Journal of
Clinical Medicine

Clinical Medicine for Healthcare and Sustainability

Edited by
Teen-Hang Meen, Yusuke Matsumoto and Kuan-Han Lee

Printed Edition of the Special Issue Published in
Journal of Clinical Medicine

www.mdpi.com/journal/jcm

MDPI

Clinical Medicine for Healthcare and Sustainability

Clinical Medicine for Healthcare and Sustainability

Editors

Teen-Hang Meen
Yusuke Matsumoto
Kuan-Han Lee

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

Editors
Teen-Hang Meen
Department of Electronic Engineering
National Formosa University
Taiwan

Kuan-Han Lee
Chia Nan University of Pharmacy & Science
Taiwan

Yusuke Matsumoto
Tokyo University of Pharmacy
and Life Sciences
Japan

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 *Journal of Clinical Medicine* (ISSN 2077-0383) (available at: https://www.mdpi.com/journal/jcm/special_issues/IEEE_ECBIOS_2019).

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**, *Article Number*, Page Range.

ISBN 978-3-03936-862-4 (Hbk)
ISBN 978-3-03936-863-1 (PDF)

Contents

About the Editors

Teen-Hang Meen was born in Tainan, Taiwan, on 1 August 1967. He received his B.S. from the Department of Electrical Engineering, National Cheng Kung University (NCKU), Tainan, Taiwan, in 1989, and M.S. and Ph.D. degrees from the Institute of Electrical Engineering, National Sun Yat-Sen University (NSYSU), Kaohsiung, Taiwan, in 1991 and 1994, respectively. He was the chairman of the Department of Electronic Engineering from 2005 to 2011 of the National Formosa University, Yunlin, Taiwan. He reeived the excellent research award from National Formosa University in 2008 and 2014. Currently, he is a distinguished professor with the Department of Electronic Engineering, National Formosa University, Yunlin, Taiwan. He is also the president of the International Institute of Knowledge Innovation and Invention (IIKII) and the chair of IEEE Tainan Section Sensors Council. He has published more than 100 SCI and SSCI papers in recent years.

Yusuke Matsumoto was born in Tokyo, Japan on 14 January 1958. He received his B.S. degree from Tokyo University of Pharmacy and Life Sciences in 1982, and Ph.D. from Showa Pharmaceutical University in 1999. He was the director of the Hachioji Pharmaceutical Association from 1995 to 2018, and he was a professor of Tokyo University of Pharmacy and Life Sciences Department of Pharmacy from 2006 to 2015. Currently, he is the Executive Director of Tokyo University of Pharmacy and Life Sciences..

Kuan-Han Lee, Department of Pharmacy, Chia Nan University of Pharmacy & Science, Tainan 717, Taiwan, E-mail: kuanhanlee@mail.cnu.edu.tw
Education
1. Ph.D. in Department of Pharmacy, Kaohsiung Medical University (1991–1994)
2. M.S. in Department of Pharmacy, Kaohsiung Medical University (1985–1987)
3. B.S. in Department of Pharmacy, Kaohsiung Medical University (1981–1985)
Professional Experience
1. 1989–1990 Lecturer Department of Pharmacy, Tajen University
2. 1990–1995 Lecturer Department of Medicinal and Applied Chemistry, Kaohsiung Medical University
3. 1995–2004 Associate Professor Department of Pharmacy, Tajen University
4. 2005–2007 Professor Department of Pharmacy, Tajen University
5. 2008–2009 Visiting Professor Department of Pharmacy, University of Southern California
6. 2008–Current Professor Department of Pharmacy, Chia Nan University of Pharmacy and Science.

Journal of
Clinical Medicine

MDPI

Editorial

Special Issue on Clinical Medicine for Healthcare and Sustainability

Teen-Hang Meen [1,*], Yusuke Matsumoto [2] and Kuan-Han Lee [3]

[1] Department of Electronic Engineering, National Formosa University, Yunlin 632, Taiwan
[2] Department of Pharmacy, Tokyo University of Pharmacy and Life Sciences, 1432-1 Horinouchi, Hachioji-city, Tokyo 192-0392, Japan; y-matsu@toyaku.ac.jp
[3] Department of Pharmacy, Chia Nan University of Pharmacy & Science, Tainan City 717, Taiwan; kuanhanlee@mail.cnu.edu.tw
* Correspondence: thmeen@nfu.edu.tw

Received: 7 July 2020; Accepted: 9 July 2020; Published: 13 July 2020

Abstract: Recently, due to the advancement of network technology, big data and artificial intelligence, the healthcare industry has undergone many sector-wide changes. Medical care has not only changed from passive and hospital-centric to preventative and personalized, but also from disease-centric to health-centric. Healthcare systems and basic medical research are becoming more intelligent and being implemented in biomedical engineering. This Special Issue on "Clinical Medicine for Healthcare and Sustainability" selected 30 excellent papers from 160 papers presented in IEEE ECBIOS 2019 on the topic of clinical medicine for healthcare and sustainability. Our purpose is to encourage scientists to propose their experiments and theoretical researches to facilitate the scientific prediction and influential assessment of global change and development.

Keywords: healthcare and sustainability; therapy of internal medicine diseases; cardiometabolic diseases

1. Introduction

Due to the development of technology and the advancements in medicine and healthcare, the average life expectancy of human beings has been on the rise for a long time. However, both the fertility rate and the mortality rate have fallen, resulting in the overall population structure rapidly aging, and it has been officially entered an advanced age society. Moreover, as the family's care function gradually fades, the pressure of personal and family care is increasing, which in turn leads to social and economic problems. Therefore, establishing a perfect long-term system of healthcare and sustainability has become one of the key factors to a complete social security system.

Therefore, the 2019 IEEE Eurasian Biomedical Engineering, Healthcare and Sustainability Conference (IEEE ECBIOS 2019) was held in Okinawa, Japan from 31 May to 3 June 2019, providing researchers in the field of biomedical engineering with a unified communication platform for healthcare and sustainability. Recently, due to the developments of computing, network technology, big data, and artificial intelligence, the healthcare industry has undergone a cross-industry transformation. Medical care has not only changed from response-centric and hospital-centric to preventative and personalized, but also from disease-centric to health-centric. Healthcare systems and basic medical research are becoming more intelligent and being implemented in biomedical engineering. This special issue of "Health Care and Sustainable Clinical Medicine" selected 30 excellent papers from 160 papers published in IEEE ECBIOS 2019, with the theme of healthcare and sustainable clinical medicine. It connects multiple disciplines, including clinical laboratory diagnosis and the treatment of medical diseases, traumatology and precision surgical techniques, clinical cancer research, neurology and psychiatry, dermatology, medical imaging, nuclear medicine, genomics, proteomics and bioinformatics, as well as medicine and women's health. Our aim is to encourage scientists to publish their experiments

and theoretical studies to promote scientific predictions and impact assessments of global change and development.

2. The Topics of Clinical Medicine for Healthcare and Sustainability

This Special Issue on "Clinical Medicine for Healthcare and Sustainability" selected 30 excellent papers from 160 papers presented in IEEE ECBIOS 2019 on the topic of clinical medicine for healthcare and sustainability. The topics of published papers are listed in Table 1.

Table 1. The topics and list of papers for the Special Issue on "Clinical Medicine for Healthcare and Sustainability".

Topics	Papers of Special Issue
Clinical Laboratory Diagnosis and Therapy of Internal Medicine Diseases	Kuwabara et al. [1], Daniel et al. [2], Lee et al. [3], Ye et al. [4], Jang [5], Lee et al. [6], Jiang et al. [7], Kapur et al. [8], Park et al. [9], Encarnación et al. [10], Wang et al. [11], Lan et al. [12], Chun et al. [13], Chen et al. [14], Jurik et al. [15], Lai et al. [16], Lin et al. [17], Caneiras et al. [18]
Traumatology and Precise Surgical Techniques	Chiu et al. [19]
Genomics, Proteomics, and Bioinformatics in Clinical Cancer Research	Kong et al. [20], Chiu et al. [21]
Neurological and Psychiatric Disorders	Kume et al. [22], Huh et al. [23], Ricardo et al. [24]
Advanced Research in Dermatology	Damiani et al. [25]
Medical Imaging and Nuclear Medicine	Lee et al. [26], Shiao et al. [27], Jo et al. [28]
Rehabilitation Medicine	Yang et al. [29]
Women's Health	Lee et al. [30]

3. Conclusions

When the domestic government, the private sector, and people in various professional fields talk about related long-term care issues, they all focus on creating a warm and home-like care institution. However, we actively emphasize the importance of community-based long-term care. While implementing the goal of "aging in place", the development of domestic non-institutional care is still in its infancy, and the satisfaction of some long-term care needs must still be completed through institutional care, and the extension or outreach of community-based care, as well as a respite service platform for the development of community-based long-term care, still rely on institutional care to help facilitate it. The development of long-term care in Taiwan is much shorter than that of Japan, Europe, the United States, and Canada. Despite years of hard work and rapid development, the long-term care resources needed to establish a complete system in terms of universalization, fairness, accessibility, and selectivity are not available. It is hoped that in the future, based on the soundness of institutional care, the outreach will move towards the goals of "community care" and "aging in place". We hope the researches of this special issue can improve the developments of clinical medicine for healthcare and sustainability.

Author Contributions: Writing and reviewing all papers, T.-H.M.; English editing, Y.M.; Checking and correcting manuscript, K.-H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The guest editors would like to thank the authors for their contributions to this special issue and all the reviewers for their constructive reviews. We are also grateful to Edwin Zhang, the Editor of the Journal of Clinical Medicine for his time and efforts on the publication of this special issue for the JCM.

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

References

1. Kuwabara, M.; Hisatome, I.; Niwa, K.; Bjornstad, P.; Carlos, A.R.J.; Ana, A.H.; Kanbay, M.; Richard, J.J.; Lanaspa, M.A. The optimal range of serum uric acid for cardiometabolic diseases: A 5-year japanese cohort study. *J. Clin. Med.* **2020**, *9*, 942. [CrossRef] [PubMed]

2. Daniel, P.F.; Concepción, B.R.L.; Concepción, P.C.C.; Cesar, H.M.; María, P.G.C.; Rafael, M.J. Prospective evaluation of intensity of symptoms, therapeutic procedures and treatment in palliative care patients in nursing homes. *J. Clin. Med.* **2020**, *9*, 750.
3. Lee, S.H.; Lim, C.M.; Koh, Y.; Hong, S.B.; Huh, J.W. Effect of an electronic medical record-based screening system on a rapid response system: 8-years' experience of a single center cohort. *J. Clin. Med.* **2020**, *9*, 383. [CrossRef] [PubMed]
4. Ye, S.; Zhang, H.; Shi, F.; Guo, J.; Wang, S.; Zhang, B. Ensemble learning to Improve the prediction of fetal macrosomia and large-for-gestational Age. *J. Clin. Med.* **2020**, *9*, 380. [CrossRef]
5. Jang, A. Postprandial hypotension as a risk factor for the development of new cardiovascular disease: A prospective cohort study with 36 month follow-up in community-dwelling elderly people. *J. Clin. Med.* **2020**, *9*, 345. [CrossRef]
6. Lee, M.; Park, S.; Lee, K.S. Relationship between morbidity and health behavior in chronic diseases. *J. Clin. Med.* **2020**, *9*, 121. [CrossRef]
7. Jiang, F.C.; Shih, C.M.; Wang, Y.M.; Yang, C.T.; Chiang, Y.J.; Lee, C.H. Decision support for the optimization of provider staffing for hospital emergency departments with a queue-based approach. *J. Clin. Med.* **2019**, *8*, 2154. [CrossRef]
8. Kapur, S.; Gehani, M.; Kammili, N.; Bhardwaj, P.; Nag, V.; Devara, S.M.; Sharad, S. Clinical validation of innovative optical-sensor-based, low-cost, rapid diagnostic test to reduce antimicrobial resistance. *J. Clin. Med.* **2019**, *8*, 2098. [CrossRef]
9. Park, K.; Nemoto, K.; Yamakawa, Y.; Yamashita, F.; Yoshida, K.; Tamura, M.; Kawaguchi, A.; Arai, T.; Sasaki, M. cerebral white matter hyperintensity as a healthcare quotient. *J. Clin. Med.* **2019**, *8*, 1823. [CrossRef]
10. Encarnación, B.R.; Valdellós, J.; Ricardo, O.R.; María, R.G.M.; Lorena, A.C.; Gabriel, A.Z.; Inmaculada, B.E. factors associated with health-related quality of life in community-dwelling older adults: A multinomial logistic analysis. *J. Clin. Med.* **2019**, *8*, 1810.
11. Wang, C.Y.; Hsien, H.H.; Hung, K.W.; Lin, H.F.; Chiou, H.Y.; Yeh, S.C.J.; Yeh, Y.J.; Shi, H.Y. Multidiscipline stroke post-acute care transfer system: Propensity-score-based comparison of functional status. *J. Clin. Med.* **2019**, *8*, 1233. [CrossRef] [PubMed]
12. Lan, S.H.; Chang, S.P.; Lai, C.C.; Lu, L.C.; Chao, C.M. The efficacy and safety of eravacycline in the treatment of complicated intra-abdominal infections: A systemic review and meta-analysis of randomized controlled trials. *J. Clin. Med.* **2019**, *8*, 866. [CrossRef] [PubMed]
13. Chun, D.I.; Kim, S.; Kim, J.; Yang, H.J.; Kim, J.H.; Cho, J.H.; Yi, Y.; Kim, W.J.; Won, S.H. Epidemiology and burden of diabetic foot ulcer and peripheral arterial disease in korea. *J. Clin. Med.* **2019**, *8*, 748. [CrossRef] [PubMed]
14. Chen, C.H.; Chen, Y.M.; Yang, Y.; Chang, Y.J.; Lin, L.J.; Yen, H.C. Re-evaluating the protective effect of hemodialysis catheter locking solutions in hemodialysis patients. *J. Clin. Med.* **2019**, *8*, 412. [CrossRef] [PubMed]
15. Jurik, R.; Stastny, P. Role of nutrition and exercise programs in reducing blood pressure: A systematic review. *J. Clin. Med.* **2019**, *8*, 1393. [CrossRef] [PubMed]
16. Lai, C.C.; Cheng, I.L.; Chen, Y.H.; Tang, H.J. The efficacy and safety of doripenem in the treatment of acute bacterial infections—A systemic review and meta-analysis of randomized controlled trials. *J. Clin. Med.* **2019**, *8*, 958. [CrossRef] [PubMed]
17. Lin, H.T.; Li, Y.I.; Hu, W.P.; Huang, C.C.; Du, Y.C. A scoping review of the efficacy of virtual reality and exergaming on patients of musculoskeletal system disorder. *J. Clin. Med.* **2019**, *8*, 791. [CrossRef] [PubMed]
18. Caneiras, C.; Jácome, C.; Sagrario, M.A.; Calvo, J.R.; Fonseca, J.A.; Escarrabill, J.; Winck, J.C. Patient experience in home respiratory therapies: Where we are and where to go. *J. Clin. Med.* **2019**, *8*, 555. [CrossRef]
19. Chiu, C.C.; Lin, W.L.; Shi, H.Y.; Huang, C.C.; Chen, J.J.; Su, S.B.; Lai, C.C.; Chao, C.M.; Tsao, C.J.; Chen, S.H.; et al. Comparison of oncologic outcomes in laparoscopic versus open surgery for non-metastatic colorectal cancer: Personal experience in a single institution. *J. Clin. Med.* **2019**, *8*, 875. [CrossRef]
20. Kong, S.; Park, H.Y.; Kang, D.; Lee, J.K.; Lee, G.; Kwon, O.J.; Shim, Y.M.; Zo, J.I.; Cho, J. Seasonal variation in physical activity among preoperative patients with lung cancer determined using a wearable device. *J. Clin. Med.* **2020**, *9*, 349. [CrossRef]
21. Chiu, C.C.; Lee, K.T.; Wang, J.J.; Sun, D.P.; Lee, H.H.; Huang, C.C.; Shi, H.Y. Preoperative health-related quality of life predicts minimal clinically important difference and survival after surgical resection of hepatocellular carcinoma. *J. Clin. Med.* **2019**, *8*, 576. [CrossRef] [PubMed]

22. Kume, Y.; Takahashi, T.; Itakura, Y.; Lee, S.; Makizako, H.; Ono, T.; Shimada, H.; Ota, H. Characteristics of mild cognitive impairment in northern japanese community-dwellers from the orange registry. *J. Clin. Med.* **2019**, *8*, 1937. [CrossRef] [PubMed]
23. Huh, Y.; Nam, G.E.; Kim, Y.H.; Lee, J.H. Relationships between multimorbidity and suicidal thoughts and plans among korean adults. *J. Clin. Med.* **2019**, *8*, 1094. [CrossRef]
24. Ricardo, J.D.O.; Teresa, M. Abuse of licit and illicit psychoactive substances in the workplace: Medical, toxicological, and forensic aspects. *J. Clin. Med.* **2020**, *9*, 770.
25. Damiani, G.; Pacifico, A.; Russo, F.; Pigatto, P.D.M.; Bragazzi, N.L.; Bonifati, C.; Morrone, A.; Watad, A.; Adawi, M. Use of secukinumab in a cohort of erythrodermic psoriatic patients: A pilot study. *J. Clin. Med.* **2019**, *8*, 770. [CrossRef] [PubMed]
26. Lee, K.S.; Jung, S.K.; Ryu, J.J.; Shin, S.W.; Choi, J. Evaluation of transfer learning with deep convolutional neural networks for screening osteoporosis in dental panoramic radiographs. *J. Clin. Med.* **2020**, *9*, 392. [CrossRef] [PubMed]
27. Shiao, E.C.; Kuo, P.L. Two-dimensional laser-align device for ultrasound-guided injection. *J. Clin. Med.* **2019**, *8*, 1048. [CrossRef]
28. Jo, M.H.; Lim, T.S.; Jeon, M.Y.; Lee, H.W.; Kim, B.K.; Park, J.Y.; Kim, D.Y.; Ahn, S.H.; Han, K.H.; Kim, S.U. Predictors of discordance in the assessment of skeletal muscle mass between computed tomography and bioimpedance analysis. *J. Clin. Med.* **2019**, *8*, 322. [CrossRef] [PubMed]
29. Yang, C.; Mustafa, R.; Scott, M.; Johan, O.; Ingmar, N. Predicting long-term health-related quality of life after bariatric surgery using a conventional neural network: A study based on the scandinavian obesity surgery registry. *J. Clin. Med.* **2019**, *8*, 2149.
30. Lee, J.H.; Go, T.H.; Lee, S.H.; Kim, J.; Huh, J.H.; Kim, J.Y.; Kang, D.R.; Jeong, S.; Koh, S.B.; Choi, J.R. Association between serum urate and risk of hypertension in menopausal women with xdh gene. *J. Clin. Med.* **2019**, *8*, 738. [CrossRef]

Journal of
Clinical Medicine

MDPI

Article

The Optimal Range of Serum Uric Acid for Cardiometabolic Diseases: A 5-Year Japanese Cohort Study †

Masanari Kuwabara [1,2,*], Ichiro Hisatome [3], Koichiro Niwa [2], Petter Bjornstad [4,5], Carlos A. Roncal-Jimenez [4], Ana Andres-Hernando [4], Mehmet Kanbay [6], Richard J. Johnson [4] and Miguel A. Lanaspa [4]

1 Intensive Care Unit and Department of Cardiology, Toranomon Hospital, Tokyo 105-8470, Japan
2 Cardiovascular Center, St. Luke's International Hospital, Tokyo 104-8560, Japan; kniwa@aol.com
3 Division of Regenerative Medicine and Therapeutics, Tottori University Graduate School of Medical Sciences, Tottori 683-8503, Japan; hisatome@med.tottori-u.ac.jp
4 Division of Renal Diseases and Hypertension, School of Medicine, University of Colorado Denver, Aurora, CO 80045, USA; Petter.bjornstad@childrenscolorado.org (P.B.), Carlos.Roncal@ucdenver.edu (C.A.R.-J.); Ana.AndresHernando@ucdenver.edu (A.A.-H.); Richard.Johnson@ucdenver.edu (R.J.J.); Miguel.LanaspaGarcia@ucdenver.edu (M.A.L.)
5 Children's Hospital Colorado and Barbara Davis Center for Childhood Diabetes, Aurora, CO 80045, USA
6 Division of Nephrology, Department of Medicine, Koc University School of Medicine, Istanbul 34450, Turkey; mkanbay@ku.edu.tr
* Correspondence: kuwamasa728@gmail.com; Tel.: +81-3-3588-1111; Fax: +81-3582-7068
† Optimal uric acid for cardiometabolic diseases.

