisation committee. An essential rule in the development of the pilot programme was that all students would actively participate in the programme with the assignment of two roles (player role and another responsibility role). The pilot programme also used various learning and curricular resources (self-designed portfolio, worksheets and reports) that had been used by other authors [70].

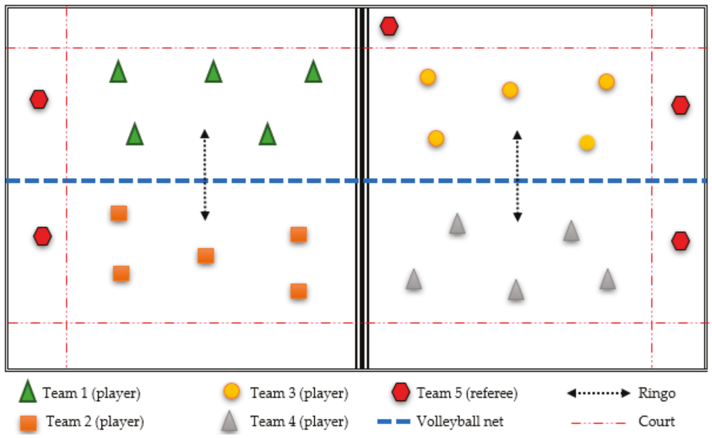


Figure 1. Practical session of the intervention programme.

For the control group, a didactic unit of traditional collective sport with a conventional teaching style was developed [71]. This traditional teaching model aimed to improve students' technical motor skills only. In the teaching–learning process, the teacher assumed a managerial role and the students adopted passive individual roles limited to following the directive instructions of the teacher. This intervention consisted of 12 55 minute sessions (two sessions per week for 6 weeks). The first nine sessions were aimed at learning the technical fundamentals of basketball (pot, dribbling, passing, throwing and receiving) through a task assignment teaching style [71]. These traditional sports sessions were based on a 10 minute warm-up, 40 minute main session that included explanations and basketball practice and a 5-minute warm down in which stretching was performed. During these sessions, all tasks were directed by the teacher without students' participation. The last three sessions were dedicated to team competition.

Two compulsory secondary education teachers, both with advanced degrees in Sports Science participated in this research. The first teacher (with a Master's of Science in Psychology) participated in the design and implementation of the pilot programme. The second teacher developed an intervention based on the traditional model. Both teachers received a 10 hour training course on the specific theoretical and practical aspects of each teaching model. In addition, supervision and tutoring was provided by a researcher expert in the sports education model and a researcher expert in the traditional education model. This tutoring consisted of: (a) session-by-session analysis during the intervention programmes; (b) telephone conversations and emails to resolve doubts, concerns and problems; and (c) weekly visits to the teaching centre. In these visits, the experts visited the centre randomly, without prior notice, with the objectives of: verifying that there were no gaps between what was planned and what was implemented, and checking that the teaching models were applied with all of their characteristics.

2.5. Data Analysis

After fulfilment of the requirements of normality and homoscedasticity was verified, we examined the distribution of the data for a univariate normality analysis. The results showed asymmetry and kurtosis values lower than 1.2. Next, we calculated the reliability coefficients (Cronbach's alpha, composite reliability, average variance extracted and McDonald's omega coefficient) to obtain reliability evidence. Then, to determine the impact of the programme, descriptive analyses (mean and SD) and analyses of variance (ANOVA) were performed with the scores collected in the pre-test stage. Subsequently, descriptive analyses and analyses of covariance (ANCOVA) were used with post-test scores to determine the impact of the programme on each of the variables. Bonferroni correction was applied for multiple comparisons. For all analyses, a p-value <0.05 was considered to indicate statistical significance. After application of Bonferroni correction, a p-value <0.012 was considered significant.

The effect size (μ^2) of the differences was calculated using partial eta-squared [72]. The effect size was analysed based on four ranges: 0–0.009, negligible; 0.010–0.089, low-effect size; 0.090–0.249, medium-effect size; and >0.250, big-effect size [72]. The data were analysed with SPSS version 24.0 (IBM Corp., Armonk, NY, USA).

3. Results

3.1. Reliability

In this study, we used well-established measures with appropriate psychometric properties (Table 3).

Table 3. Reliability evidence of the instruments used (n = 113).

Measures	α	*CR*	*AVE*	Ω
KIDSCREEN	0.91	0.89	0.674	0.92
PANASN-PA	0.70	0.77	0.502	0.72
PANASN-NA	0.74	0.76	0.519	0.77
TEIQue-ASF	0.71	0.70	0.503	0.79
SAS-T	0.85	0.80	0.687	0.87

Notes: α = Cronbach's alpha; *CR* = composite reliability, *AVE* = average variance extracted; Ω = McDonald's omega index.

3.2. Effects of the Programme

The pre-test MANOVA results did not reveal statistically significant differences between the groups prior to the intervention, Wilks' Lambda, $\Lambda = 0.571$, F (5, 108) = 0.739, $p = 0.333$, with a small effect size ($\eta 2 = 0.062$, $r = 0.11$).

The ANOVA using the pre-test scores (Table 4) revealed no statistically significant differences in any of the dependent variables before the programme began, except for a significantly higher score for trait emotional intelligence in the experimental group compared with the control group. The size of the effect was low for trait emotional intelligence ($\mu^2 = 0.009$). Applying Bonferroni correction showed no significant differences in any of the variables.

Results from the pre-test–post-test MANCOVA revealed significant differences between the two conditions, Wilks' Lambda, $\Lambda = 0.899$, F (5, 108) = 5.295, $p = 0.003$, with an average effect size ($\eta 2 = 0.267$, $r = 0.32$).

Next, we performed ANCOVA for the dependent variables using the post-test scores. To assess the magnitude of these differences, the effect size for each variable was calculated by partial eta-squared (Table 4).

3.2.1. Effects on Subjective Well-Being

There was no significant improvement in post-test health-related quality of life in the experimental group (Table 4). The experimental group did not show a significant increase in positive affect scores after testing. We confirmed a significant decrease in post-test negative affect scores in the experimental group (Table 4), with a medium effect size ($\mu^2 = 0.123$; partial eta-squared).

3.2.2. Effects on Trait Emotional Intelligence

The analysis revealed significant improvements in trait emotional intelligence in the experimental group after the programme, with a medium effect size (μ^2 = 0.241) (Table 4).

3.2.3. Effects on Social Anxiety

There were no significant differences between the experimental and the control groups in SAS-T scores (Table 4).

Table 4. Mean, standard deviation, analysis of variance, analysis of covariance and effect size for differences in means (partial eta-squared) as a function of the experimental and control groups at pre-test and post-test.

Measures	PRE-TEST					POST-TEST				
	EXPERIMENTAL	CONTROL				EXPERIMENTAL	CONTROL			
	Mean (SD)	*Mean (SD)*	*F*	*p*	μ^2	*Mean (SD)*	*Mean (SD)*	*F*	*p*	μ^2
KIDSCREEN										
HRQL	35.18 (5.84)	35.09 (6.01)	1.414	0.697	0.001	36.94 (6.14)	35.04 (5.99)	1.975	0.018	0.072
PANASN										
PA	21.43 (3.90)	20.02 (4.21)	3.293	0.07	0.008	21.95 (2.78)	20.33 (2.88)	5.438	0.017	0.144
NA	11.23 (3.54)	11.58 (3.58)	0.251	0.62	0.004	9.96 (2.95)	11.62 (3.86)	7.044	0.010	0.123
TEIQUE-ASF										
TEI	4.82 (0.60)	4.58 (0.64)	4.368	0.04	0.009	5.02 (0.63)	4.52 (0.53)	16.394	0.000	0.241
SAS										
SAS-T	2.61 (0.67)	2.64 (0.68)	0.040	0.84	−0.001	2.38 (0.63)	2.58 (0.62)	3.419	0.062	0.014

Note: HRQL = health-related quality of life; PA = positive affect; NA = negative affect; TEI = trait emotional intelligence; SAS-T = Total Social Anxiety Scale; SD = standard deviation.

4. Discussion

This study evaluated the impact of a pilot programme based on the sports education model on compulsory secondary education students' subjective well-being, trait emotional intelligence and social anxiety.