Received: 14 February 2020; Accepted: 27 March 2020; Published: 30 March 2020

Abstract: The optimal range of serum uric acid (urate) associated with the lowest risk for developing cardiometabolic diseases is unknown in a generally healthy population. This 5-year cohort study is designed to identify the optimal range of serum urate. The data were collected from 13,070 Japanese between ages 30 and 85 at the baseline (2004) from the Center for Preventive Medicine, St. Luke's International Hospital, Tokyo. We evaluated the number of subjects (and prevalence) of those free of the following conditions: hypertension, diabetes, dyslipidemia, and chronic kidney disease (CKD) over 5 years for each 1 mg/dL of serum urate stratified by sex. Furthermore, the odds ratios (ORs) for remaining free of these conditions were calculated with multiple adjustments. Except for truly hypouricemic subjects, having lower serum urate was an independent factor for predicting the absence of hypertension, dyslipidemia, and CKD, but not diabetes. The OR of each 1 mg/dL serum urate decrease as a protective factor for hypertension, dyslipidemia, and CKD was 1.153 (95% confidence interval, 1.068–1.245), 1.164 (1.077–1.258), and 1.226 (1.152–1.306) in men; 1.306 (1.169–1.459), 1.121 (1.022–1.230), and 1.424 (1.311–1.547) in women, respectively. Moreover, comparing serum urate of 3–5 mg/dL in men and 2–4 mg/dL in women, hypouricemia could be a higher risk for developing hypertension (OR: 4.532; 0.943–21.78) and CKD (OR: 4.052; 1.181–13.90) in women, but not in men. The optimal serum urate range associated with the lowest development of cardiometabolic diseases was less than 5 mg/dL for men and 2–4 mg/dL for women, respectively.

Keywords: uric acid; risk factor; epidemiology; cardiometabolic diseases; hypertension

1. Introduction

Both epidemiologically and in animal models, hyperuricemia is strongly associated with the development and progression of cardiovascular disease [1,2]. In this regard, a recent meta-analysis report estimated a 12% increase in coronary heart disease per 1 mg/dL elevation in serum uric acid

(urate) levels. However, it is important to note that other studies could not support serum urate as a truly independent risk factor for cardiovascular diseases [3–5]. Furthermore, some studies have shown a J-shaped relationship between serum urate and cardiovascular disease supporting the idea that both low and high urate are associated with greater risk of developing cardiovascular events. As an example, according to the PIUMA study, serum urate levels lower than 4.5 mg/dL in men and 3.2 mg/dL in women with essential hypertension but also higher than 5.2 mg/dL in men and 3.9 mg/dL in women are associated with increased rates of cardiovascular disease and cardiovascular disease-related deaths. The J-curve phenomenon relating serum urate to cardiovascular events was also reported in the Syst-Eur trial [6]. The variability in the findings from the cross-sectional studies and the evidence supporting a J-shaped cause-effect between urate and cardiovascular disease suggests the need to identify the proper level range for which urate could be relatively safe, or in contrast, could exert a substantial deleterious role in the pathogenesis of cardiovascular disease. In this regard, little information is known on the optimal serum urate range associated with the lowest risk of disease in a generally healthy population.

Therefore, we aimed to evaluate whether a J-shaped curve exists for serum urate with cardiometabolic diseases (hypertension, diabetes, dyslipidemia, and chronic kidney disease (CKD)) and to identify the optimal range of serum urate that is associated with the lowest risk for developing these conditions by a longitudinal design.

2. Materials and Methods

2.1. Study Design and Study Subjects

This study included a single-center, large-scale, cross-sectional study and a 5-year longitudinal cohort study in Japan. We reviewed and used the database of health records from the Center for Preventive Medicine, St. Luke's International Hospital, Tokyo, Japan between 2004 and 2009. The study subjects came to the center to have annual regular health check-up by themselves, and also provided a general history for comorbidities. Every subject and/or their companies paid for the examinations and each subject had identical physical and laboratory examinations including blood pressure. In general, the patients with any symptoms go to their hospital or clinic with Japanese government insurance, and therefore this study population was defined as a 'generally healthy population'. Some of the findings from this database have already been published [7–15]. There were 30,227 subjects (15,263 men) who underwent an annual health check-up at the center in 2004. In this study, we enrolled 13,201 subjects who underwent health checks both in 2004 and in 2009. The background demographics between all the subjects in 2004 and this cohort study subjects were similar as shown in our previous manuscript [8]. The prevalence of hyperuricemia between the two groups showed no significant differences, including in men and in women, respectively [8]. For this study, we included subjects between the ages 30 and 85 in 2004 whose data were available at both 2004 and 2009. Subjects younger than 30 years of age were excluded due to their very modest risk for hypertension and cardiovascular diseases, while subjects aged 85 years old and above were excluded due to the substantial risk of death during a five-year follow-up. Out of the 13,201 subjects, only 121 subjects were less than 30 years old and 10 subjects were 85 years old and above in 2004, and 13,070 subjects were enrolled (mean age: 51.1 ± 11.3 years). Of these, 6367 were men (52.4 ± 11.5 years) and 6733 women (49.9 ± 11.1 years). For our first analysis (a cross-sectional study), we checked the prevalence of cardiometabolic diseases, like hypertension, diabetes, dyslipidemia, and CKD in each 1 mg/dL of serum urate range in each sex at baseline (2004). Then, we excluded the subjects with each cardiometabolic condition at baseline, and we checked whether they remained free of these conditions over the following five years, as it related to each serum urate quartile separated by sex (a 5-year cohort study). We also calculated the odds ratios (ORs) of each 1 mg/dL increase of serum urate for each cardiometabolic diseases after multiple adjustments for well-known factors associated with cardiovascular diseases including age, body mass index (BMI), smoking and drinking habits, serum urate, and the presence of other cardiometabolic diseases as

detailed in Figure 1. When we assessed the protective factors for the development of hypertension, diabetes, dyslipidemia, and CKD, we excluded 2599 subjects with hypertensions, 575 subjects with diabetes, 5118 subjects with dyslipidemia, and 492 subjects with CKD at the baseline in each analysis, respectively. Moreover, we compared the cumulative incidence of any of these cardiometabolic diseases between hypouricemic subjects (serum urate of less than 3 mg/dL in men and less than 2 mg/dL in women) and normouricemic subjects (serum urate of 3–5 mg/dL in men and 2–4 mg/dL in women) to detect whether hypouricemia is a risk for developing cardiometabolic diseases compared with normouricemia. We also calculated the ORs of hypouricemia for each cardiometabolic diseases compared with normouricemia after multiple adjustments.

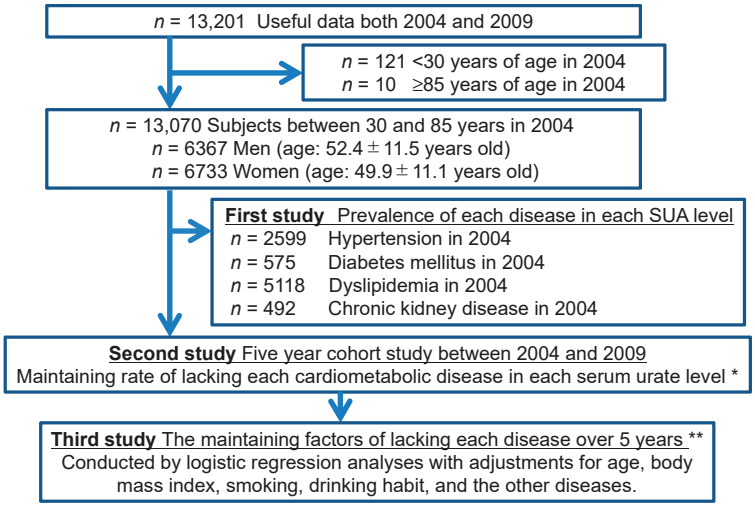


Figure 1. Flow diagram of study enrollment. All the analyses were stratified by sex. * Each cardiometabolic disease means hypertension, diabetes, dyslipidemia and chronic kidney disease. ** The number of subjects depends on the excluded subjects having the corresponding disorders at baseline.

2.2. Definition of Hypertension, Diabetes, Dyslipidemia, CKD, and Hypouricemia

Hypertension is defined as a condition when subjects are on current antihypertensive medication and/or systolic blood pressure of more than or equal to 140 mmHg and/or diastolic blood pressure of more than or equal to 90 mmHg [16,17]. Blood pressure readings were obtained using an automatic brachial sphygmomanometer (OMRON Corporation, Kyoto, Japan), which was upper arm blood pressure measuring and had passed validation. Two blood pressure examinations were taken after the participants were seated and rested quietly for more than five minutes with their feet on the ground and their back supported. The mean systolic and diastolic blood pressure of each of the subjects were calculated from the recorded measurements. Diabetes is defined as current diabetes mellitus on medication use and/or HbA1c (National Glycohemoglobin Standardization Program) more than or equal to 6.5%, according to International Expert Committee. Dyslipidemia is defined as current medication use for dyslipidemia and/or low-density lipoprotein cholesterol more than or equal to 140 mg/dL, high-density lipoprotein cholesterol less than 40 mg/dL, and/or triglyceride more than or equal to 150 mg/dL, according to Japan Atherosclerosis Society guidelines [18]. CKD is defined as estimated glomerular filtration rate (eGFR) is less than 60 mL/min/1.73m^2. We calculated eGFR using the Japanese GFR equation: eGFR (mL/min/1.73m^2) = 194 × serum creatinine $^{-1.094}$ × age $^{-0.287}$ (×0.739 if woman) [19]. Hypouricemia is defined as serum urate level lower than 3.0 mg/dL in men and 2.0 mg/dL in women in this study [20]

2.3. Statistical Analysis

All the statistical analyses were performed using the SPSS Statistics software (IBM SPSS Statistics version 22 for Windows; IBM, New York, NY, USA). The statistically significant level was set at probability $p < 0.05$ (two-tailed). Data are expressed as mean ± standard deviation or as percent frequency unless otherwise specified. Comparisons between two groups were performed with student *t*-tests for normally distributed variables, and χ^2 analyses for categorical data. The maintaining factors for lacking hypertension, diabetes, dyslipidemia, and CKD in the period of over five years were evaluated both by crude models and by multivariable logistic regression models with adjustments of the age, BMI, smoking and drinking habits, serum urate, and the other cardiometabolic diseases. We also calculated odds ratios (ORs) in each group. When we analyzed logistic regression analyses in the longitudinal study, we excluded hypouricemic subjects because there was not a linear association between serum urate levels and the maintaining rate of lacking prevalence of these cardiometabolic diseases only in hypouricemic subjects. Moreover, we compared cumulative incidence of each cardiometabolic disease between hypouricemia and normouricemia to clarify whether *J.* curve phenomenon exists or not. In this analysis, we used propensity score matching to combine the other factors (age, BMI, smoking and drinking habits, and cardiometabolic feathers; hypertension, diabetes, dyslipidemia, and CKD) into one parameter because the number of hypouricemic subjects were small (45 hypouricemic subjects).

2.4. Ethical Considerations

We adhered to the principles of the Declaration of Helsinki. All data were collected and compiled in a protected computer database. Individual data were anonymous without identifiable personal information. Informed consent was obtained from all subjects by a comprehensive agreement method provided by St. Luke's International Hospital. St. Luke's International Hospital Ethics Committee approved the protocol for this study (approval number: 16-R025).

3. Results

3.1. Demographics of this Study's Subjects

Table 1 shows the demographics of this study for men and women. In general, women were significantly older, and had lower BMI, lower blood pressure, less smoking and drinking habits, lower prevalence of hypertension, diabetes, dyslipidemia, and CKD, and lower serum urate compared to men.

Table 1. Demographics of study subjects at baseline (2004).

	Women	**Men**	***p***
Number of Subjects	**6733**	**6337**	
Age	49.9 ± 11.1	42.4 ± 11.5	<0.001
Body mass index (kg/m^2)	21.3 ± 3.0	23.8 ± 2.9	<0.001
Systolic blood pressure (mmHg)	114.5 ± 17.5	124.2 ± 17.2	<0.001
Diastolic blood pressure (mmHg)	70.8 ± 10.9	77.9 ± 10.9	<0.001
Pulse rate (bpm)	75.2 ± 10.8	71.6 ± 10.5	<0.001
Smoking	16.3%	63.0%	<0.001
Drinking habits	26.0%	61.5%	<0.001
Hypertension	13.4%	26.8%	<0.001
Diabetes mellitus	2.1%	6.8%	<0.001
Dyslipidemia	29.6%	49.3%	<0.001
Hypouricemia	0.22%	0.47%	0.016
Chronic kidney disease	2.3%	5.3%	<0.001
eGFR (mL/min/1.73m^2)	88.2 ± 15.7	82.6 ± 15.5	<0.001
Serum uric acid (mg/dL)	4.49 ± 0.95	6.24 ± 1.23	<0.001

bpm, beats per minute; *p*, probability. Data are presented as mean ± standard deviation.

3.2. Prevalence of Cardiometabolic Disease in Each Serum Urate Level (A Cross-Sectional Study)

Figure 2 shows the prevalence of hypertension, diabetes, dyslipidemia, and CKD for each 1mg/dL of serum urate range at baseline (2004). As shown in the figure, serum urate lower than 4 mg/dL were associated with the lowest prevalence of hypertension, dyslipidemia, and CKD in women while the range of serum urate between 2 and 4 mg/dL corresponded with the lowest prevalence of diabetes. In men, the range of serum urate associated with the lowest prevalence of these conditions was more variable. Interestingly, the prevalence of diabetes in men decreased with increasing serum urate, which may be due to the effect of glycosuria to cause uricosuria and decrease serum urate levels. As a result, serum urate levels in men ranging from 2 to 6 mg/dL corresponded with the lowest prevalence of dyslipidemia and CKD while levels ranging from 3 to 6 mg/dL were associated with the lowest prevalence of hypertension, respectively. However, it is important to note that this cross-sectional analysis at baseline did not account for medication for each disease, raising the possibility of a potential medication bias. Therefore, we conducted a 5-year cohort study to evaluate the odds for remaining free of these disease conditions over time.

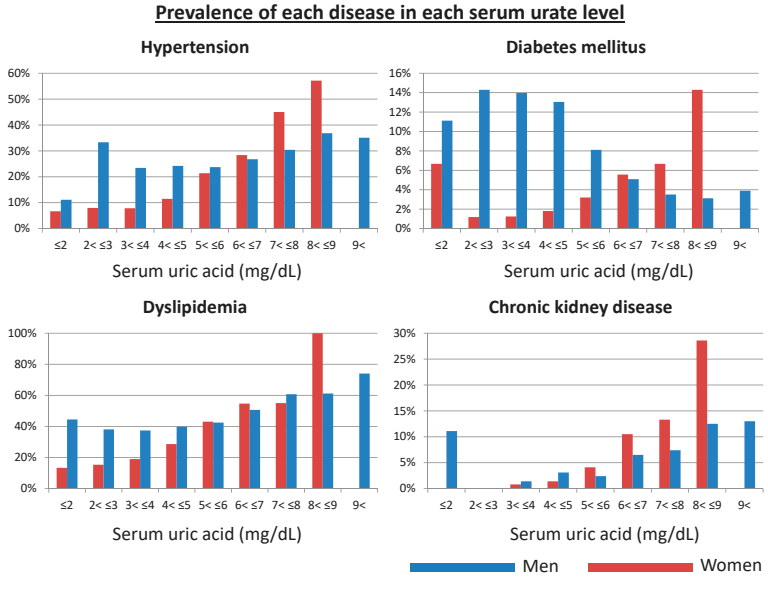


Figure 2. Prevalence of hypertension, diabetes, dyslipidemia, and chronic kidney disease in each serum urate at baseline (2004). Blue bars showed men and red bars showed women.

3.3. Rate of Being Free of Various Cardiometabolic Conditions According to Serum Urate Levels over Five Years (A Longitudinal Study)

The number of subjects with the new development of hypertension, diabetes, dyslipidemia, and CKD over 5 years was 1108/10,471 (10.6%), 318/12,495 (2.5%), 1454/7952 (18.3%), and 1961/12,578 (15.6%), respectively. Figure 3 shows the relative risk for being free of cardiometabolic disease (hypertension, diabetes, dyslipidemia, and CKD) for each serum urate group over a five-year period. There is a linear association between serum urate levels and the rate of being free of cardiometabolic disease except for subjects with hyporuricemia. and we excluded hypouricemic subjects. The multivariable analyses showed that serum urate levels were protective for developing hypertension, diabetes, dyslipidemia, and CKD irrespective of sex, except for hypouricemic subjects. Even when accounting for a J-curve phenomenon with hypouricemic subjects, hyperuricemic subjects had lower maintaining rates with respect to lacking hypertension, diabetes, dyslipidemia, and CKD than hypouricemic subjects.

We conducted additional analyses using four categories of serum urate levels; 2 or less (hypouricemia), from 2 to 4 mg/dL, from 4 to 6 mg/dL, and more than 6 mg/dL (hyperuricemia). We compared the relative risk for being free of cardiometabolic disease (hypertension, diabetes, dyslipidemia, and CKD) among these four serum urate categories over a five-year period. Figure 4 shows that the group with serum urate from 2 to 4 mg/dL exhibited maintaining rates of lacking hypertension or CKD compared to the other urate categories. The group with serum urate of 2 mg/dL or less had the highest maintaining rate with respect to lacking diabetes or dyslipidemia.