The preliminary results obtained in this study revealed significant improvement of a specific indicator of subjective well-being (NA) in the experimental group after the pilot programme. The experimental group did not show a significant improvement in health-related quality of life when compared with the control group. Our results are consistent with the findings reported in other studies [32], which did not confirm significant benefits for life satisfaction among adolescents following a sport education-based experience. Our findings also indicated that the programme showed a significant decrease in negative affect, improving an indicator of the affective component of subjective well-being (i.e., a decrease in negative emotions) [39,40]. These results partially verify Hypothesis 1, and highlight the importance of further research in this context.

This is consistent with the findings reported in other studies [32], as previous studies established a connection between physical activity and subjective well-being [73]. These findings are also consistent with research that argues that active, inclusive and effective teaching and learning processes applied within a quality physical education framework fosters a motivating school climate in affective and psychological terms [1,74]. Pedagogical and methodological aspects highlighted by this intervention pilot programme (e.g., cooperative learning, a feeling of membership to a team, positive interdependence and self-management or autonomy/use of responsibility roles) could have influenced these results. Furthermore, a motivating school context, enabled by the implementation of the sport education model [31], may also strengthen affective bonding in adolescents [12].

Significant improvement in trait emotional intelligence was observed in the experimental group after the programme, which confirmed Hypothesis 2. These results are also consistent with those reported by other authors [31]. The relationship between trait emotional intelligence and subjective well-being [38,45], as well as that between trait emotional intelligence and physical and psychological health [44], may trigger these improvements in adolescents. This indicates that good trait emotional intelligence promotes positive emotional states and a reduction of negative moods, thereby positively impacting well-being and health [45].

We found no significant improvement in students' social anxiety, meaning that we could not confirm Hypothesis 3. Although this is similar to the results obtained by different authors for social relationships variables [23,24], our results contradicted the findings of other studies [25,26,28,30]. Further research on the effects of the sport education model is therefore necessary, especially given the theoretical specificity of social anxiety and its incidence in social relationships among adolescents. In addition, social anxiety can present opposing consequences. It can have positive effects on social relationships for some individuals, whereas it can have negative effects on others, characterised by anguish and social avoidance [65].

However, in our opinion this study has been very exhaustive in the evaluation methodology of the intervention program (including the Bonferroni corrections). In this sense, we have not found any study on the effectiveness of the sport education model that uses these statistical corrections. This fact could be influencing the comparison of our results with those obtained in other similar researches on the sport education model as a teaching model.

Despite these promising results, this study had some limitations. First, the sampling procedure was chosen for reasons of convenience and not by random procedures. However, allocating students to either the experimental or control group was performed randomly based on the class group to which they belonged. Second, it would be necessary to increase the sample size to minimise potential biases in the results and increase the generalisability. Third, the instruments used were self-reported, and the results might have been influenced by bias related to social desirability in adolescents. It would be necessary to use high-performance tests or hetero-evaluation to minimise such bias. Similarly, differences in the trait emotional intelligence pre-test scores between the experimental and control groups might have had an impact on our results. Finally, it is necessary to highlight the difficulties encountered when following all of the recommendations for the implementation of the sport education model [19]. Similarly, it would be necessary to include session analysis procedure in order to evaluate if the main principles of the model were followed by the teachers [75].

Several aspects can be suggested regarding future lines of investigation, such as increasing the number of participants and diversifying their sociocultural background. It may also be worthwhile analysing the impact of the programme on other variables, such as academic performance and social and school adjustment. Similarly, to study the effects on depression with the use of biological correlates (HPA markers, cortisol immunitarian parameters, etc.). In addition, it would be interesting to conduct a follow-up evaluation to assess the sustainability of the effects of the programme.

This study presents innovative contributions at both theoretical and practical levels. The theoretical contribution is related to fact that the lack of physical activity can have a harmful effect on individual's health and is currently an important public health concern [76]. In this respect, United Nations Educational, Scientific and Cultural Organization (UNESCO) [1] emphasised the importance of fostering and promoting active behaviours [6] in all contexts, especially at schools. Consequently, it should be noted that there is a positive connection between health and physical activity: sedentarism is a major risk factor for mortality, which gives rise to concern about the prevalence of sedentarism and socio-educative patterns of inactivity, especially in school contexts. At a practical level, our findings may help teaching staff in their tasks at school, as they provide a tool that may be used in teaching practice. In addition, the findings open up interesting fields of research in terms of the application of sport education, especially in terms of its impact on psychological variables.

5. Conclusions

In conclusion, our findings suggest that the pilot programme stimulated some improvement in adolescents' subjective well-being and trait emotional intelligence, but did not impact social anxiety. Therefore, on the basis of quality physical education and the social and emotional learning approach, the implementation of such programmes is recommended given the possible psychological benefits for adolescents in the educational context. A commitment to sports and other physical-sport activity options within a quality physical education framework, efficiently applied by means of relevant pedagogical models (such as sport education), may play an important role in students' integral development [77,78].

Author Contributions: P.L., J.G. and J.C. conceived and designed the experiments; P.L., J.G. and J.C. performed the experiments; P.L., J.G. and J.C. analysed the data; P.L., J.G. and J.C. contributed reagents/materials/analysis tools; and P.L., J.G. and J.C. wrote the paper.

Funding: This research received no external funding.

Acknowledgments: We thank Edanz Group (www.edanzediting.com/ac) for editing a draft of this manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. United Nations Educational, Scientific and Cultural Organization (UNESCO). *Educación Física de Calidad: Guía Para los Responsables Politicos*; UNESCO: Paris, France, 2015. Available online: http://unesdoc.unesco.org/images/0023/002313/231340s.pdf (accessed on 4 October 2018).
2. Association for Physical Education: Health Position Paper. 2015. Available online: http://www.afpe.org.uk/physical-education/wp-content/uploads/afPE_Health_Position_Paper_Web_Version2015.pdf (accessed on 4 October 2018).
3. Penney, D.; Brooker, R.; Hay, P.; Gillespie, L. Curriculum, pedagogy and assessment: three message systems of schooling and dimensions of quality physical education. *Sport. Educ. Soc.* **2009**, *14*, 421–442. [CrossRef]
4. Le Masurier, G.; Corbin, B. Top 10 reasons for quality physical education. *Am. J. Health Educ.* **2006**, *77*. [CrossRef]
5. Keegan, R.J.; Keegan, S.L.; Daley, S.; Ordway, C.; Edwards, A. *Getting Australia Moving: Establishing a Physically Literate and Active Nation (Game Plan)*; University of Canberra: Canberra, Australia, 2013; ISBN 978-1740883719.
6. Whitehead, M. *Physical Literacy throughout the Lifecourse*; Routledge: London, UK, 2010; ISBN 978-0415487436.
7. Whitehead, M. The Concept of Physical Literacy. *Eur. J. Phys. Educ.* **2001**, *6*, 127–138. [CrossRef]
8. Hallal, P.C.; Andersen, L.B.; Bull, F.C.; Guthold, R.; Haskell, W.; Ekelund, U. Lancet Physical Activity Series Working Group. Global physical activity levels: Surveillance progress, pitfalls and prospects. *Lancet* **2012**, *380*, 247–257. [CrossRef]
9. World Health Organization (WHO). *Recomendaciones Mundiales Sobre Actividad Física Para la Salud*; World Health Organization: Geneva, Switzerland, 1948. Available online: http://apps.who.int/iris/bitstream/handle/10665/44441/9789243599977_spa.pdf;jsessionid=D0530DDFA8910637132BC6D75745E29C?sequence=1 (accessed on 5 October 2018).
10. National Association for Sport and Physical Education (NASPE) y American Heart Association (AHA). *Shape of the Nation Report: Status of Physical Education in the USA*; NASPE: Reston, VA, USA; AHA: Dallas, TX, USA, 2012. Available online: http://www.shapeamerica.org/advocacy/son/2012/upload/2012-Shape-of-Nation-full-report-web.pdf (accessed on 4 October 2018).
11. Armour, K.; Sandford, R.; Duncombe, R. Positive youth development and physical activity/sport interventions: Mechanisms leading to sustained impact. *Phys. Educ. Sport Pedag.* **2013**, *18*, 256–281. [CrossRef]
12. Bailey, R.; Armour, K.; Kirk, D.; Jess, M.; Pickup, I.; Sandford, R. The educational benefits claimed for physical education and school sport: An academic review. *Res. Pap. Educ.* **2009**, *24*, 1–27. [CrossRef]
13. Knowles, A.; Wallhead, T.L.; Readdy, T. Exploring the Synergy Between Sport Education and In-School Sport Participation. *J. Teach. Phys. Educ.* **2018**, *37*, 113–122. [CrossRef]
14. Spittle, M.; Byrne, K. The influence of Sport Education on student motivation in physical education. *Phys. Educ. Sport Pedag.* **2009**, *14*, 253–266. [CrossRef]
15. Decorby, K.; Halas, J.; Dixon, S.; Wintrup, L.; Janzen, H. Classroom Teachers and the Challenges of Delivering Quality Physical Education. *J. Educ. Res.* **2005**, *98*, 208–221. [CrossRef]
16. Metzler, M.W. *Instructional Models for Physical Education*, 3rd ed.; Holcomb Hathaway: Scottsdale, AZ, USA, 2011.