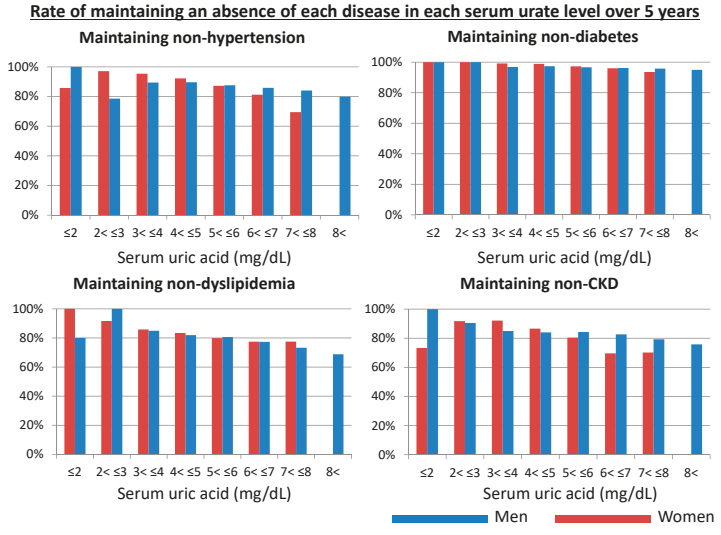


Figure 3. Maintaining rate of lacking hypertension, diabetes, dyslipidemia, and chronic kidney disease in each serum urate over five years. CKD, chronic kidney disease. Blue bars showed men and red bars showed women.

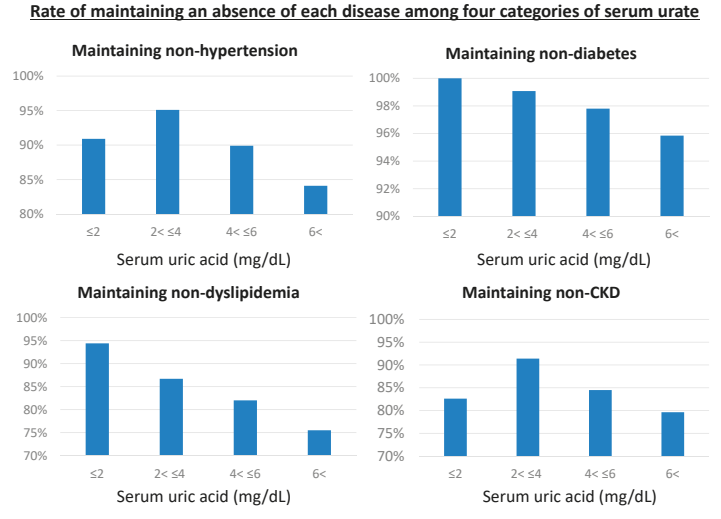


Figure 4. Maintaining rate of lacking hypertension, diabetes, dyslipidemia, and chronic kidney disease among four serum urate categories over five years. CKD, chronic kidney disease.

3.4. Optimal Serum Urate Range Associated with the Lowest Risk of Cardiometabolic Diseases

To determine the optimal range of serum urate to prevent the development of cardiometabolic disease, we conducted a multivariable logistic regression analysis and calculated ORs for maintaining conditions without hypertension, diabetes, dyslipidemia, and CKD after excluding 45 hypouricemic subjects (30 men and 15 women) since we intended to exclude the effects of *J.* curve phenomenon.

To evaluate factors that predict continued normotension, we analyzed 10,471 subjects after excluding 2599 subjects with hypertensions at baseline. After multivariable adjustments for age, BMI, smoking and drinking habits, diabetes, dyslipidemia, and CKD, lower serum urate was an independent factor that protects against the development of hypertension both in men (OR per 1 mg/dL decrease: 1.153; 95% CI, 1.068–1.245) and women (OR: 1.306; 95% CI, 1.169–1.459) (Table 2, Hypertension).

When we assessed the factors that protected against the development of diabetes, we analyzed 12,495 subjects after excluding 575 subjects with diabetes at the baseline. After multiple adjustments age, BMI, smoking and drinking habits, hypertension, dyslipidemia, and CKD, lower serum urate showed a tendency as a protective factor for the development of diabetes in women (OR per 1 mg/dL decrease: 1.206; 95% CI, 0.969–1.500), but it did not reach the significant ($p = 0.093$). In contrast, lower serum urate was not an independent protective factor for the development of diabetes in men ($p = 0.24$) (Table 2, Diabetes).

When we assessed the protective factor for the development of dyslipidemia, we analyzed 7952 subjects after excluding 5118 subjects with dyslipidemia at baseline. After multiple adjustments age, BMI, smoking and drinking habits, hypertension, diabetes, and CKD, lower serum urate was an independent protective factor for the development of dyslipidemia both in men (OR per 1 mg/dL decrease: 1.164; 95% CI, 1.077–1.258) and women (OR per 1 mg/dL decrease: 1.121; 95% CI, 1.022–1.230) (Table 2, Dyslipidemia).

When we assessed the factors that protected against the development of CKD, we analyzed 12,578 subjects after excluding 492 subjects with CKD at the baseline. After multiple adjustments age, BMI, smoking and drinking habits, hypertension, diabetes, and dyslipidemia, lower serum urate was an independent protective factor for the development of CKD both in men (OR per 1 mg/dL decrease: 1.226; 95% CI, 1.152–1.306) and women (OR per 1 mg/dL decrease: 1.424; 95% CI, 1.311–1.547) (Table 2, Chronic kidney disease).

We also compared the ORs for hypertension, diabetes, dyslipidemia, and CKD among four categories of serum urate levels. We referenced the group with serum urate from 2 to 4 mg/dL as shown in Table 3. Belonging to the group with serum urate from 2 to 4 mg/dL conferred protection from developing hypertension, dyslipidemia, and CKD when compared with the group with serum urate more than 4 mg/dL, but not diabetes (Table 3).

3.5. Hypouricemia as a Risk of Cardiometabolic Diseases Compared with Normouricemia

We compared the cumulative incidence of cardiometabolic diseases between hypouricemic subjects (30 men and 15 women) and normouricemic subjects (3–5 mg/dL for 958 men and 2–4 mg/dL for 2192 women). The number of hypouricemic subjects were small, and we could not analyze ORs of diabetes both in men and women and dyslipidemia in women. After multiple adjustments age, BMI, smoking and drinking habits, diabetes, dyslipidemia, and CKD, hypouricemia tends to be higher risk for the development of CKD in women (OR: 4.532; 95% CI, 0.943–21.78), but not reach significance ($p = 0.059$) (Table 4, Hypertension). After multiple adjustments that included age, BMI, smoking and drinking habits, hypertension, diabetes, and CKD, hypouricemia continued to show a higher risk for the development of CKD in women (OR: 4.052; 95% CI, 1.181–13.90), but not in men. (Table 4, Chronic kidney disease). The cumulative incidence of hypertension, dyslipidemia, and CKD in men was not significantly different between hypouricemic and normouricemic groups.

Table 2. Lower serum urate as a protective factor for the development of hypertension, diabetes, dyslipidemia and chronic kidney disease over 5 years.

Maintaining without Hypertension		**Crude**			**Adjusted ***		
Women		OR	95% CI	*p*	OR	95% CI	*p*
Serum uric acid	per 1 mg/dL decreased	1.755	1.587–1.941	<0.001	1.306	1.169–1.459	<0.001
Men							
Serum uric acid	per 1 mg/dL decreased	1.180	1.099–1.266	<0.001	1.153	1.068–1.245	<0.001
Maintaining without diabetes mellitus		**Crude**			**Adjusted †**		
Women		OR	95% CI	*p*	OR	95% CI	*p*
Serum uric acid	per 1 mg/dL decreased	1.822	1.507–2.202	<0.001	1.206	0.969–1.500	0.093
Men							
Serum uric acid	per 1 mg/dL decreased	1.160	1.037–1.298	0.010	1.074	0.953–1.210	0.24
Maintaining without dyslipidemia		**Crude**			**Adjusted ‡**		
Women		OR	95% CI	*p*	OR	95% CI	*p*
Serum uric acid	per 1 mg/dL decreased	1.311	1.202–1.429	<0.001	1.121	1.022–1.230	0.015
Men							
Serum uric acid	per 1 mg/dL decreased	1.205	1.120–1.296	<0.001	1.164	1.077–1.258	<0.001
Maintaining without chronic kidney disease		**Crude**			**Adjusted ¶**		
Women		OR	95% CI	*p*	OR	95% CI	*p*
Serum uric acid	per 1 mg/dL decreased	1.655	1.545–1.795	<0.001	1.424	1.311–1.547	<0.001
Men							
Serum uric acid	per 1 mg/dL decreased	1.144	1.082–1.210	<0.001	1.226	1.152–1.306	<0.001

OR, odds ratio; CI, confidence interval; *p*, probability. * Data adjusted for age, body mass index, smoking and drinking habits, diabetes mellitus, dyslipidemia, chronic kidney disease, and serum uric acid. † Data adjusted for age, body mass index, smoking and drinking habits, hypertension, dyslipidemia, chronic kidney disease, and serum uric acid. ‡ Data adjusted for age, body mass index, smoking and drinking habits, hypertension, diabetes mellitus, chronic kidney disease, and serum uric acid. ¶ Data adjusted for age, body mass index, smoking and drinking habits, hypertension, diabetes mellitus, dyslipidemia, and serum uric acid.

Table 3. Lower serum urate as a protective factor for the development of hypertension, diabetes, dyslipidemia and chronic kidney.

Maintaining without hypertension	**Crude**			**Adjusted ***		
Serum uric acid	OR	95% CI	*P*	OR	95% CI	*p*
2 mg/dL to 4 mg/dL	Reference			Reference		
2 mg/dL and less	1.922	0.444–8.328	0.38	1.705	0.385–7.564	0.48
4 mg/dL to 6 mg/dL	2.170	1.756–2.682	<0.001	1.543	1.237–1.926	<0.001
more than 6 mg/dL	3.630	2.920–4.512	<0.001	2.031	1.570–2.628	<0.001
Maintaining without diabetes mellitus	**Crude**			**Adjusted †**		
Serum uric acid	OR	95% CI	*P*	OR	95% CI	*p*
2 mg/dL to 4 mg/dL	Reference			Reference		
2 mg/dL and less	–	–	–	–	–	–
4 mg/dL to 6 mg/dL	2.404	1.530–3.779	<0.001	1.405	0.881–2.238	0.153
more than 6 mg/dL	4.634	2.957–7.262	<0.001	1.571	0.947–2.606	0.080
Maintaining without dyslipidemia	**Crude**			**Adjusted ‡**		
Serum uric acid	OR	95% CI	*P*	OR	95% CI	*p*
2 mg/dL to 4 mg/dL	Reference			Reference		
2 mg/dL and less	0.384	0.051–2.896	0.35	0.346	0.046–2.622	0.30
4 mg/dL to 6 mg/dL	1.437	1.233–1.674	<0.001	1.259	1.073–1.478	0.005
more than 6 mg/dL	2.116	1.784–2.508	<0.001	1.568	1.267–1.940	<0.001
Maintaining without chronic kidney disease	**Crude**			**Adjusted ¶**		
Serum uric acid	OR	95% CI	*p*	OR	95% CI	*p*
2 mg/dL to 4 mg/dL	Reference			Reference		
2 mg/dL and less	2.236	0.754–6.634	0.15	2.368	0.752–7.459	0.14
4 mg/dL to 6 mg/dL	1.949	1.665–2.281	<0.001	1.579	1.337–1.864	<0.001
more than 6 mg/dL	2.716	2.307–3.199	<0.001	2.345	1.927–2.854	<0.001

OR, odds ratio; CI, confidence interval; *p*, probability. * Data adjusted for age, body mass index, smoking and drinking habits, diabetes mellitus, dyslipidemia, chronic kidney disease, and serum uric acid. † Data adjusted for age, body mass index, smoking and drinking habits, hypertension, dyslipidemia, chronic kidney disease, and serum uric acid. ‡ Data adjusted for age, body mass index, smoking and drinking habits, hypertension, diabetes mellitus, chronic kidney disease, and serum uric acid. ¶ Data adjusted for age, body mass index, smoking and drinking habits, hypertension, diabetes mellitus, dyslipidemia, and serum uric acid.

Table 4. Hypouricemia (serum urate less than 3 mg/dL in men and less than 2 mg/dL in women) as a risk factor for the development of hypertension, diabetes, dyslipidemia and chronic kidney disease over 5 years compared with normouricemia (serum urate of 3–5 mg/dL in men and 2–4 mg/dL in women).

Hypertension		**Crude**			**Adjusted ***		
Women	**Reference**	OR	95% CI	*p*	OR	95% CI	*p*
Hypouricemia (*n* = 14)	Normouricemia (*n* = 2020)	3.659	0.807–16.599	0.093	4.532	0.943–21.78	0.059
Men							
Hypouricemia (*n* = 22)	Normouricemia (*n* = 728)	1.355	0.392–4.684	0.545	1.141	0.319–4.075	0.84
Diabetes		**Crude**			**Adjusted**		
Women	**Reference**	OR	95% CI	*p*	OR	95% CI	*p*
Hypouricemia (*n* = 14)	normouricemia (*n* = 2165)	–			–		
Men							
Hypouricemia (*n* = 26)	normouricemia (*n* = 831)	–			–		
Dyslipidemia		**Crude**			**Adjusted †**		
Women	**Reference**	OR	95% CI	*p*	OR	95% CI	*p*
Hypouricemia (*n* = 13)	normouricemia (*n* = 1789)	–			–		
Men							
Hypouricemia (*n* = 18)	normouricemia (*n* = 582)	0.28	0.037–2.129	0.219	0.238	0.031–1.847	0.17
Chronic kidney disease		**Crude**			**Adjusted ‡**		
Women	**Reference**	OR	95% CI	*p*	OR	95% CI	*p*
Hypouricemia (*n* = 15)	normouricemia (*n* = 1795)	4.212	1.327–13.37	0.015	4.052	1.181–13.90	0.026
Men							
Hypouricemia (*n* = 29)	normouricemia (*n* = 932)	0.396	0.093–1.681	0.209	0.303	0.068–1.351	0.117

OR, odds ratio; CI, confidence interval; *p*, probability; N/A, not available for analysis. * Data adjusted for age, body mass index, smoking and drinking habits, diabetes mellitus, dyslipidemia, chronic kidney disease, and serum uric acid. † Data adjusted for age, body mass index, smoking and drinking habits, hypertension, diabetes mellitus, chronic kidney disease, and serum uric acid. ‡ Data adjusted for age, body mass index, smoking and drinking habits, hypertension, diabetes mellitus, dyslipidemia, and serum uric acid.

4. Discussion

The primary goal of our study was to identify the range of serum urate associated with the lowest risk for developing cardiometabolic diseases in a healthy Japanese population. Except for truly hypouricemic subjects (defined as ≤3 mg/dL in men and ≤2 mg/dL in women), our study indicates that lower serum urate level is an independent protective factor for the development of cardiometabolic disease. We show that in heathy subjects, for each 1 mg/dL decrease of serum urate in men, there was an 18% increment in the protection from developing hypertension, a 16% increment against dyslipidemia, and a 23% increment against CKD. Compared to men, lower serum urate in women conferred greater odds for preventing the appearance of cardiometabolic diseases. Specifically, for each 1 mg/dL decrease of serum urate in women, there was a 31% increment in the protection from developing hypertension, a 12% increment against dyslipidemia, and a 42% increment against CKD.

We also compared the cumulative incidence of cardiometabolic diseases over 5 years between hypouricemic subjects and normouricemic subjects (3–5 mg/dL for men and 2–4 mg/dL for women). The number of hypouricemic subjects was small (30 men and 15 women), and it might be difficult to apply the results to every population because of less power to analyze. However, our results suggest that hypouricemia could be a risk for development of hypertension and CKD in women, but not in men. Accounting for these results, we could see the *J.* curve phenomenon only in women, and the optimal serum urate range associated with the less development of cardiometabolic diseases could be less than 5 mg/dL for men and 2–4 mg/dL for women in a generally healthy population.

Other studies have also showed an inverse correlation between serum urate and the incidence of cardiovascular diseases in subjects with serum urate levels lower than 4.5 mg/dL in men and 3.2 mg/dL in women [6,21,22]. This phenomenon is observed primarily in those subjects with low serum urate levels. Of note, the study subjects in these previous reports often were hypertensive, diabetic or receiving medication against these conditions [6,21,22]. In our study, we can see the similar J-curve phenomenon in hypertension and CKD in women, but the serum urate levels required for this J-shape phenomenon were much lower compared to those reported in previous studies [6,21,22]

Our study also showed that hypouricemic subjects demonstrated greater risk for developing hypertension and CKD than normouricemic subjects in women. We postulate that the higher cumulative incidence of cardiometabolic diseases in hypouricemic women could well relate to the relatively frequent genetic loss of the urate transporter (URAT) in the Japanese population. Potential mechanisms for why this increases the risk for these conditions might relate to the marked uricosuria that may increase the risk for kidney disease, or potentially the possibility that a low serum urate may reduce antioxidant activity in the patients [23,24]. Importantly, there are no studies to determine whether lowering serum urate levels to very low levels with xanthine oxidase inhibitors increases cardiovascular risk compared to untreated controls.

Our study points out the necessity of addressing the risk of hypouricemia in addition to hyperuricemia in the pathogenesis of cardiovascular disease. Our published data demonstrated that hypouricemia is associated with endothelium dysfunction [23]. Consistently, a large-scale cross-sectional study showed that hypouricemic men had higher rates of kidney disease compared to non-hypouricemic subjects. However, the rates of other diseases including diabetes and urinary stones were not significantly different between hypouricemic and non-hypouricemic subjects [20]. In this regard, our longitudinal study showed that hypouricemia did not carry the lowest risk for developing cardiometabolic diseases. Since excess serum urate not only has an adverse effect, but also acts preferably as a reducing substance, this dual nature needs to be considered clinically.

This study showed a positive association between serum urate and cardiometabolic diseases, but most Mendelian randomization studies or meta-analyses suggested that elevated serum urate was only associated with gout [25–29]. However, Mendelian studies are often limited by not considering other influencing conditions, such as life habits including food, alcohol, and fructose intake. Most hyperuricemia is mainly acquired by life habits except for some genetic diseases [30], and it is therefore difficult to apply the results from Mendelian studies to most acquired hyperuricemic subjects. The gap

of results between clinical studies and genetic studies suggest that acquired hyperuricemia may cause more cardiometabolic diseases than genetic hyperuricemia.