17. Siedentop, D.; Hastie, P.A.; Van der Mars, H. *Complete Guide to Sport Education*, 3rd ed.; Human Kinetics: Champaign, IL, USA, 2019; ISBN 978-1492562511.
18. Hastie, P.A.; Wallhead, T. Models-Based Practice in Physical Education: The Case for Sport Education. *J. Hum. Kinet.* **2019**, *35*, 390–399. [CrossRef]
19. Hastie, P.A.; Casey, A. Fidelity in Models-Based Practice Research in Sport Pedagogy: A Guide for Future Investigations. *J. Teach. Phys. Educ.* **2014**, *33*, 422–431. [CrossRef]
20. Hastie, P.; Sluder, J.B.; Buchanan, A.M.; Wadsworth, D.D. The Impact of an Obstacle Course Sport Education Season on Students' Aerobic Fitness Levels. *Res. Q. Exerc. Sport* **2009**, *80*, 788–791. [CrossRef] [PubMed]
21. Araujo, R.; Mesquita, I.; Hastie, P.; Pereira, C. Students' game performance improvements during a hybrid sport education–step-game-approach volleyball unit. *Eur. Phy. Educ. Rev.* **2016**, *22*, 185–200. [CrossRef]
22. Walh-Alexander, Z.; Chomentowski, P. Impact of a university physical conditioning sport education season on students' fitness levels. *Health Educ. J.* **2018**, *77*, 828–836. [CrossRef]
23. Méndez-Giménez, A.; Fernández-Río, J.; Méndez-Alonso, D. Sport Education model versus Traditional model: Effects on motivation and sportsmanship. *Rev. Int. Med. Cienc. Act.* **2015**, *15*, 449–466. [CrossRef]
24. Cuevas, R.; García, L.M.; Contreras, O. Influence of the Sport Education Model in the psychological basic needs. *Cuad. Psicol. Deporte* **2015**, *15*, 155–162. [CrossRef]
25. Kao, C.C. Development of Team Cohesion and Sustained Collaboration Skills with the Sport Education Model. *Sustainability* **2019**, *11*, 2348. [CrossRef]
26. Menéndez-Santurio, J.I.; Fernández-Río, J. Violencia, responsabilidad, amistad y necesidades psicológicas básicas: Efectos de un programa de Educación Deportiva y Responsabilidad Personal y Social. *Rev. Psicodidact.* **2016**, *21*, 245–260. [CrossRef]
27. Sinelnikov, O.A.; Hastie, P.A. A motivational analysis of a season of Sport Education. *Phys. Educ. Sport Pedag.* **2010**, *15*, 55–69. [CrossRef]
28. García López, L.M.; Gutiérrez, D.; González, S.; Valero, A. Cambios en la Empatía, la Asertividad y las Relaciones Sociales por la Aplicación del Modelo de Instrucción Educación Deportiva. *Rev. Psicol. Deporte* **2012**, *21*, 321–330.
29. Méndez-Giménez, A.; Martínez de Ojeda, D.; Valverde, J.J. Students and teachers assessment of conventional and self-made material: Longitudinal crossover study in Physical Education. *Retos* **2016**, *30*, 20–25.
30. Wallhead, T.L.; Garn, A.C.; Vidoni, C. Effect of a Sport Education Program on Motivation for Physical Education and Leisure-Time Physical Activity. *Res. Q. Exerc. Sport* **2014**, *85*, 478–487. [CrossRef]
31. Méndez-Giménez, A.; Martínez de Ojeda, D.; Valverde-Pérez, J.J. Emotional Intelligence and Motivational Mediators in a Season of Sport Education Mime. *Ágora Educ. Fis. Deport.* **2017**, *19*, 52–72. [CrossRef]
32. Puente-Maxera, F.; Méndez-Giménez, A.; Martínez de Ojeda, D.; Liarte, J.P. El modelo de Educación Deportiva y la orientación. Efectos en la satisfacción con la vida, las inteligencias múltiples, las necesidades psicológicas básicas y las percepciones sobre el modelo de los adolescentes. *Sport-Tk Rev. Euroam. Cienc. Deporte* **2018**, *7*, 115–128. [CrossRef]
33. Hastie, P.; Martínez de Ojeda, D.; Calderón, A. A review of research on Sport Education: 2004 to the present. *Phys. Educ. Sport Pedagog.* **2011**, *16*, 103–132. [CrossRef]
34. World Health Organization (WHO). *Estrategia Mundial Sobre Régimen Alimentario, Actividad Física y Salud*; World Health Organization: Geneva, Switzerland, 1948. Available online: https://www.who.int/dietphysicalactivity/strategy/eb11344/strategy_spanish_web.pdf (accessed on 5 October 2018).
35. World Health Organization (WHO). *Carta de Ottawa Para la Promoción de la Salud*; World Health Organization: Geneva, Switzerland, 1948. Available online: https://www.paho.org/hq/dmdocuments/2013/Carta-de-ottawa-para-la-apromocion-de-la-salud-1986-SP.pdf (accessed on 8 October 2018).
36. Fredrickson, B.L. *Positivity*; Crown: New York, NY, USA, 2009; ISBN 978-0307393746.
37. Buss, D.M. The evolution of happiness. *Am. Psychol.* **2000**, *55*. [CrossRef]
38. Sánchez-Álvarez, N.; Extremera, N.; Fernández-Berrocal, P. The relation between emotional intelligence and subjective well-being: A meta-analytic investigation. *J. Posit. Psychol.* **2015**, *11*, 276–285. [CrossRef]
39. Diener, E. Subjective Well-Being: The science of Happiness and a Proposal for a National Index. *Am. Psychol.* **2000**, *55*, 34–43. [CrossRef]
40. Diener, E.; Suh, E.M.; Lucas, R.E.; Smith, H.L. Subjective well-being: Three decades of progress. *Psychol. Bull.* **1999**, *125*, 276–302. [CrossRef]