Our study has several limitations. First, this study is a retrospective single center study, which may have introduced selection bias. However, single center studies had some advantages of the similarity of the methodology. Second, we could not check the additional and withdrawal medication or gouty attacks over the periods. Some hyperuricemic subjects with gouty attacks might have medication especially non-steroidal anti-inflammatory drugs (NSAIDs), which might cause CKD or hypertension. However, our definition of each disease included medication use. Moreover, we did not exclude the subjects on medication for hyperuricemia or gout intentionally, because the serum urate levels on medication could be useful to evaluate the effects of serum urate on cardiometabolic diseases. However, there is a possibility of the influence of urate-lowering medications on the development (or prevention) of other cardiometabolic diseases, which thus may bias the present results. We additionally conducted the sensitivity analyses that excluded 373 (2.9%) subjects with urate-lowering medications (Supplementary Table S1). The results showed the same results, thus supporting our main results more robustly. However, this study was not able to show whether urate-lowering medications could prevent cardiometabolic diseases or not because this study was an observational study. We had to adjust the patient backgrounds between the medication group and the control group to show the efficacy of urate-lowering medications for hyperuricemia to prevent cardiometabolic diseases. Third, this longitudinal study lacks time-to-event data, which precluded survival analysis. Fourth, we measured serum urate only once, and blood pressure only at the center. Serum urate can fluctuate for natural or iatrogenic causes. Moreover, some hypertensive subjects might have white-coat hypertension and some non-hypertensive subjects might have masked hypertension. Measuring serum urate many times and ambulatory blood pressure monitoring are the best to evaluate serum urate and blood pressure precisely, but it is difficult in practice in the setting of an annual medical examination. Fifth, the number of hypouricemic subjects was small, and it might be less power to analyze the significant difference. Therefore, it is difficult to discuss the J-curve phenomenon precisely. Finally, causality cannot be inferred, because this is an observational study. Interventional studies are needed to further clarify whether the treatments for hyperuricemia are useful for preventing the development of cardiometabolic diseases.

5. Conclusions

Even in the normal range, having higher serum urate could be a risk for hypertension, dyslipidemia, and CKD. The optimal serum urate range, which conferred the lowest risk for developing cardiometabolic diseases, could be less than 5 mg/dL for men and 2–4 mg/dL for women in a generally healthy population. These findings suggest that routine screening of serum urate is useful as a predictor for cardiometabolic diseases in primary care settings.

Supplementary Materials: The following are available online at http://www.mdpi.com/2077-0383/9/4/942/s1, Table S1. The sensitivity analyses that excluded 373 (2.9%) subjects with urate-lowering medications: Lower serum

Author Contributions: Conceptualization, M.K., I.H., R.J.J., and M.A.L.; Data Curation, M.K. and K.N.; Formal Analysis, M.K.; Investigation, M.K. and K.N.; Methodology, M.K., I.H., R.J.J. and M.A.L.; Software, M.K.; Validation, M.K. and K.N.; Visualization, M.K.; Writing—Original Draft M.K. and P.B.; Preparation, M.K.; Writing—Review and Editing, M.K., I.H., K.N., P.B., C.A.R.-J., A.A.-H., M.K., R.J.J., and M.A.L.; Supervision, M.K., I.H., K.N., P.B., C.A.R.-J, A.A.-H., M.K., R.J.J., and M.A.L. All authors have read and agreed to the published version of the manuscript.

Funding: This study is supported by the research grant by Toranomon Hospital, Okinaka Memorial Institute for Medical Research, and Gout Research Foundation in Japan.

Acknowledgments: All the authors of this paper fulfill the criteria of authorship. The authors thank the study subjects and all staff in Center for Preventive Medicine, St. Luke's International Hospital, for assistance with data collection.

Conflicts of Interest: Kuwabara reports the research grant by Toranomon Hospital, Okinaka Memorial Institute for Medical Research, and Gout Research Foundation in Japan. Bjornstad reports personal fees and non-financial support from Boehringer-Ingelheim, non-financial support from Sanofi, personal fees from Bayer, grants from Horizon Pharma, other from XORTX, personal fees from NovoNordisk, personal fees from BMS, non-financial support from AstraZenica, outside the submitted work. Johnson reports personal fees from Horizon Pharma, other from Colorado Research Partners LLC, other from XORT Therapeutics, outside the submitted work. In addition, Johnson has a patent US patent 8,557,831 licensed, and a patent US Patent 9,155,740 B licensed. The remaining authors have nothing to disclose.

References

1. Freedman, D.S.; Williamson, D.F.; Gunter, E.W.; Byers, T. Relation of serum uric acid to mortality and ischemic heart disease. The NHANES I Epidemiologic Follow-up Study. *Am. J. Epidemiol.* **1995**, *141*, 637–644. [CrossRef] [PubMed]
2. Ioachimescu, A.G.; Brennan, D.M.; Hoar, B.M.; Hazen, S.L.; Hoogwerf, B.J. Serum uric acid is an independent predictor of all-cause mortality in patients at high risk of cardiovascular disease: A preventive cardiology information system (PreCIS) database cohort study. *Arthritis Rheumatol.* **2008**, *58*, 623–630. [CrossRef] [PubMed]
3. Culleton, B.F.; Larson, M.G.; Kannel, W.B.; Levy, D. Serum uric acid and risk for cardiovascular disease and death: The Framingham Heart Study. *Ann. Intern. Med.* **1999**, *131*, 7–13. [CrossRef] [PubMed]
4. Moriarity, J.T.; Folsom, A.R.; Iribarren, C.; Nieto, F.J.; Rosamond, W.D. Serum uric acid and risk of coronary heart disease: Atherosclerosis Risk in Communities (ARIC) Study. *Ann. Epidemiol.* **2000**, *10*, 136–143. [CrossRef]
5. Sakata, K.; Hashimoto, T.; Ueshima, H.; Okayama, A.; Group, N.D.R. Absence of an association between serum uric acid and mortality from cardiovascular disease: NIPPON DATA 80, 1980-1994. National Integrated Projects for Prospective Observation of Non-communicable Diseases and its Trend in the Aged. *Eur. J. Epidemiol.* **2001**, *17*, 461–468. [CrossRef] [PubMed]
6. De Leeuw, P.W.; Thijs, L.; Birkenhager, W.H.; Voyaki, S.M.; Efstratopoulos, A.D.; Fagard, R.H.; Leonetti, G.; Nachev, C.; Petrie, J.C.; Rodicio, J.L.; et al. Prognostic significance of renal function in elderly patients with isolated systolic hypertension: Results from the Syst-Eur trial. *J. Am. Soc. Nephrol.* **2002**, *13*, 2213–2222. [CrossRef]
7. Kuwabara, M.; Hisatome, I.; Niwa, K.; Hara, S.; Roncal-Jimenez, C.A.; Bjornstad, P.; Nakagawa, T.; Andres-Hernando, A.; Sato, Y.; Jensen, T.; et al. Uric Acid Is a Strong Risk Marker for Developing Hypertension From Prehypertension: A 5-Year Japanese Cohort Study. *Hypertension* **2018**, *71*, 78–86. [CrossRef]
8. Kuwabara, M.; Niwa, K.; Hisatome, I.; Nakagawa, T.; Roncal-Jimenez, C.A.; Andres-Hernando, A.; Bjornstad, P.; Jensen, T.; Sato, Y.; Milagres, T.; et al. Asymptomatic Hyperuricemia Without Comorbidities Predicts Cardiometabolic Diseases: Five-Year Japanese Cohort Study. *Hypertension* **2017**, *69*, 1036–1044. [CrossRef]
9. Kuwabara, M.; Kuwabara, R.; Hisatome, I.; Niwa, K.; Roncal-Jimenez, C.A.; Bjornstad, P.; Andres-Hernando, A.; Sato, Y.; Jensen, T.; Garcia, G.; et al. "Metabolically Healthy" Obesity and Hyperuricemia Increase Risk for Hypertension and Diabetes: 5-year Japanese Cohort Study. *Obesity (Silver Spring)* **2017**, *25*, 1997–2008. [CrossRef]
10. Kuwabara, M.; Hisatome, I.; Roncal-Jimenez, C.A.; Niwa, K.; Andres-Hernando, A.; Jensen, T.; Bjornstad, P.; Milagres, T.; Cicerchi, C.; Song, Z.; et al. Increased Serum Sodium and Serum Osmolarity Are Independent Risk Factors for Developing Chronic Kidney Disease; 5 Year Cohort Study. *PLoS ONE* **2017**, *12*, e0169137. [CrossRef]
11. Kuwabara, M.; Bjornstad, P.; Hisatome, I.; Niwa, K.; Roncal-Jimenez, C.A.; Andres-Hernando, A.; Jensen, T.; Milagres, T.; Sato, Y.; Garcia, G.; et al. Elevated Serum Uric Acid Level Predicts Rapid Decline in Kidney Function. *Am. J. Nephrol.* **2017**, *45*, 330–337. [CrossRef] [PubMed]
12. Kuwabara, M.; Niwa, K.; Nishihara, S.; Nishi, Y.; Takahashi, O.; Kario, K.; Yamamoto, K.; Yamashita, T.; Hisatome, I. Hyperuricemia is an independent competing risk factor for atrial fibrillation. *Int. J. Cardiol.* **2017**, *231*, 137–142. [CrossRef] [PubMed]
13. Kuwabara, M.; Niwa, K.; Nishi, Y.; Mizuno, A.; Asano, T.; Masuda, K.; Komatsu, I.; Yamazoe, M.; Takahashi, O.; Hisatome, I. Relationship between serum uric acid levels and hypertension among Japanese individuals not treated for hyperuricemia and hypertension. *Hypertens. Res.* **2014**, *37*, 785–789. [CrossRef] [PubMed]
14. Kuwabara, M.; Motoki, Y.; Sato, H.; Fujii, M.; Ichiura, K.; Kuwabara, K.; Nakamura, Y. Low frequency of toothbrushing practices is an independent risk factor for diabetes mellitus in male and dyslipidemia in female: A large-scale, 5-year cohort study in Japan. *J. Cardiol.* **2017**, *70*, 107–112. [CrossRef] [PubMed]

15. Kuwabara, M.; Motoki, Y.; Ichiura, K.; Fujii, M.; Inomata, C.; Sato, H.; Morisawa, T.; Morita, Y.; Kuwabara, K.; Nakamura, Y. Association between toothbrushing and risk factors for cardiovascular disease: A large-scale, cross-sectional Japanese study. *BMJ Open* **2016**, *6*, e009870. [CrossRef] [PubMed]
16. Shimamoto, K.; Ando, K.; Fujita, T.; Hasebe, N.; Higaki, J.; Horiuchi, M.; Imai, Y.; Imaizumi, T.; Ishimitsu, T.; Ito, M.; et al. The Japanese Society of Hypertension Guidelines for the Management of Hypertension (JSH 2014). *Hypertens. Res.* **2014**, *37*, 253–390. [CrossRef]
17. Black, H.R.; Sica, D.; Ferdinand, K.; White, W.B. Eligibility and Disqualification Recommendations for Competitive Athletes With Cardiovascular Abnormalities: Task Force 6: Hypertension: A Scientific Statement from the American Heart Association and the American College of Cardiology. *J. Am. Coll. cardiol.* **2015**, *66*, 2393–2397. [CrossRef]
18. Teramoto, T.; Sasaki, J.; Ishibashi, S.; Birou, S.; Daida, H.; Dohi, S.; Egusa, G.; Hiro, T.; Hirobe, K.; Iida, M.; et al. Executive summary of the Japan Atherosclerosis Society (JAS) guidelines for the diagnosis and prevention of atherosclerotic cardiovascular diseases in Japan-2012 version. *J. Atheroscler. Thromb.* **2013**, *20*, 517–523. [CrossRef]
19. Matsuo, S.; Imai, E.; Horio, M.; Yasuda, Y.; Tomita, K.; Nitta, K.; Yamagata, K.; Tomino, Y.; Yokoyama, H.; Hishida, A.; et al. Revised equations for estimated GFR from serum creatinine in Japan. *Am. J. Kidney Dis.* **2009**, *53*, 982–992. [CrossRef]
20. Kuwabara, M.; Niwa, K.; Ohtahara, A.; Hamada, T.; Miyazaki, S.; Mizuta, E.; Ogino, K.; Hisatome, I. Prevalence and complications of hypouricemia in a general population: A large-scale cross-sectional study in Japan. *PLoS ONE* **2017**, *12*, e0176055. [CrossRef]
21. Verdecchia, P.; Schillaci, G.; Reboldi, G.; Santeusanio, F.; Porcellati, C.; Brunetti, P. Relation between serum uric acid and risk of cardiovascular disease in essential hypertension. The PIUMA study. *Hypertension* **2000**, *36*, 1072–1078. [CrossRef] [PubMed]
22. Mazza, A.; Zamboni, S.; Rizzato, E.; Pessina, A.C.; Tikhonoff, V.; Schiavon, L.; Casiglia, E. Serum uric acid shows a J-shaped trend with coronary mortality in non-insulin-dependent diabetic elderly people. The CArdiovascular STudy in the ELderly (CASTEL). *Acta Diabetol.* **2007**, *44*, 99–105. [CrossRef] [PubMed]
23. Sugihara, S.; Hisatome, I.; Kuwabara, M.; Niwa, K.; Maharani, N.; Kato, M.; Ogino, K.; Hamada, T.; Ninomiya, H.; Higashi, Y.; et al. Depletion of Uric Acid Due to SLC22A12 (URAT1) Loss-of-Function Mutation Causes Endothelial Dysfunction in Hypouricemia. *Circ. J.* **2015**, *79*, 1125–1132. [CrossRef] [PubMed]
24. Waring, W.S.; McKnight, J.A.; Webb, D.J.; Maxwell, S.R. Uric acid restores endothelial function in patients with type 1 diabetes and regular smokers. *Diabetes* **2006**, *55*, 3127–3132. [CrossRef]
25. Li, X.; Meng, X.; Timofeeva, M.; Tzoulaki, I.; Tsilidis, K.K.; Ioannidis, J.P.; Campbell, H.; Theodoratou, E. Serum uric acid levels and multiple health outcomes: Umbrella review of evidence from observational studies, randomised controlled trials, and Mendelian randomisation studies. *BMJ* **2017**, *357*, j2376. [CrossRef] [PubMed]
26. Jordan, D.M.; Choi, H.K.; Verbanck, M.; Topless, R.; Won, H.H.; Nadkarni, G.; Merriman, T.R.; Do, R. No causal effects of serum urate levels on the risk of chronic kidney disease: A Mendelian randomization study. *PLoS Med.* **2019**, *16*, e1002725. [CrossRef]
27. Rasheed, H.; Hughes, K.; Flynn, T.J.; Merriman, T.R. Mendelian randomization provides no evidence for a causal role of serum urate in increasing serum triglyceride levels. *Circ. Cardiovasc. Genet.* **2014**, *7*, 830–837. [CrossRef]
28. Kottgen, A.; Albrecht, E.; Teumer, A.; Vitart, V.; Krumsiek, J.; Hundertmark, C.; Pistis, G.; Ruggiero, D.; O'Seaghdha, C.M.; Haller, T.; et al. Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. *Nat. Genet.* **2013**, *45*, 145–154. [CrossRef]
29. Yang, Q.; Kottgen, A.; Dehghan, A.; Smith, A.V.; Glazer, N.L.; Chen, M.H.; Chasman, D.I.; Aspelund, T.; Eiriksdottir, G.; Harris, T.B.; et al. Multiple genetic loci influence serum urate levels and their relationship with gout and cardiovascular disease risk factors. *Circ. Cardiovasc. Genet.* **2010**, *3*, 523–530. [CrossRef]
30. Johnson, R.J.; Segal, M.S.; Sautin, Y.; Nakagawa, T.; Feig, D.I.; Kang, D.H.; Gersch, M.S.; Benner, S.; Sanchez-Lozada, L.G. Potential role of sugar (fructose) in the epidemic of hypertension, obesity and the metabolic syndrome, diabetes, kidney disease, and cardiovascular disease. *Am. J. Clin. Nutr.* **2007**, *86*, 899–906. [CrossRef]

Journal of
Clinical Medicine

MDPI

Article

Prospective Evaluation of Intensity of Symptoms, Therapeutic Procedures and Treatment in Palliative Care Patients in Nursing Homes

Daniel Puente-Fernández [1], Concepción B. Roldán-López [2], Concepción P. Campos-Calderón [3,*], Cesar Hueso-Montoro [4], María P. García-Caro [4] and Rafael Montoya-Juarez [4]

[1] Doctoral Program of Clinical Medicine and Public Health, University of Granada, 18071 Granada, Spain; danielpuentefdz@correo.ugr.es
[2] Department of Statistics and Operational Research, Faculty of Medicine, University of Granada, 1016 Granada, Spain; iroldan@ugr.es
[3] Alicante Biomedical Research Institute (ISABIAL), 03010 Alicante, Spain
[4] Department of Nursing, Faculty of Health Sciences, Mind, Brain and Behaviour Research Institute, University of Granada, 18016 Granada, Spain; cesarhueso@ugr.es (C.H.-M.); mpazgc@ugr.es (M.P.G.-C.); rmontoya@ugr.es (R.M.-J.)
* Correspondence: concha_campos@hotmail.com; Tel.: +34-958244871

Received: 22 January 2020; Accepted: 8 March 2020; Published: 10 March 2020

Abstract: The aim of the study is to evaluate the intensity of symptoms, and any treatment and therapeutic procedures received by advanced chronic patients in nursing homes. A multi-centre prospective study was conducted in six nursing homes for five months. A nurse trainer selected palliative care patients from whom the sample was randomly selected for inclusion. The Edmonton Symptoms Assessment Scale, therapeutic procedures, and treatment were evaluated. Parametric and non-parametric tests were used to evaluate month-to-month differences and differences between those who died and those who did not. A total of 107 residents were evaluated. At the end of the follow-up, 39 had (34.6%) died. All symptoms ($p < 0.050$) increased in intensity in the last week of life. Symptoms were more intense in those who had died at follow-up ($p < 0.05$). The use of aerosol sprays ($p = 0.008$), oxygen therapy ($p < 0.001$), opioids ($p < 0.001$), antibiotics ($p = 0.004$), and bronchodilators ($p = 0.003$) increased in the last week of life. Peripheral venous catheters ($p = 0.022$), corticoids ($p = 0.007$), antiemetics ($p < 0.001$), and antidepressants ($p < 0.05$) were used more in the patients who died. In conclusion, the use of therapeutic procedures (such as urinary catheters, peripheral venous catheter placement, and enteral feeding) and drugs (such as antibiotics, anxiolytics, and new antidepressant prescriptions) should be carefully considered in this clinical setting.

Keywords: palliative care; nursing homes; symptom assessment; drug therapy; therapeutics; longitudinal studies

1. Introduction

The World Health Organization (WHO) [1] and the European Association of Palliative Care (EAPC) [2] encourage an increase in the quality of dying in long-term care settings. In fact, several articles call for more research on end-of-life interventions in these centres, in order to improve care practice [3,4]. Meanwhile, nursing homes have become a plausible alternative in situations where the home is not the most suitable place for the end of life, due to clinical complexity or lack of resources [5].