41. Pavot, W.; Diener, E. The satisfaction with life scale and the emerging construct of life satisfaction. *J. Posit. Psychol.* **2008**, *3*, 137–152. [CrossRef]
42. Gable, S.L.; Haidt, J. What (and why) is positive psychology? *Rev. Gen. Psychol.* **2005**, *9*, 103–110. [CrossRef]
43. Seligman, M.E.P.; Csikszentmihalyi, M. Positive psychology: An introduction. *Am. Psychol.* **2000**, *55*, 5–14. [CrossRef]
44. Martins, A.; Ramalho, N.; Morin, E. A comprehensive meta-analysis of the relationship between emotional intelligence and health. *Pers. Individ. Dif.* **2010**, *49*, 554–564. [CrossRef]
45. Zeidner, M.; Matthews, G.; Roberts, R. The emotional intelligence, health, and well-being nexus: What have we learned and what have we missed? *Appl. Psychol. Health Well-Being* **2012**, *4*, 1–30. [CrossRef]
46. Wersebe, H.; Lieb, R.; Meyer, A.H.; Miche, M.; Mikoteit, T.; Imboden, C.; Hoyer, J.; Bader, K.; Hatzinger, M.; Gloster, A.T. Well-being in major depression and social phobia with and without comorbidity. *Int. J. Clin. Health Psychol.* **2018**, *18*, 201–208. [CrossRef]
47. American Psychiatric Association (APA). *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed.; American Psychiatric Publishing: Arlington, VA, USA, 2013.
48. Beidel, D.C.; Turner, S.M.; Young, B.J.; Ammerman, R.T.; Sallee, F.R.; Crosby, L. Psychopathology of Adolescent Social Phobia. *J. Psychopathol. Behav. Assess.* **2007**, *29*, 46–53. [CrossRef]
49. Inglés, C.J.; Delgado, B.; García-Fernández, J.M.; Ruiz-Esteban, C.; Díaz-Herrero, A. Sociometric Types and Social Interaction Styles in a Sample of Spanish Adolescents. *Span. J. Psychol.* **2010**, *13*, 730–740. [CrossRef]
50. Cejudo, J.; Rodrigo-Ruiz, D.; López-Delgado, M.L.; Losada, L. Emotional intelligence and its relationship with levels of social anxiety and stress in adolescents. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1073. [CrossRef]
51. World Health Organization (WHO). *Social Determinants of Mental Health*; WHO Press: Geneva, Switzerland, 2014; ISBN 978-92-4-150680-9.
52. Domitrovich, C.E.; Durlak, J.A.; Staley, K.C.; Weissberg, R.P. Social-emotional competence: An essential factor for promoting positive adjustment and reducing risk in school children. *Child. Dev.* **2017**, *88*, 408–416. [CrossRef]
53. Kimber, B.; Sandell, R.; Bremberg, S. Social and emotional training in Swedish schools for the promotion of mental health: An effectiveness study of 5 years of intervention. *Health Educ. Res.* **2008**, *23*, 931–940. [CrossRef]
54. Ecclestone, K. From emotional and psychological well-being to character education: Challenging policy discourses of behavioural science and 'vulnerability'. *Res. Pap. Educ.* **2012**, *27*, 463–480. [CrossRef]
55. Baker, J.A.; Dilly, L.J.; Aupperlee, J.L.; Patil, S.A. The developmental context of school satisfaction: Schools as psychologically healthy environments. *Sch. Psychol. Q.* **2003**, *18*. [CrossRef]
56. Araújo, R.; Mesquita, I.; Hastie, P.A. Review of the Status of Learning in Research on Sport Education: Future Research and Practice. *J. Sports Sci. Med.* **2014**, *13*, 846–858.
57. Browne, T.B.J.; Carlson, T.B.; Hastie, P.A. A comparison of rugby seasons presented in traditional and sport education formats. *Eur. Phys. Educ. Rev.* **2004**, *10*, 199–214. [CrossRef]
58. Ravens-Sieberer, U.; European, K.G. *The Kidscreen Questionnaires: Quality of Life Questionnaires for Children and Adolescents (Handbook)*; Papst Science Publisher: Lengerich, Germany, 2006; ISBN 978-3-89967-334-0.
59. Aymerich, M.; Berra, S.; Guillamón, I.; Herdman, M.; Alonso, J.; Ravens-Sieberer, U.; Rajmil, L. Desarrollo de la versión en español del KIDSCREEN, un cuestionario de calidad de vida para la población infantil y adolescente. *Gac. Sanit.* **2005**, *19*, 93–102. [CrossRef]
60. Watson, D.; Clark, L.A.; Tellegen, A. Development and validation of brief measures of positive and negative affect: The PANAS scales. *J. Pers. Soc. Psychol.* **1998**, *54*, 1063–1070. [CrossRef]
61. Sandín, B. Escalas PANAS de afecto positivo y negativo para niños y adolescentes (PANASN). *Rev. Psicopatología Psicol. Clín.* **2003**, *8*, 173–182. [CrossRef]
62. Petrides, K.V.; Sangareau, Y.; Furnham, A.; Frederickson, N. Trait Emotional Intelligence and Children's PeerRelations at School. *Soc. Dev.* **2006**, *15*, 537–547. [CrossRef]
63. Ferrando, M.; Prieto, M.D.; Almeida, L.S.; Ferrándiz, C.; Bermejo, R.; López-Pina, J.A.; Hernández, D.; Sáinz, M.; Fernández, M.C. Trait Emotional Intelligence and Academic Performance: Controlling for the Effects of IQ, Personality, and Self-Concept. *J. Psychoeduc. Assess.* **2011**, *29*, 150–159. [CrossRef]
64. Petrides, K.V.; Furnham, A. Trait emotional intelligence: Psychometric investigation with reference toestablished trait taxonomies. *Eur. J. Personal.* **2001**, *15*, 425–448. [CrossRef]