Recent studies have indicated that there is a high prevalence of physical and psychological symptoms in nursing homes [3,6–8]. All of these symptoms increase in intensity and prevalence as the end of life approaches [4,9]. Most of the studies that have evaluated end-of-life symptoms in nursing homes are retrospective studies [3,6,9–11]. They may exhibit selection bias and problems caused by

poorly recorded or unrecorded data. Prospective studies may be very helpful to properly assess the changes in symptom control when is death is about to occur.

Hospices in Spain are not widely developed, so end-of-life care must be provided by other institutions. In the case of elderly patients, this care is mostly provided by nursing homes. In these centres, most of the beds are privately funded (71%) [12], although some are partially government-funded. In Andalusia, only nursing homes with more than 60 beds are required to offer twenty-four hour nursing services and their own medical care [13].

Beyond this, little is known regarding routine therapeutic procedures and pharmacological treatments in palliative patients in nursing homes. In a recent retrospective study in Spanish hospitals, patients who were at the end stage of their lives received similar therapeutic and diagnosis procedures to acute care patients [14]. This is congruent with other papers published previously: procedures such as catheter insertion, the use of aspirators, and other actions that are common for patient care in a general hospital can make the difference between comfort and discomfort for end-of-life patients [8,15,16].

Regarding pharmacological treatments, a recent review highlighted that many patients continue to receive medications that are not prescribed as palliative treatments or for symptom control, despite being in the end stage of life [17]. A previous review [18] pointed out that few studies focus on pharmacological de-prescription in end-of-life and concluded that life expectancy is not often used as a criterion for medication discontinuation, even though unnecessary drugs might cause side effects that may increase suffering for patients.

In this context, the European Association of Palliative Care [19] emphasises that, in Spain, there are no specific documents on palliative care in long-term care facilities, nor publications regarding the provision of palliative care in this type of centres in Spain.

The purpose of this study is to prospectively explore perception of symptom control, pharmacological treatments, and therapeutic procedures received by palliative patients admitted to nursing homes in the last six months of life. This is one of the first studies to use a prospective approach, and the first one to show the end-of-life situation in Spanish nursing homes with this methodology. We hypothesize that, when death is near, intensity of symptoms and pharmacological treatments linked to symptom control will increase, whereas the frequency of routine therapeutic procedures will be the same as in previous months.

2. Experimental Section

2.1. Design

This is a multi-centre prospective study which has been conducted in nursing homes in Spain.

2.2. Sample

Six nursing homes were selected for convenience based on their institutional characteristics: Presence of a multidisciplinary team, the possible involvement of professionals, and the presence of both public and private beds. All centres included in the study have more than 50 beds. In each centre, one or two nurses with close knowledge of the patient who have been working at the nursing home for at least 6 months were responsible for data collection. All of the nurses that participated signed an informed consent form and received training prior to data collection. In order to control bias and to produce reliable data for the research, these professionals completed a training course designed to explain the study, to ensure that the same data collection methods were followed, and to avoid the dropout of patients at the follow-up stage. The research team was in contact with them via email, and they visited the centres regularly, i.e., at least once a month.

2.3. Recruitment

Each nursing home nurse recruited residents with chronic diseases that met the following criteria according to the Spanish Society of Palliative Care (SECPAL):

- Advanced, progressive, and incurable disease
- Little to no possibility of response to any specific treatment
- Presence of numerous problems or intense, multiple, multifactorial and changing symptoms
- Great emotional impact on patient, family, and staff
- Life expectancy limited to 6 months.

Within each nursing home, twenty patients were randomly selected among all the patients that met the aforementioned inclusion criteria. They were observed and the data of interest were recorded without interfering with the natural course of events. Data were collected between June 2016 and January 2017. All participants, patients, or representatives of patients (in the case of cognitively impaired patients) were fully informed and signed informed consent forms.

2.4. Data Collection Procedure

Nurses collected demographic and clinical information from the clinical records of the patients. A structured questionnaire was used to collect socio-demographic (age, gender, years in the centre, marital status, and number of children) and clinical (medical diagnosis, Charlson Comorbidity Index) data from patient records.

For the systematic symptom assessment, we used the Edmonton Symptom Assessment System (ESAS) [20]. The ESAS has been validated for both patient and care partner report in different settings, including those with older people with multiple morbidities [21]. ESAS was used regularly in all the nursing homes that participated in the study for symptom assessment. The patient version of the ESAS was self-administered by cognitively intact patients. For cognitively impaired residents, the professional version of the ESAS was completed by trained nurses. The relatives of the patients were not involved in data collection. Cognitive impairment was defined as the patient making three or more mistakes in the Pfeiffer test. The Pfeiffer test was used in all the nursing homes that participated in this study.

The prescription of therapeutic procedures such as urinary catheterisation, enteral feeding, peripheral venous catheter placement, use of aerosol sprays, oxygen therapy, and pharmacological treatments such as non-opioid analgesics, opioid analgesics, antibiotics, bronchodilators, corticosteroids, antiemetics, antihistamines, antidepressants, anxiolytics, hypnotics, and barbiturates was also evaluated.

Data were collected between June 2016 and January 2017. For this study, outcome data were collected from clinical records of the first month (T1) and of the following months (T2, T3, T4, and T5) if residents were still alive. For all the residents who died within these six months, the same data were collected from the clinical records of the last week before death (CT = closure test). All participants, residents, and the care partner were fully informed and signed informed consent forms.

2.5. Statistical Analysis

A descriptive analysis was carried out to describe the main characteristics of the study sample. Numerical variables were described with the mean and standard deviation (SD) and the median and interquartile range (P25-P75). Categorical variables were described using absolute frequencies and percentages. Quantitative data were assessed for normality using the Kolmogorov-Smirnov test, and all of the quantitative data collected were found to deviate significantly from the normal distribution ($p < 0.001$). Due to this, non-parametric inferential tests were used. Pearson's chi-squared test was used to evaluate between-group differences and McNemar's test was used to compare the prevalence rates. Wilcoxon's signed-rank test was used in order to compare month-by-month the symptoms reported using the ESAS for nursing home residents. Statistical analyses were conducted using IBM SPSS v.24. p-values of less than 0.05 were considered to be statistically significant.

2.6. Statement of Ethics

All participants (or when appropriate, a representative) signed a form to give their informed consent. The study received the approval of the Research Ethics Committee (PI-0619-2011). In compliance with Spanish Law (Article 16, Law 41/2002), patients' data were anonymised.

3. Results

Thirteen patients dropped out of the study. Two of them moved to another nursing home. Ten patients or representatives of patients refused to give their informed consent during follow-up. One of the residents died before the beginning of the follow-up. As a result, the final sample consisted of 107 residents. Most of them were women (63.6%) and they had a mean age of 84.6 (SD = 7.4) years. The characteristics can be seen in Table 1.

Table 1. Socio-demographic and clinical characteristics of the patients.

Socio-Demographic and Clinical Characteristics of the Patients.	Total Sample n = 107		Dying within 6 month n = 39		Alive ≥ 6 month n = 68		p
Age, md (P25-P75)	84	(81–89)	86	(83–95)	84	(78.5–89)	0.011[1]
Female, *n* (%)	68	(63.6)	24	(64.8)	44	(62.9)	0.835[1]
Years in the centre, md (P25-P75)	2	(1–4)	2	(0.6–5)	2	(1.3–4)	0.946[2]
Marital status widower, *n* (%)	63	(60.8)	25	(67.6)	38	(54.3)	0.012[2]
Number of children, md (P25-P75)	2	(0–3)	2	(0–2)	3	(2–5)	0.222[1]
CCI, md (P25-P75)	3.5	(2–6)	4	(4–6)	3	(2–5)	0.007[1]
Primary diagnosis							
Myocardial infarction, *n* (%)	6	(5.6)	2	(5.3)	4	(6.0)	1.000[2]
Heart failure, *n* (%)	28	(26.2)	12	(31.6)	16	(23.3)	0.492[2]
Peripheral vascular disease, *n* (%)	10	(9.3)	9	(23.7)	1	(1.5)	0.000[2]
Thromboembolic disease, *n* (%)	7	(6.5)	6	(15.8)	1	(1.5)	0.009[2]
Stroke or other cerebral lesions, *n* (%)	45	(42.1)	22	(57.9)	23	(34.3)	0.024[2]
Hemiplegia, *n* (%)	14	(13.1)	7	(18.4)	7	(10.4)	0.370[2]
Arterial hypertension, *n* (%)	63	(58.9)	15	(38.5)	41	(60.3)	0.044[2]
Dementia	51	(47.7)	20	(52.6)	43	(64.2)	0.301[2]
COPD, *n* (%)	25	(23.4)	12	(31.6)	13	(19.4)	0.233[2]
Arrhythmia, *n* (%)	21	(19.6)	14	(36.8)	7	(10.4)	0.002[2]
Renal disease, *n* (%)	19	(17.8)	6	(15.8)	13	(19.4)	0.794[2]
Diabetes, *n* (%)	34	(31.8)	10	(27.8)	24	(38.1)	0.380[2]
Tumour, *n* (%)	17	(15.9)	8	(20.5)	9	(13.2)	0.308[2]
Solid tumour with metastasis, *n* (%)	10	(9.3)	4	(10.5)	6	(9.0)	1.000[2]

Charlson Comorbidity Index, CCI; [1] Mann-Whitney *U*-test; [2] Pearson's chi-squared; COPD: chronic obstructive pulmonary disease.

Residents who died within the follow-up period (n = 39, 34.6%) were generally older and widowers, had a higher Charlson comorbidity index (CCI), and had more peripheral vascular and thromboembolic diseases, stroke or other cerebral lesions, arterial hypertension, and arrhythmia.

3.1. Perception of Symptom Intensity

In the comparison from T1 to T5, the perception of intensity of all symptoms was scored as moderate, except nausea and dyspnoea, which were scored as mild. No statistical differences were found in symptom intensity between T1 and T2 to T5 (Table 2). However, all differences were found to be statistically significant between T1 and symptoms in the last week of life (CT). In the comparison with CT, the median ratings for nausea ($p = 0.040$) and depression ($p = 0.033$) increased by up to 2 points; pain ($p = 0.026$), fatigue ($p = 0.003$), drowsiness ($p \leq 0.001$), dyspnoea ($p \leq 0.001$), and insomnia ($p = 0.032$) increased by up to 3 points; anxiety ($p = 0.001$), poor appetite ($p \leq 0.001$), and malaise ($p = 0.004$) increased in intensity by up to 4 points. In this case, all symptoms were scored as moderate except nausea, which was scored as mild, and fatigue and malaise, which were scored as severe.

Table 2. Month-by-month comparison of symptoms using Edmonton Symptom Assessment System (ESAS) for residents in nursing homes.

Symptoms	T1 vs. T2 (n = 102)		T1 vs. T3 (n = 95)		T1 vs. T4 (n = 84)		T1 vs. T5 (n = 82)		T1 vs. CT (n = 39)	
	md (P25-P75)	p^1	md (P25-P75)	p^1	md (P25-P75)	p^1	md (P25-P75)	p^1	md (P25-P75)	p^1
Pain										
T1	4 (2–6)	0.563	4 (2–6)	0.934	5 (2–7)	0.718	4 (2–6.5)	0.741	4 (2–7)	0.026
T(2-5) or CT	3.5 (2–6)		3.5 (2–6)		5 (2–7)		5 (2–7.5)		7 (2–9.5)	
Fatigue										
T1	5.5 (3–7)	0.225	5 (2–7)	0.485	5.5 (3–7)	0.443	4 (2.5–8)	0.559	5 (3–7)	0.003
T(2-5) or CT	5.5 (3–8)		5 (2–8)		5.5 (3–8)		5 (2.5–9)		8 (3.5–9)	
Nausea										
T1	0 (0–3)	0.721	0 (0–2)	0.728	0 (0–2.5)	0.228	0 (0–2)	0.836	0 (0-3)	0.040
T(2-5) or CT	0 (0–3)		0 (0-2)		0 (0–1)		1 (0–3)		2 (0–7)	
Depression										
T1	3 (0–6)	0.773	3 (0–6)	0.833	3 (0–6)	0.654	3 (0–6)	0.589	3 (0–7)	0.033
T(2-5) or CT	3 (0–5.5)		2 (0–7)		3.5 (1–6)		3.5 (1–7)		4.5 (1–9)	
Anxiety										
T1	3 (0–6)	0.298	3 (0–5.5)	0.470	3 (0–6)	0.889	3 (0–6)	0.553	3 (0–6)	0.001
T(2-5) or CT	2 (0–5)		2 (0–6)		3 (0–6)		4 (0–6)		7 (1–9)	
Drowsiness										
T1	2 (4–7)	0.357	4 (2–7)	0.777	3.5 (2–7)	0.985	3 (0.5–5)	0.850	4 (3–6)	<0.001
T(2-5) or CT	2 (1–7)		4 (1–6)		4 (1–7)		4 (1–5)		7 (6–10)	
Poor appetite										
T1	3 (0–6)	0.624	3 (0–6)	0.479	3 (0–7)	0.332	3 (0–6)	0.473	3 (1–7)	<0.001
T(2-5) or CT	4 (0–6)		2 (0–5.5)		2 (0–6.5)		2 (0–4)		7 (3–10)	
Malaise										
T1	5 (0–7)	0.114	5 (0–7)	0.284	5 (0–7)	0.210	4 (0–7)	0.357	5 (0–7)	0.004
T(2-5) or CT	4.5 (0–6)		4 (0–6)		5 (1–8)		4 (3–7)		9 (2–9.5)	
Dyspnoea										
T1	1 (0–6)	0.522	1 (0–6)	0.765	1 (0–6)	0.602	1 (0–6)	0.187	4 (0–6)	<0.001
T(2-5) or CT	1 (0–5)		1 (0–6)		0 (0–6.5)		0 (0–5.5)		7 (5–9)	
Insomnia										
T1	2.5 (0–6)	0.991	2 (0–6)	0.480	2.5 (0–7)	0.119	2 (0–6)	0.955	3 (0–7)	0.032
T(2-5) or CT	2 (0–6)		2 (0–5)		3 (0–6)		3 (0–6)		6 (1-9)	

Wilcoxon's signed-rank test[1]; T1: Initial follow-up time; T2, T3, T4, T5: Different follow-up times; CT: Closure Test. Week before death; P25: 25th percentile; P75: 75th percentile.

Residents who died during the follow-up period rated symptom intensity as higher for all symptoms, compared to those who were alive for the entire duration (Table 3).

Table 3. Comparison of symptoms using ESAS in residents of nursing homes who died with those who did not die.

Symptoms	Dying within 6 months n = 39 n(Range)	Alive ≥ 6 months n = 68 n(Range)	p^1
Pain, md (P25-P75)	7 (2–9)	4 (2–6)	0.012
Fatigue, md (P25-P75)	8 (3.5–9)	6 (3–8)	0.005
Nauseas, md (P25-P75)	1 (0–7)	0 (0–1)	0.003
Depression, md (P25-P75)	4 (1–9)	3 (0–6)	0.050
Anxiety, md (P25-P75)	4 (1–9)	3 (0–6)	0.002
Drowsiness, md (P25-P75)	7 (1–9)	4 (0–7)	< 0.001
Poor appetite, md (P25-P75)	7 (6–10)	4 (2–7)	< 0.001
Malaise, md (P25-P75)	9 (2–9.5)	5 (2–7)	< 0.001
Dyspnoea, md (P25-P75)	7 (5–9)	1 (0–6)	< 0.001
Insomnia, md (P25-P75)	6 (1–9)	2 (0–6)	0.011

[1]Mann-Whitney *U*-test.

3.2. Therapeutic Procedures and Pharmacological Treatments

No statistical differences were found in the comparison of therapeutic procedures between T1 and T2 to T5. Nevertheless, the analysis showed significant differences between T1 and CT (Table 4). The most repeated procedures (oxygen therapy ($p \leq 0.001$), use of aerosol sprays ($p = 0.008$), and peripheral venous catheter placement ($p = 0.039$)) had an increase of between 20 and 40 percentage points. Despite

this, the percentage of procedures related to urinary catheters ($p = 1000$) and enteral feeding ($p = 0.221$) was not significantly different between T1 and CT.

Table 4. Comparison of therapeutic procedures and pharmacological treatments by months for patients in nursing homes.

Therapeutic Procedures/Pharmacological Treatments	T1 vs. T2 ($n = 102$)		T1 vs. T3 ($n = 95$)		T1 vs. T4 ($n = 84$)		T1 vs. T5 ($n = 82$)		T1 vs. CT ($n = 37$)		
	%	p^1	%	p^1	%	p^1	%	p^1	%	p^1	95% CI[2]
Therapeutic procedures											
Urinary catheter											
T1	14.7	1.000	14.7	1.000	11.9	0.752	14.6	0.267	21.1	1.000	
T(2-5) or CT	13.7		14.7		14.3		8.5		23.7		
Peripheral venous catheter placement											
T1	26.5	0.860	24.2	0.522	22.9	0.502	24.4	0.480	25.6	0.039	4.1–39.9
T(2-5) or CT	24.5		28.4		18.6		19.5		48.7		
Enteral feeding											
T1	11.8	1.000	14.0	1.000	15.5	1.000	14.6	0.789	5.3	0.221	
T(2-5) or CT	11.8		14.0		15.5		17.1		15.8		
Aerosol sprays											
T1	23.5	0.789	18.9	0.248	19.3	0.267	18.3	0.248	28.2	0.008	11.5–51.9
T(2-5) or CT	21.6		26.4		25.3		22.0		61.5		
Oxygen therapy											
T1	30.4	0.803	28.4	0.450	27.4	0.511	24.4	0.343	36.9	<0.001	17.6–65.3
T(2-5) or CT	32.4		31.9		33.3		29.3		79.5		
Pharmacological treatments											
Non-opioid analgesics											
T1	58.8	0.263	54.7	0.345	54.8	0.404	51.2	0.170	71.8	0.628	
T(2-5) or CT	64.7		61.1		60.7		61.0		64.1		
Opioid analgesics											
T1	12.7	0.375	11.6	0.131	8.3	0.445	12.2	1.000	17.9	<0.001	25.6-57.3
T(2-5) or CT	15.7		16.8		11.9		11.0		61.5		
Antibiotics											
T1	21.6	0.185	20.0	0.109	21.4	0.136	17.1	0.211	30.8	0.004	14.9–53.4
T(2-5) or CT	29.4		29.5		31.0		25.6		66.7		
Bronchodilators											
T1	27.5	0.302	26.3	0.114	28.6	0.814	32.9	0.505	25.6	0.003	12.4–41.2
T(2-5) or CT	32.4		32.6		31.0		29.3		53.8		
Corticosteroids											
T1	20.6	1.000	18.9	1.000	15.5	0.453	17.1	1.000	28.2	0.267	
T(2-5) or CT	21.6		20.0		20.2		18.3		41.0		
Antiemetics											
T1	7.8	1.000	9.5	1.000	6.0	1.000	8.5	0.505	17.9	0.227	
T(2-5) or CT	7.8		9.5		7.1		12.2		30.8		
Antihistamines											
T1	7.8	1.000	6.3	0.500	6.0	1.000	3.7	1.000	10.3	0.248	
T(2-5) or CT	7.8		8.4		7.1		3.7		2.6		
Antidepressants											
T1	33.3	1.000	34.7	0.617	32.5	0.479	33.7	0.131	28.2	0.114	
T(2-5) or CT	33.3		32.6		30.1		27.7		12.8		
Anxiolytics											
T1	32.4	0.773	35.0	0.181	28.6	0.752	29.3	0.386	41.0	0.150	
T(2-5) or CT	30.4		28.4		25.7		34.1		25.6		
Hypnotics/barbiturates											
T1	49.0	0.267	46.3	1.000	41.7	0.383	50.0	0.424	51.3	1.000	
T(2-5) or CT	44.1		46.3		47.6		43.9		71.8		

[1]McNemar's test; [2]Agresti Min 95% confidence interval for p2-p1.; T1: Initial follow-up time.; T2, T3, T4, T5: Different follow-up times.

Regarding pharmacological treatments, no significant differences were found between T1 and T2 to T5. However, some statistical differences were found between T1 and CT (Table 4). Opioid analgesics ($p \leq 0.001$), antibiotics ($p = 0.004$), bronchodilators ($p = 0.003$) had a significant increase in usage, that increase being of 45, 35, and 29 percentage points, respectively.

CT: Closure Test. Week before death. Statistical differences were found in the use of peripheral venous catheters ($p = 0.022$), aerosol sprays ($p = 0.001$), and oxygen therapy ($p = 0.001$) between the patients who died in the follow-up and those who survived (Table 5).

Table 5. Comparison of therapeutic procedures and pharmacological treatments in nursing home patients who died with those who did not die.

Therapeutic Procedures/Pharmacological Treatments	Dying within 6 months, $n = 39$	Alive ≥ 6 months, $n = 68$	p*	OR (95% CI)
Therapeutic procedures				
Urinary catheter, (%)	23.7	13.2	0.176	
Peripheral venous catheter placement, (%)	48.7	25.0	0.022	2.850 (1.238–6.562)
Enteral feeding, (%)	15.8	14.7	0.867	
Aerosol sprays, (%)	61.5	20.6	<0.001	6.171 (2.578–14.771)
Oxygen therapy, (%)	79.5	26.5	<0.001	10.764 (4.181–27.713)
Pharmacological treatments				
Non-opioid analgesics	65.8	54.4	0.350	
Opioid analgesics	63.2	10.3	<0.001	14.939 (5.372–41.546)
Antibiotics	65.8	16.2	<0.001	9.965 (3.930–25.268)
Bronchodilators	55.3	16.2	<0.001	6.401 (2.580–15.880)
Corticosteroids	42.1	16.2	0.007	3.769 (1.514–9.379)
Antiemetics	31.6	4.4	<0.001	10.000 (2.607–38.359)
Antihistamines	2.6	7.4	0.417	
Antidepressants	13.2	33.3	0.026	.278 (0.096–0.805)
Anxiolytics	26.3	29.4	0.909	
Hypnotics/barbiturates	52.6	48.5	0.839	

*Pearson's chi-squared; OR (95% CI), odds ratio (95% confidence interval of the odds ratio).

Finally, the comparison of pharmacological treatments showed differences for use of antibiotics ($p < 0.001$), bronchodilators ($p < 0.0001$), opioids ($p < 0.001$), corticosteroids ($p = 0.007$), antiemetics ($p < 0.001$), and antidepressants ($p = 0.026$) between those who died and the survivors (Table 5). The administration of these treatments was significantly greater in all deceased patients than in those who survived, except for antidepressants, whose usage was significantly lower.

4. Discussion

This is one of the first studies that prospectively describes the last months of life of nursing home residents, and the first that has been conducted in Spain. Our results suggest that there is a sudden increase in symptoms, therapeutic procedures, and pharmacological treatments in the last week of life, in comparison with previous follow-up times. In addition, an increasing number of invasive therapeutic procedures, which may result in decreased comfort for residents, was observed, such as urinary catheter placement, peripheral venous catheter placement, and enteral feeding. Similarly, increased drug use, such as antibiotics, anxiolytics, and new antidepressant prescriptions was also observed.

The perception of the intensity of symptoms remains stable between T1 and the following months, but increases substantially between T1 and the last week of life. This finding is consistent with the previous literature, which details a worsening of symptoms in the last days of life [6,9]. Nevertheless, it is necessary to point out that the consulted studies used prevalence, not intensity, to assess symptoms. Thompson et al. [10] conducted a prospective study in which they assessed pain in the last six months of life of residents in nursing homes, showing that the intensity of their pain remained stable during a short follow-up period, except in the last days of life, when it increased.

In the same way, in relation to therapeutic procedures, there are significant differences in the use of peripheral venous catheters, oxygen therapy, enteral feeding, and aerosol sprays in the last week of life compared to at T1. These differences are greater if we compare the therapeutic procedures between patients who died within the follow-up period and survivors. Regarding oxygen therapy and the use of aerosol sprays, Hendriks et al. [4] highlighted that, unlike what the results of the present study show, there was a decrease in the use of these procedures when death was near. Similarly, a retrospective study conducted in four Spanish hospitals [14] showed that oxygen therapy was a very frequently used intervention at the end of life. This study also reported that there is an increase in the use of peripheral venous catheters during the last days of life [14].

Enteral feeding is another intervention that might be considered to be futile [22], as this does not improve the wellbeing of patients in a significant way and may even be prolonging the dying process [22]. One of the factors that may influence the continuation of enteral feeding is that some professionals and relatives consider this intervention to be a measure of comfort that should not be removed [23].

With respect to urinary catheters, previous studies are not clear about the use of these interventions at the end of life. A literature review by Farrington et al. [24], which included clinical practice guidelines, pointed out that, even though some of the studies reviewed stated that urinary catheterisation could be used to provide comfort to patients, this procedure may cause or increase patient discomfort [24].

This could be interpreted as the performance of futile interventions in the last week of life in the nursing homes analysed.

As expected, the use of some medications linked to symptom control such as opioids, corticosteroids, and antiemetics increased in the last days of life. Opioids were one of the most used drugs in this study, which corresponds to what is described in the literature [4,9]. In relation to the use of non-opioid analgesics, Jansen et al., [25] unlike our study, reported an increase in the use of this group of drugs at the end of life.

On the other hand, there is a decrease in the use of antidepressants in the last week of life, although the consumption of other psychotropic drugs remains stable, compared to in previous months. The use of this kind of drug in end-of-life care is controversial: Some of them could be considered futile since they are not used to improve symptoms typical of the end of life [26]. The time delay before certain antidepressants have a noticeable effect is long (usually 4–6 weeks), so their usage may be considered futile for this reason. In fact, although psychotropic drugs may be indicated for the control of psycho-emotional symptoms, authors point out that they can cause undesirable side effects in the geriatric population and an increased risk of mortality [27].

Regarding the use of antibiotics at the end of life, our results indicate an increase in the last week of life. This may be due to the high percentage of patients with dementia in the sample, in whom infections are a common cause of death. Although, previous studies indicate that the use of antibiotics improves the prognosis of patients and the relief of symptoms [28–30], other studies provide evidence that not administering antibiotics improves comfort [31]. Furthermore, using antibiotics is not without risk in fragile patients with chronic diseases, due to drug reactions, drug-drug interactions, and infections [32]. Even so, there is no consensus as to whether or not they should be used at the end of life.

Furthermore, there has been no decrease in the prescription of drugs for any of the drugs evaluated. According to the consulted bibliography, one of those that would be expected to decrease according to current recommendations would be anxiolytics [33]. In our sample, the prescription of anxiolytics did not diminish at the end of life. According to Westbury [33], 'these psychotropic agents should be prescribed cautiously, at the lowest therapeutic doses for as short a time as possible, and be monitored regularly'. The literature consulted shows that identification of the terminal state increases the likelihood of a de-prescription occurring [34]. In the case of nursing homes, this identification is critical for facilitating patients' access to palliative care and, consequently, for improving the quality of care they receive, their satisfaction with it, and their symptoms [35]. Our results may be due to the lack of use of predictive survival tools that could be used in these centres. Therefore, in the absence of a prediction of the end of life, professionals do not question the utility of the interventions that can be carried out.

The present article tried to demonstrate part of the reality of the care provided by Spanish nursing homes, the study of which has had its importance emphasised by institutions, such as the EAPC. It would have been interesting to have assessed patient comfort, in order to clarify the suitability of controversial interventions, due to their possible futility in an end-of-life context. This work is a first approach to the end of life in Spanish nursing homes, being the stepping stone on which it can be developed into an intervention programme to improve end-of-life care in these centres. At the same

time, it could well help to validate specific tools, in order to assess the quality of the dying process and to improve the detection of palliative care needs.

It should also be highlighted that some limitations of this study may affect the reliability of our results. It should be noted that the sample size is small in comparison with other published studies, so it has not been possible to complete further analyses. Furthermore, characteristics of this study's sample are similar to those in other studies conducted in nursing homes regarding age, sex, and diseases [4,6,9,36], so the results should be extrapolated carefully.

In this study, SECPAL criteria were used for case selection. Our results pointed out that only the 36.4% of patients of the sample have died, so a discussion on whether these criteria are the most appropriate is needed, particularly the limitation of a life expectancy of six months.

Several tools have been proposed to identify palliative care needs and prognosis [37]. For instance, White et al. [38] highlighted in a meta-analysis that the accuracy of the 'Surprise Question' referring to a one-year period was higher than 70% in trained professionals. For further studies, a year-long follow-up period could be considered.

5. Conclusions

In this prospective study, intensity of end-of-life symptoms increased in the last week of life. There is also an increase in therapeutic procedures and pharmacological treatments, but not all the procedures and drugs are linked to symptom management. Interventions (such as urinary catheters, peripheral venous catheter placement, and enteral feeding) and drugs (such as antibiotics, anxiolytics, and new antidepressant prescriptions) should be carefully considered in this clinical setting, in order to improve patient comfort and avoid futile treatments.

Primary care workers and stakeholders might support nursing home professionals in order to provide better symptom control and decide which interventions and drugs are to be recommended in the last days of life.

Author Contributions: Conceptualization, C.H.-M., M.P.G.-C. and R.M.-J.; data curation, D.P.-F. and C.P.C.-C.; formal analysis, C.B.R.-L.; investigation, D.P.-F. and C.B.R.-L.; methodology, C.H.-M., M.P.G.-C. and R.M.-J.; project administration, C.P.C.-C..; supervision, C.B.R.-L. and R.M.-J.; writing-original draft, D.P.-F., C.B.R.-L., C.P.C.-C., C.H.-M. and R.M.-J.; writing-review and editing, D.P.-F., C.B.R.-L., C.P.C.-C., C.H.-M. and R.M.-J. All authors have read and agreed to the published version of the manuscript.

Funding: This paper has been partially supported by the Junta de Andalucía, by project FQM-235 and the Andalusian CICYE project AP-0105-2016. The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

Acknowledgments: We would like to thank all the patients, family members and professionals who have made this study possible. This manuscript is part of a PhD of Daniel Puente-Fernandez.

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

References

1. Hall, S.; Petkova, H.; Tsouros, A.D.; Costantini, M.; Higginson, I.J.; World Health Organization. *Palliative Care for Older People: Better Practices*; WHO Regional Office for Europe: Copenhagen, Denmark, 2011; p. 67.
2. Froggatt, K.A.; Reitinger, E.; Heimerl, K.; Hockley, J.; Brazil, K. Palliative Care in Long-Term Care Settings for Older People: EAPC Taskforce. *Eur. J. Palliat. Care* **2013**, *20*, 251–253.
3. Hendriks, S.A.; Smalbrugge, M.; Hertogh, C.M.P.M.; Van Der Steen, J.T. Dying with dementia: Symptoms, treatment, and quality of life in the last week of life. *J. Pain Symptom Manag.* **2014**, *47*, 710–720. [CrossRef] [PubMed]
4. Hendriks, S.A.; Smalbrugge, M.; Galindo-Garre, F.; Hertogh, C.M.P.M.; van der Steen, J.T. From Admission to Death: Prevalence and Course of Pain, Agitation, and Shortness of Breath, and Treatment of These Symptoms in Nursing Home Residents with Dementia. *J. Am. Med. Dir. Assoc.* **2015**, *16*, 475–481. [CrossRef]
5. Costa, V.; Earle, C.C.; Esplen, M.J.; Fowler, R.; Goldman, R.; Grossman, D.; Levin, L.; Manuel, D.G.; Sharkey, S.; Tanueputro, P.; et al. The determinants of home and nursing home death: A systematic review and meta-analysis. *BMC Palliat. Care* **2016**, *15*, 8. [CrossRef]

6. Estabrooks, C.A.; Hoben, M.; Poss, J.W.; Chamberlain, S.A.; Thompson, G.N.; Silvius, J.L.; Norton, P.G. Dying in a Nursing Home: Treatable Symptom Burden and its Link to Modifiable Features of Work Context. *J. Am. Med. Dir. Assoc.* **2015**, *16*, 515–520. [CrossRef]
7. Sandvik, R.K.; Selbaek, G.; Bergh, S.; Aarsland, D.; Husebo, B.S. Signs of Imminent Dying and Change in Symptom Intensity During Pharmacological Treatment in Dying Nursing Home Patients: A Prospective Trajectory Study. *J. Am. Med. Dir. Assoc.* **2016**, *17*, 821–827. [CrossRef] [PubMed]
8. Hoben, M.; Chamberlain, S.A.; Knopp-Sihota, J.A.; Poss, J.W.; Thompson, G.N.; Estabrooks, C.A. Impact of Symptoms and Care Practices on Nursing Home Residents at the End of Life: A Rating by Front-line Care Providers. *J. Am. Med. Dir. Assoc.* **2016**, *17*, 155–161. [CrossRef] [PubMed]
9. Koppitz, A.; Bosshard, G.; Schuster, D.H.; Hediger, H.; Imhof, L. Type and course of symptoms demonstrated in the terminal and dying phases by people with dementia in nursing homes. *Zeitschrift für Gerontol und Geriatr.* **2015**, *48*, 176–183. [CrossRef] [PubMed]
10. Thompson, G.N.; Doupe, M.; Reid, R.C.; Baumbusch, J.; Estabrooks, C.A. Pain Trajectories of Nursing Home Residents Nearing Death. *J. Am. Med. Dir. Assoc.* **2017**, *18*, 700–706. [CrossRef] [PubMed]
11. Smedbäck, J.; Öhlén, J.; Årestedt, K.; Alvariza, A.; Fürst, C.J.; Håkanson, C. Palliative care during the final week of life of older people in nursing homes: A register-based study. *Palliat. Support Care* **2017**, *15*, 417–424. [CrossRef]
12. Campos-Calderón, C.; Montoya-Juárez, R.; Hueso-Montoro, C.; Hernández-López, E.; Ojeda-Virto, F.; García-Caro, M.P. Interventions and decision-making at the end of life: The effect of establishing the terminal illness situation. *BMC Palliat. Care* **2016**, *15*, 91. [CrossRef] [PubMed]
13. Centro de Investigaciones Sociológicas. Informe Envejecimiento en Red. Centro de Investigaciones Sociológicas. Madrid. 2018. Available online: http://envejecimiento.csic.es/documentos/documentos/enred-estadisticasresidencias2017.pdf (accessed on 10 March 2020).
14. *Decreto 168/2007 de 12 de Junio, por el que se Regula el Procedimiento para el Reconocimiento de la Situación de Dependencia y del Derecho a las Prestaciones del Sistema para la Autonomía y Atención a la Dependencia, así como los órganos Competentes para su Valoración. Boletín Oficial de la Junta de Andalucía*. núm. 119. 2007, pp. 38–42. Available online: https://www.juntadeandalucia.es/boja/2007/119/3 (accessed on 10 March 2020).
15. Verhofstede, R.; Smets, T.; Cohen, J.; Eecloo, K.; Costantini, M.; Van Den Noortgate, N.; Deliens, L. End-of-Life Care and Quality of Dying in 23 Acute Geriatric Hospital Wards in Flanders, Belgium. *J. Pain Symptom Manag.* **2017**, *53*, 693–702. [CrossRef] [PubMed]
16. Li, Q.; Zheng, N.T.; Temkin-Greener, H. Quality of end-of-life care of long-term nursing home residents with and without dementia. *J. Am. Geriatr. Soc.* **2013**, *61*, 1066–1073. [CrossRef] [PubMed]
17. Poudel, A.; Yates, P.; Rowett, D.; Nissen, L.M. Use of Preventive Medication in Patients with Limited Life Expectancy: A Systematic Review. *J. Pain Symptom Manag.* **2017**, *53*, 1097–1110. [CrossRef]
18. Tjia, J.; Velten, S.J.; Parsons, C.; Valluri, S.; Briesacher, B.A. Studies to Reduce Unnecessary Medication Use in Frail Older Adults: A Systematic Review. *Drugs Aging* **2013**, *30*, 285–307. Available online: http://link.springer.com/10.1007/s40266-013-0064-1 (accessed on 10 March 2020). [CrossRef]
19. Arias-Casais, N.; Garralda, E.; Rhee, J.; Lima, L.; Pons-Izquierdo, J.; Clark, D.; Hasselaar, J.; Ling, J.; Mosoiu, D.; Centeno-Cortes, C. *EAPC Atlas of Palliative Care in Europe*; EAPC Press: Vilvoorde, Belgium, 2019.
20. Carvajal, A.; Centeno, C.; Watson, R.; Bruera, E. A comprehensive study of psychometric properties of the Edmonton Symptom Assessment System (ESAS) in Spanish advanced cancer patients. *Eur. J. Cancer* **2011**, *47*, 1863–1872. [CrossRef]
21. Nekolaichuk, C.; Watanabe, S.; Beaumont, C. The Edmonton Symptom Assessment System: A 15-year retrospective review of validation studies (1991–2006). *Palliat. Med.* **2008**, *22*, 111–122. [CrossRef]
22. Krishna, L. Nasogastric feeding at the end of life: A virtue ethics approach. *Nurs. Ethics.* **2011**, *18*, 485–494. [CrossRef]
23. Van der Riet, P.; Good, P.; Higgins, I.; Sneesby, L. Palliative care professionals' perceptions of nutrition and hydration at the end of life. *Int. J. Palliat. Nurs.* **2008**, *14*, 145–151. [CrossRef]
24. Farrington, N.; Fader, M.; Richardson, A. Managing urinary incontinence at the end of life: An examination of the evidence that informs practice. *Int. J. Palliat. Nurs.* **2019**, *9*, 449–456. [CrossRef]
25. Jansen, K.; Schaufel, M.A.; Ruths, S. Drug treatment at the end of life: An epidemiologic study in nursing homes. *Scand. J. Prim. Health Care* **2014**, *32*, 187–192. [CrossRef] [PubMed]