65. La Greca, A.M.; López, N. Social anxiety among adolescents: Linkages with peer relations and friendships. *J. Abnorm. Child. Psychol.* **1998**, *26*, 83–94. [CrossRef]
66. Olivares, J.; Ruiz, J.; Hidalgo, M.D.; García-López, L.J.; Rosa, A.I.; Piqueras, J.A. Social Anxiety Scale for Adolescents (SAS-A): Psychometric properties in a spanish-speaking population. *I. J. Clin. Health Psychol.* **2005**, *5*, 85–97.
67. Mahedero, P.; Calderón, A.; Arias-Estero, J.L.; Hastie, P.A.; Guarino, A.J. Effects of student skill level on knowledge, decision making, skill execution and game performance in a mini-volleyball Sport Education season. *J. Teach. Phys. Educ.* **2015**, *34*, 626–641. [CrossRef]
68. Méndez-Giménez, A. Una iniciación deportiva de calidad con materiales autoconstruidos. El ejemplo del ringo en el marco de un modelo comprensivo estructural. *Tándem* **2005**, *18*, 61–69.
69. Méndez-Giménez, A. *Modelos Actuales de Iniciación Deportiva: Unidades Didácticas Sobre Juegos y Deportes de Cancha Dividida*; Wanceulen: Sevilla, Spain, 2011; ISBN 978-84-9823-827-3.
70. García López, L.M.; Gutiérrez, D. The effects of a sport education season on empathy and assertiveness. *Phys. Educ. Sport Pedag.* **2013**, *20*, 1–16. [CrossRef]
71. Delgado-Noguera, M.A. Teaching styles of Physical Education and Sport through 40 years of professional life. *Retos* **2015**, *28*, 20–27.
72. Tabachnick, B.G.; Fidell, L.S. *Using Multivariate Statistics*, 5th ed.; Allyn and Bacon: New York, NY, USA, 2007.
73. Biddle, S.J.H.; Fox, K.R.; Boutcher, H. *Physical Activity and Psychological Wellbeing*; Routledge: London, UK, 2000; ISBN 978-0415234818.
74. Standage, M.; Gillison, F. Students' motivational responses toward school physical education and their relationship to general self-esteem and health-related quality of life. *Psychol. Sport Exerc.* **2007**, *8*, 704–721. [CrossRef]
75. Práxedes, A.; Del Villar Álvarez, F.; Moreno, A.; Gil-Arias, A.; Davids, K. Effects of a nonlinear pedagogy intervention programme on the emergent tactical behaviours of youth footballers. *Phys. Educ. Sport Peda.* **2019**, *24*, 1–12. [CrossRef]
76. Haskell, W.L.; Blair, S.N.; Hill, J.O. Physical activity: Health outcomes and importance for public health policy. *Prev. Med.* **2009**, *49*, 280–282. [CrossRef] [PubMed]
77. Eley, D.; Kir, D. Developing Citizenship through Sport: The Impact of a Sport-Based Volunteer Programme on Young Sport Leaders. *Sport Educ. Soc.* **2002**, *7*, 151–166. [CrossRef]
78. Russell, I.M. *A National Framework for Youth Action and Engagement*; HMSO: London, UK, 2005.

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
Tel. +41 61 683 77 34
Fax +41 61 302 89 18
www.mdpi.com

International Journal of Environmental Research and Public Health Editorial Office
E-mail: ijerph@mdpi.com
www.mdpi.com/journal/ijerph

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland

Tel: +41 61 683 77 34
Fax: +41 61 302 89 18

www.mdpi.com

ISBN 978-3-03943-714-6