26. McNeil, M.J.; Kamal, A.H.; Kutner, J.S.; Ritchie, C.S.; Abernethy, A.P. The Burden of Polypharmacy in Patients near the End of Life. *J. Pain Symptom Manag.* **2016**, *51*, 178–183. [CrossRef] [PubMed]
27. Park, Y.; Franklin, J.M.; Schneeweiss, S.; Levin, R.; Crystal, S.; Gerhard, T.; Huybrechts, K.F. Antipsychotics and mortality: Adjusting for mortality risk scores to address confounding by terminal Illness. *J. Am. Geriatr. Soc.* **2015**, *63*, 516–523. [CrossRef] [PubMed]
28. Rosenberg, J.H.; Albrecht, J.S.; Fromme, E.K.; Noble, B.N.; McGregor, J.C.; Comer, A.C.; Furuno, J.P. Antimicrobial use for symptom management in patients receiving hospice and palliative care: A systematic review. *J. Palliat. Med.* **2013**, *16*, 1568–1574. [CrossRef]
29. Lam, P.T.; Chan, K.S.; Tse, C.Y.; Leung, M.W. Retrospective analysis of antibiotic use and survival in advanced cancer patients with infections. *J. Pain Symptom Manag.* **2005**, *30*, 536–543. [CrossRef]
30. Oh, D.Y.; Kim, J.H.; Kim, D.W.; Im, S.A.; Kim, T.Y.; Heo, D.S.; Kim, N.K. Antibiotic use during the last days of life in cancer patients. *Eur. J. Cancer Care (Engl.)* **2006**, *15*, 74–79. [CrossRef]
31. Givens, J.L.; Jones, R.N.; Shaffer, M.L.; Kiely, D.K.; Mitchell, S.L. Survival and comfort after treatment of pneumonia in advanced dementia. *Arch. Intern Med.* **2010**, *170*, 1102–1107. [CrossRef]
32. Juthani-Mehta, M.; Malani, P.N.; Mitchell, S.L. Antimicrobials at the End of Life: An Opportunity to Improve Palliative Care and Infection Management. *JAMA* **2015**, *314*, 2017–2018. [CrossRef]
33. Westbury, J.L.; Gee, P.; Ling, T.; Brown, D.T.; Franks, K.H.; Bindoff, I.; Peterson, G.M. RedUSe: Reducing antipsychotic and benzodiazepine prescribing in residential aged care facilities. *Med. J. Aust.* **2018**, *208*, 398–403. [CrossRef]
34. Van Den Noortgate, N.J.; Verhofstede, R.; Cohen, J.; Piers, R.D.; Deliens, L.; Smets, T. Prescription and Deprescription of Medication during the Last 48 Hours of Life: Multicenter Study in 23 Acute Geriatric Wards in Flanders, Belgium. *J. Pain Symptom Manag.* **2016**, *51*, 1020–1026. [CrossRef]
35. Stephens, C.E.; Hunt, L.J.; Bui, N.; Halifax, E.; Ritchie, C.S.; Lee, S.J. Palliative Care Eligibility, Symptom Burden, and Quality-of-Life Ratings in Nursing Home Residents. *JAMA Intern Med.* **2018**, *178*, 141–142. [CrossRef] [PubMed]
36. Brandt, H.E.; Ooms, M.E.; Deliens, L.; van der Wal, G.; Ribbe, M.W. The last two days of life of nursing home patients-a nationwide study on causes of death and burdensome symptoms in the Netherlands. *Palliat. Med.* **2006**, *20*, 533–540. [CrossRef] [PubMed]
37. Simmons, C.P.L.; McMillan, D.C.; McWilliams, K.; Sande, T.A.; Fearon, K.C.; Tuck, S.; Fallon, M.T.; Laird, B.J. Prognostic Tools in Patients with Advanced Cancer: A Systematic Review. *J. Pain Symptom Manag.* **2017**, *53*, 962–970. [CrossRef] [PubMed]
38. White, N.; Kupeli, N.; Vickerstaff, V.; Stone, P. How accurate is the 'Surprise Question' at identifying patients at the end of life? A systematic review and meta-analysis. *BMC Med.* **2017**, *15*, 139. [CrossRef]

Journal of
Clinical Medicine

MDPI

Article

Evaluation of Transfer Learning with Deep Convolutional Neural Networks for Screening Osteoporosis in Dental Panoramic Radiographs

Ki-Sun Lee [1,2,3,*], Seok-Ki Jung [4], Jae-Jun Ryu [5], Sang-Wan Shin [6,7] and Jinwook Choi [1,8,*]

1 Department of Biomedical Engineering, College of Medicine, Seoul National University, Seoul 03080, Korea
2 Department of Clinical Dentistry, College of Medicine, Korea University, Seoul 02841, Korea
3 Department of Prosthodontics, Korea University An-san Hospital, Gyung-gi do 15355, Korea
4 Department of Orthodontics, Korea University Ansan Hospital, Gyung-gi do 15355, Korea; jgosggg@korea.ac.kr
5 Department of Prosthodontics, Korea University Anam Hospital, Seoul 02841, Korea; koprosth@unitel.co.kr
6 Department of Advanced Prosthodontics, Graduate School of Clinical Dentistry, Korea University, Seoul 02841, Korea; swshin@korea.ac.kr
7 Institute of Clinical Dental Research, Korea University, Seoul 02841, Korea
8 Institute of Medical & Biological Engineering, Medical Research Center, Seoul 03080, Korea
* Correspondence: kisuns@gmail.com (K.-S.L.); jinchoi@snu.ac.kr (J.C.)

Received: 25 December 2019; Accepted: 30 January 2020; Published: 1 February 2020

Abstract: Dental panoramic radiographs (DPRs) provide information required to potentially evaluate bone density changes through a textural and morphological feature analysis on a mandible. This study aims to evaluate the discriminating performance of deep convolutional neural networks (CNNs), employed with various transfer learning strategies, on the classification of specific features of osteoporosis in DPRs. For objective labeling, we collected a dataset containing 680 images from different patients who underwent both skeletal bone mineral density and digital panoramic radiographic examinations at the Korea University Ansan Hospital between 2009 and 2018. Four study groups were used to evaluate the impact of various transfer learning strategies on deep CNN models as follows: a basic CNN model with three convolutional layers (CNN3), visual geometry group deep CNN model (VGG-16), transfer learning model from VGG-16 (VGG-16_TF), and fine-tuning with the transfer learning model (VGG-16_TF_FT). The best performing model achieved an overall area under the receiver operating characteristic of 0.858. In this study, transfer learning and fine-tuning improved the performance of a deep CNN for screening osteoporosis in DPR images. In addition, using the gradient-weighted class activation mapping technique, a visual interpretation of the best performing deep CNN model indicated that the model relied on image features in the lower left and right border of the mandibular. This result suggests that deep learning-based assessment of DPR images could be useful and reliable in the automated screening of osteoporosis patients.

Keywords: osteoporosis screening; artificial intelligence; convolutional neural networks; dental panoramic radiographs

1. Introduction

Osteoporosis is a systemic disease characterized by low bone mineral density (BMD) and micro-architectural deterioration of bone structure, thereby leading to compromised bone strength and, consequently, an increased risk of fracture [1]. Hip, spine, and wrist fractures caused by osteoporosis often lead to disorders that reduce the quality of life of the patient and, in severe cases, increase the risk of mortality [2,3]. With fast population aging and an increase in life expectancy, osteoporosis is increasingly becoming a global public health issue; it has been estimated that more than 200 million

people are suffering from osteoporosis [4]. According to recent statistics from the International Osteoporosis Foundation, approximately one in three women over the age of 50 will experience osteoporotic fractures, as will one in five men over the age of 50 [4–7]. Moreover, it is expected that more people will be affected by osteoporosis in the future and, consequently, the rate of osteoporotic fractures will increase [8]. This is because the disease initially develops without any symptoms, remains undiagnosed due to scarce symptomatology, and its first manifestation is often a low-energy fracture of long bones or vertebrae [9].

Generally, osteoporosis is diagnosed by evaluating bone mineral density (BMD) measurements (expressed as a T-score) using dual-energy X-ray absorptiometry (DXA), which is considered as the reference-standard examination for BMD assessment [10,11]. However, this technique is complex, expensive, and the availability is limited for overall population diagnosis [12]. Recently, digital images of dental panoramic radiographs (DPRs) have been evaluated as cost-effective and important image data for osteoporosis screening. This is because the widespread use of panoramic radiation in dental care for elderly patients with increased life expectancy and a number of studies have demonstrated the feasibility of BMD estimation and osteoporosis screening using panoramic radiography [13–23].

However, previous approaches primarily relied on manually categorized feature indexes [13–23], such as the Gonion index, mandibular cortical index, mental index, and panoramic mandibular index, and traditional classifier called machine learning (ML) algorithms, such as support vector machine (SVM) [22] and fuzzy classifiers [23], for screening osteoporosis. Although the previously handcrafted feature indexes provided sufficient evidence for assisting osteoporosis screening using panoramic radiographs, these methods for discriminating features are of a low order and do not fully characterize the heterogeneous pattern in radiographic images. In addition, most previous studies require tedious and manual operations, such as extensive preprocessing, image normalization, and region of interest (ROI) segmentation, which can significantly affect the repeatability of the classification method.

In the last few years, deep learning algorithms, particularly deep convolutional neural networks (CNNs) architecture, have been widely recognized as a reliable approach to learn the classification of the characteristics of features directly from original medical images [24,25]. As opposed to ML approaches that rely on explicitly classified features, deep CNNs are a class of deep neural networks that can learn high dimensional features to maximize the networks ability to discriminate abnormalities among images [26]. There are many different CNN architectures that have been designed to perform image classifications and recognitions. Each of these architectures differ in specific aspects, including the number and size of layers, the connections between these layers, and the overall network depth. Because different network architectures are best suited for different problems, and it is difficult to know in advance which architecture is the right choice for a given task, empirical examination is often recognized as the best way to make these decisions [27].

Although deep CNNs have been recognized as efficient tools for image classification, they require a large amount of training data, which can be difficult to apply to medical radiographic image data. When the target dataset is significantly smaller than the base dataset, transfer learning is considered a powerful technique for training deep CNNs without overfitting [28,29]. The general process of transfer learning is performed through the use of pretrained models in a two-step method as follows: First, copying the first n layers of pretrained base network on a general large dataset to the first n layers of a target network and secondly, the remaining layers of the target network are then randomly initialized and trained on a small local dataset toward the target task [28]. On the basis of the transfer learning techniques, several state-of-the-art results showed outperformance in both general image classification [30–32] and medical image classification [33–36]. However, a few studies have been done to develop and evaluate transfer learning-based deep CNN models for predicting osteoporosis in DPRs.

The aim of this study is to develop and evaluate the deep learning approaches for screening osteoporosis with DPR images. Using the classified panoramic radiograph images based on the BMD value (T-score), this study evaluated several different CNN models based on osteoporosis

discriminating accuracy. In addition, we quantitatively evaluated the effect of transfer learning and fine-tuning of a deep CNN model on classifying performance.

2. Patients and Methods

2.1. Patients

The study was done on a total of 680 panoramic radiograph images from 680 different patients who visited the Korea University Ansan Hospital. The patients simultaneously underwent skeletal BMD examinations and digital panoramic radiography evaluations within four months, between 2009 and 2018. The subjects were classified into a non-osteoporosis group (T-score ≥ -2.5) or osteoporosis group (T-score < -2.5), according to the World Health Organization criteria [37], into which 380 and 300 subjects were assigned, respectively. The dataset was divided into training and test sets as follows: The radiographs were selected randomly, and 136 radiographs (20% of the total), 68 each from the osteoporosis and non-osteoporosis groups, were set aside as a test set. This ensured that the testing data set only contained images of novel radiographs that had not been encountered by the model during training. The remaining 544 radiographs were used for the training and validation set. This study protocol was approved by the institutional review board of the Korea University Ansan Hospital (no. 2019AS0126).

2.2. Data Preprocessing

The dimensions of the collected dental X-ray images varied from 1348 to 2820 pixels in width and 685 to 1348 pixels in height. For consistency of image preprocessing, the images were downsampled to a uniform size of 1200 × 630 pixels, using bilinear interpolation. The final ROI was restricted to the lower part of the mandible, below the teeth-containing alveolar bone, for an image size of 700 × 140 pixels (Figure 1). This included the most ROI areas of previous studies [13–23] that applied various classification techniques by detailed and specifically indexing the image feature characteristics of the limited small region of mandible. By setting the ROI to include most of the mandible instead of the specific area of it, this study evaluated the area that plays the most distinctive role in osteoporosis classification through explainable deep learning techniques.

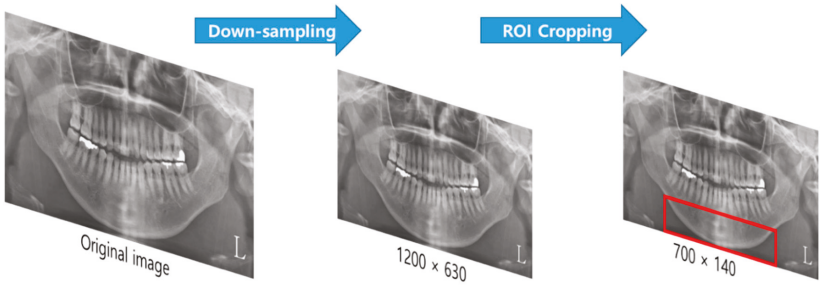


Figure 1. Image preprocessing for this study. The original DPRs were downsampled, and the ROI is restricted to the mandibular region below the teeth (region inside the bounding box). DPR, dental panoramic radiograph; ROI, region of interest.

2.3. Convolutional Neural Networks

This study employed four study groups of CNN as follows: a basic CNN model with three convolutional layers (CNN3), a visual geometry group deep CNN model with no pre-trained weights (VGG16), a transfer learning model from VGG16 with pre-trained weights (VGG16-TF), and a transfer learning and fine-tuning model from VGG16 with pre-trained weights (VGG16-TF-FT). The preceding

architectures, along with the four variant CNN models (CNN3, VGG16, VGG16-TR, and VGG16-TR-FT) used in this study, are depicted in the block diagram in Figure 2.

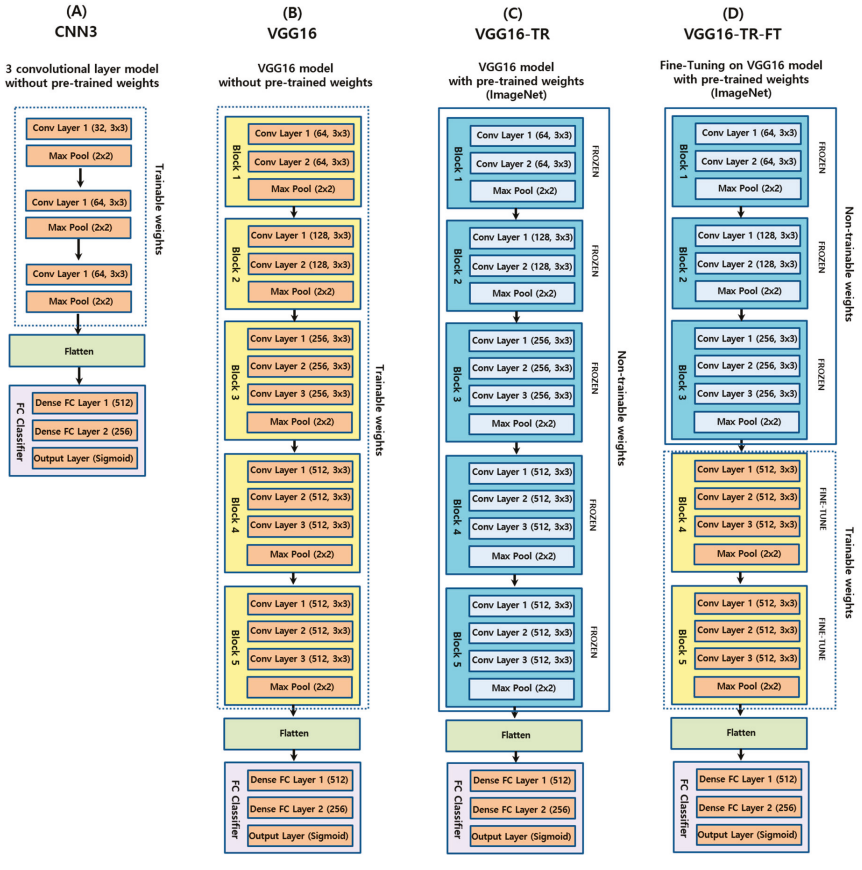


Figure 2. Schematic diagrams of the four convolutional neural networks (CNN) architectures evaluated in this study.

The reason for choosing VGG16 [31] architecture was that it had been widely adopted and recognized as state-of-the-art in both general and medical image classification tasks [24]. Additionally, it has been trained on large-scale datasets, so that a transfer learning approach could be adopted for large-scale image recognition [38]. For the VGG16 architecture under consideration, the following three different experimental groups were evaluated: the native group (VGG16), transfer learning group (VGG16-TR), and transfer learning with fine-tuning group (VGG16-TR-TF). In the native version, model weights were randomly initialized, and training was conducted using only the DPR data described in this study. In the transfer learning version, model weights were fixed, based on pre-training with a general image dataset, except for the final, fully connected layers, which were randomly initialized. In the transfer learning with fine-tuning version, model weights were initialized based on pre-training on a general image dataset, the same as previous versions, except that some of the last blocks were unfrozen so that their weights were updated in each training step. In this study, the last two transfer learning version models (VGG16-TR and VGG16-TR-FT) employed pre-trained weights using the ImageNet database [38]. ImageNet is an image dataset containing thousands of different objects used to train and evaluate image classification models.

2.4. Model Training

The 544 images selected as the training dataset were randomly divided into five folds. This was done to perform 5-fold cross validation to evaluate the model training, while avoiding overfitting or bias [39]. Within each fold, the dataset was partitioned into independent training and validation sets, using an 80 to 20 percentage split. The selected validation set was a completely independent fold from the other training folds and it was used to evaluate the training status during the training. After one model training step was completed, the other independent fold was used as a validation set and the previous validation set was reused, as part of the training set, to evaluate the model training. An overview of the 5-fold cross validation performed in this study is presented in Figure 3.

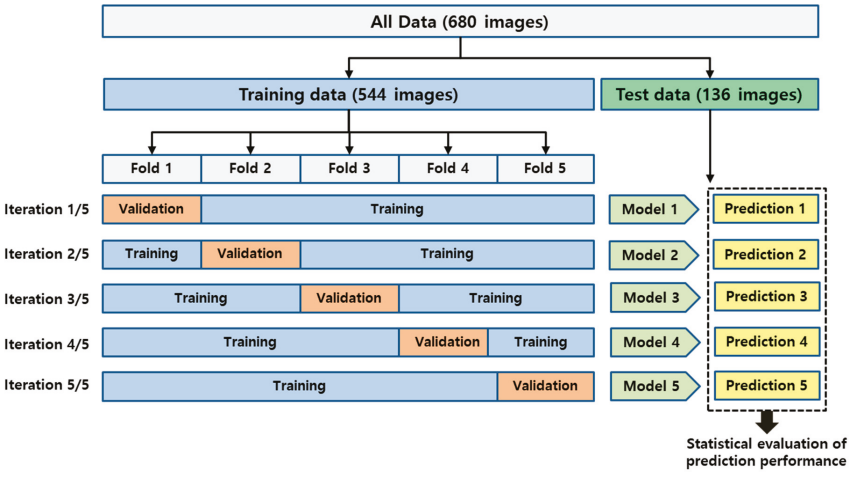


Figure 3. The overview of the performed 5-fold cross validation in this study.

This process was repeated for each architecture (CNN3, VGG16, VGG16-TR, and VGG16-TR-FT). All models were trained and evaluated on a 64-bit Windows 10 operating system, with 64 GB memory and an NVIDIA Quadro P4000 GPU. Building, training, validation, and prediction of deep learning models were performed using the Keras [40] library and TensorFlow [41] backend engine.

2.5. Performance Evaluation

The evaluation of the screening performance of the CNN models was performed with the independent test dataset in each cross-validation fold. To comprehensively evaluate the screening performance on the test dataset, the accuracy, sensitivity, specificity, receiver operating characteristic (ROC) curve, and precision recall (PR) curve were calculated. The accuracy, sensitivity, and specificity score can be calculated as follows:

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FN + FP}$$

$$\text{sensitivity} = \frac{TP}{TP + FN}$$

$$\text{specificity} = \frac{TN}{TN + FP}$$

TP and FP are the number of correctly and incorrectly predicted images, respectively. Similarly, TN and FN represent the number of correctly and incorrectly predicted images, respectively. The area under the ROC curve (AUC) was also calculated.

2.6. Visualizing Model Decisions

Deep learning models have often been referred to as non-interpretable black boxes because it is difficult to know the process by which they make predictions. To know the decision-making process of the model, and which features are most important for the model to screen osteoporosis in DPR images, this study employed the gradient-weighted class activation mapping technique (Grad-CAM) [42] and the most significant regions for screening osteoporosis in DPR images were highlighted.

3. Results

3.1. Baseline Clinical and Demographic Characteristics of the Subjects

The patients were 565 female and 115 male, with an age range from 27 to 90 years (mean age of 63.0 years). There were 380 patients (mean age 58.5) without osteoporosis (T-score ≥ −2.5) and 300 patients (mean age 68.6) with osteoporosis (T-score < −2.5). The clinical characteristics of the DPR dataset used in this study are summarized in Table 1.

Table 1. Clinical and demographic characteristics of the dental panorama radiographs (DPRs) dataset in this study.

Parameter	Without Osteoporosis (T-Score ≥ −2.5)	With Osteoporosis (T-Score < −2.5)	Total
Number of patients	380	300	680
Number of female/male	332/48	233/67	565/115
Mean age (±SD)	58.5 (±11.8)	68.4 (±8.4)	63.0 (±11.6)

3.2. Prediction Performance

The CNN models of this study were trained using a cross-entropy loss function on the selected training image dataset. The screening performances of the four CNN models tested in this study are displayed in Table 2. It was observed that the transfer learning and fine tuning VGG16 model with pre-trained weights (VGG16-TR-FT) achieved the top performance, with the highest AUC of 0.858 (95% CI 0.865 to 0.850), sensitivity of 0.900 (95% CI 0.919 to 0.881), specificity of 0.815 (95% CI 0.847 to 0.783), and accuracy of 0.840 (95% CI 0.857 to 0.822). The screening performances of the other models that applied transfer learning techniques, but did not apply fine tuning, one with pre-trained weights (VGG-TR) and the other without pre-trained weights (VGG16), were slightly degraded. The arbitrarily established model with three convolutional layers (CNN3) achieved the lowest performance, with an AUC of 0.667 (95% CI 0.708 to 0.626), sensitivity of 0.684 (95% CI 0.889 to 0.480), specificity of 0.649 (95% CI 0.813 to 0.486), and accuracy of 0.660 (95% CI 0.725 to 0.594).

Table 2. Osteoporosis screening accuracy of convolutional neural network models in this research.

Model	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Accuracy (95% CI)
CNN3	0.667 (±0.041)	0.684 (±0.204)	0.649 (±0.164)	0.660 (±0.066)
VGG16	0.742 (±0.018)	0.674 (±0.048)	0.811 (±0.034)	0.771 (±0.018)
VGG16-TR	0.782 (±0.006)	0.737 (±0.046)	0.828 (±0.052)	0.802 (±0.024)
VGG16-TR-TF	0.858 (±0.008)	0.900 (±0.019)	0.815 (±0.032)	0.840 (±0.018)

Figure 4 shows the ROC curves of all tested models. The VGG16-TR-FT models achieved the highest AUC of 0.86, while the CNN3 model achieved the lowest AUC of 0.61. Figure 5 shows the PR curves of the tested CNN models. It was also observed that the VGG16-TR-FT models achieved the highest PR of 0.86, while the CNN3 model achieved the lowest PR of 0.61.

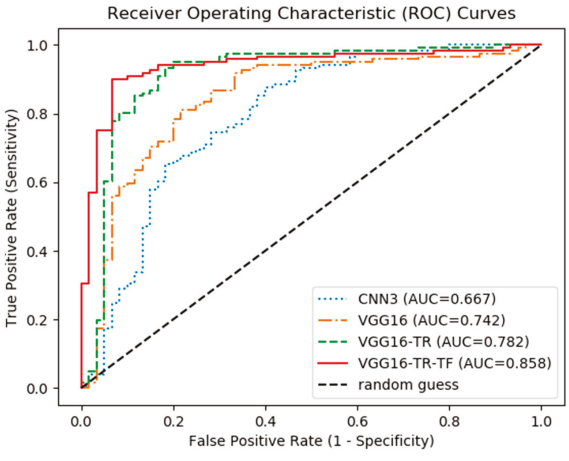


Figure 4. Mean ROC curves of each CNN models for screening osteoporosis on DPR images in this study.

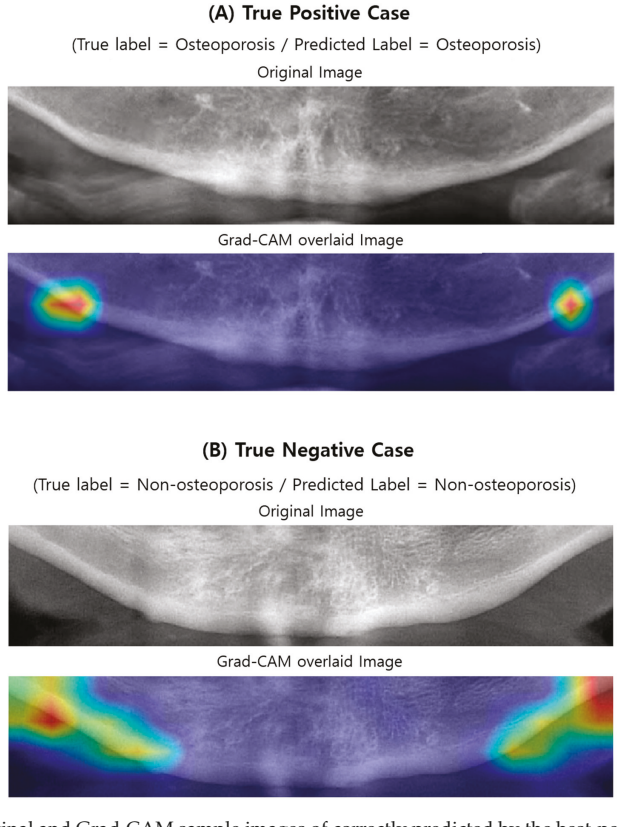


Figure 5. Original and Grad-CAM sample images of correctly predicted by the best-performing deep CNN model (VGG16-TR-TF) for DPR image-based osteoporosis screening are illustrated. Below each original sample images, a Grad-CAM image is superimposed over the original image. The bright red in each Grad-CAM image indicate the region that has the greatest impact on screening osteoporosis patients.

3.3. Visualizing Model Decisions

Figures 5 and 6 illustrate the case examples of predictions using the best predictive VGG16-TR-FT model as compared with ground truth. Each case example employed a Grad-CAM technique to perform a visual interpretation to determine which areas affected the deep CNN's class classification. In the case of screening correctly for osteoporosis (Figure 5A), the region showing the weak lower border of the mandibular cortical bone and the less dense, spongy bone texture at its periphery was extracted as the main image feature of the classification. In correctly screened cases of no osteoporosis (Figure 5B), the region showing the strong lower boundary of the mandible cortical bone and the dense texture around its periphery was extracted as the main image feature of the classification. However, in the case of incorrectly screened cases, i.e., the non-osteoporosis case predicted as osteoporosis (Figure 6A) or the osteoporosis case predicted as non-osteoporosis (Figure 6B), the central region of the mandible or the ghost images of the hyoid bone was extracted as the main image feature.

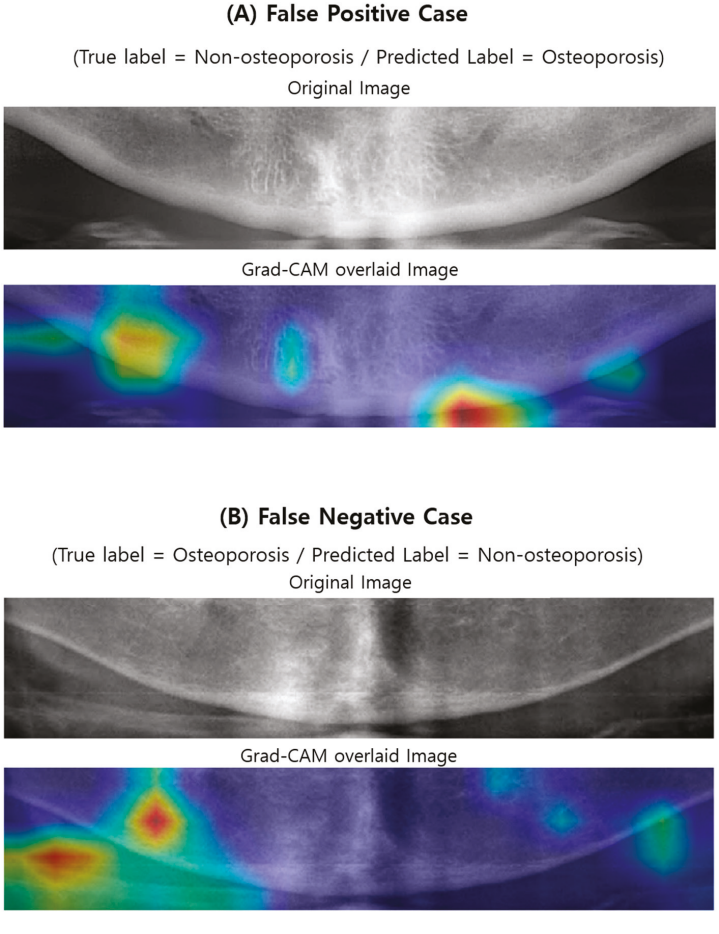


Figure 6. Original and Grad-CAM sample images of incorrectly predicted by the best-performing deep CNN model (VGG16-TR-TF) for DPR image-based osteoporosis screening are illustrated. Below each original sample images, a Grad-CAM image is superimposed over the original image. The bright red in each Grad-CAM image indicate the region that has the greatest impact on screening osteoporosis patients.

4. Discussion

Although DPRs are commonly performed for the evaluation of dentition and adjacent structures of the jaw, some clinical assistant diagnosis (CAD) systems based on DPRs have been suggested for screening systemic diseases, such as osteoporosis and carotid artery calcification [13–23,43]. However, the approaches of most previous studies are only valid when image features are accurately extracted, using sophisticated and manual image preprocessing algorithms or techniques. If a DPR image is imported from an unfamiliar environment or unexpected noise is added to the image, the prediction can easily be distorted. The neural network algorithm can resolve this problem. All the knowledge necessary for diagnosis is established only with the given training image data, without complicated or sophisticated image preprocessing. In recent years, a cutting-edge neural network technology, called deep learning, has been applied to medical imaging analysis and has shown a level of performance that is equal to or better than a clinician. As mentioned above, most previous CAD system studies, which used manual or sophisticated image preprocessing and machine learning algorithms for the screening of osteoporosis based on DPRs, presented variable diagnostic performances, in terms of sensitivity and specificity [13–23]. Recently, a deep learning-based osteoporosis prescreening study, which resulted in a very high AUC score (0.9763 to 0.9991) and accuracy (92.5% to 98.5%), was published [44]. However, in that study, osteoporosis labeling was subjectively performed by dental specialists, rather than BMD score (T-score) which is the gold standard for diagnosing osteoporosis. In addition, the study did not visually interpret the decision of the trained CNN model, and using five arbitrarily established convolutional layers, there is a limitation to the reproducibility of the deep CNN model.

The first major findings of the present study showed that applying appropriate transfer learning and fine-tuning techniques on pre-trained deep CNN architectures had an equivalent DPR-based osteoporosis screening level of previous studies, even with small image datasets, without complex image preprocessing and image ROI settings. According to Table 2 and Figure 4, the CNN3 group, having only arbitrary established three convolutional layers, showed the lowest true-positive screening performance and accuracy among the experimental groups. On the basis of these results, it can be estimated that a CNN model with a small number of convolutional layers can have limitation in learning the true data distribution from a small number of datasets.

Comparing models that used pre-trained weights (VGG16-TR and VGG16-TR-FT) to those that did not (VGG16), also revealed that deep CNNs initialized with large-scale pre-trained weights outperformed those directly learnt from small-scale data, with AUC improvements between 7% to 11%. Thus, in the case of having a small-scale image dataset, this study also suggests that the use of transfer learning on deep CNN models with pre-trained weights can be an efficient solution for the classification of medical images, instead of learning a deep neural network from scratch.

Moreover, as shown in Table 2 and Figure 7, the results of this study also indicated an improvement in screening performance when using fine-tuning on some convolutional blocks in deep CNN layers. In general, the deep CNN model learned from pre-trained deep neural networks on a large natural image dataset could be used to classify common images but cannot be well utilized for specific classifying tasks of medical images (Figure 8A). However, according to a previous study that described the effects and mechanisms of fine tuning on deep CNNs [45], when certain convolutional blocks of a deep CNN model were fine-tuned, the deep CNN model could be further specialized for specific classifying tasks (Figure 8B). More specifically, earlier layers of a deep CNN contain generic features that should be useful to many classification tasks, but later layers progressively contain more specialized features to the details of the classes contained in the original dataset (i.e., the large natural image dataset on which the deep CNN was originally trained). Using this property, when the parameters of the early layers are preserved and the parameters in later layers are updated during training new datasets, the deep CNN model can be effectively used in new classification tasks. In conclusion, fine-tuning uses the parameters learned from a previous training of the network on a large dataset and, then, adjusts the parameters in later layers from the new dataset, improving the performance and accuracy in the new classification task. As with the previous study, the fine-tuning technique, which freezes

the weight parameters of some initial convolutional blocks in the deep CNN model called VGG16, and, then, updates the weight parameters of the later convolutional blocks (Figure 8B), show higher performance than other experimental groups. The conceptual diagram of the fine-tuning technique mentioned above can be seen in Figure 8.

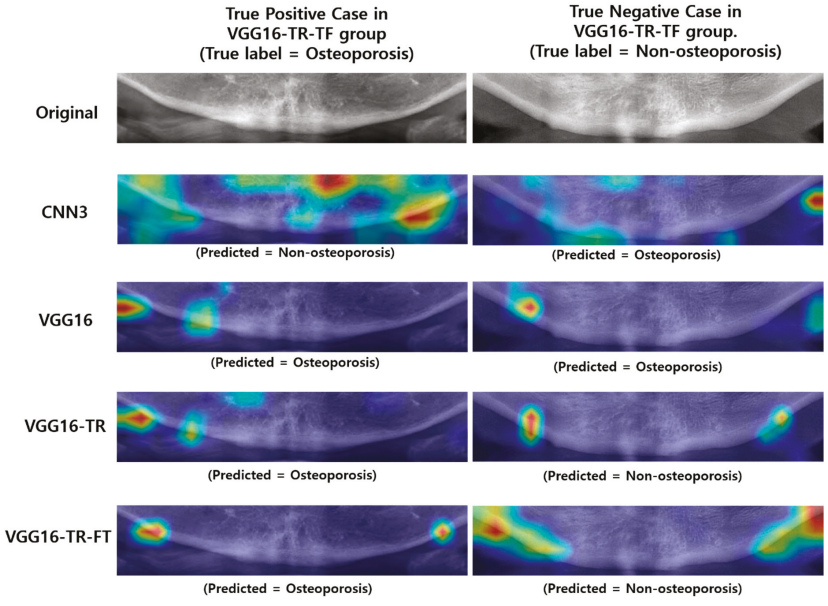


Figure 7. Comparison of grad-CAM images from other groups against some original images showing true positive and true negative in the best performing VGG16-TR-TF group.

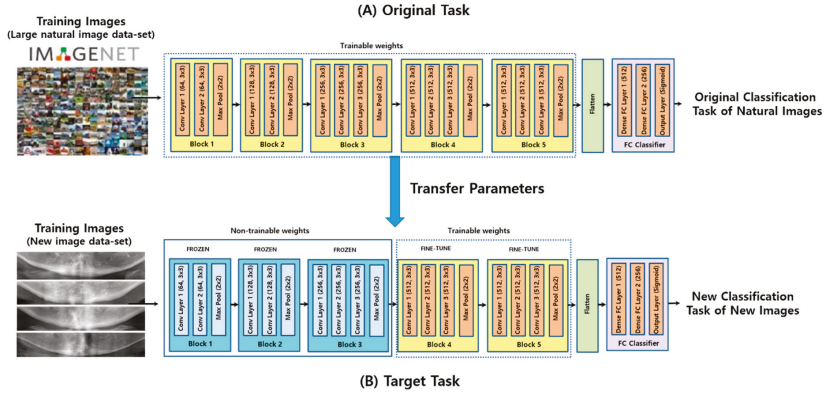


Figure 8. The conceptual diagram of the fine-tuning technique in the transfer learning of a deep CNN.

The second major result of this study was to identify areas where image feature differences occurred when screening osteoporosis in DPR images using the Grad-CAM technique. To understand and visualize the decision of deep CNN models, some samples of the correctly and incorrectly screened examples were reviewed (Figures 5 and 6). For additional insight to model decisions, a Grad-CAM technique was performed in this study. This technique identified the areas of input images that had the greatest impact on model classification. According to this additional review, the model

does seem to identify the feature characteristics of osteoporosis in DPR images (e.g., cortical bone thinning). According to the Grad-CAM evaluation of this study, DPR-based screening performances of osteoporosis were high when the image features were specified in the middle region of the left and right side of the mandibular lower border. This region is also consistent with the regions used to discriminate osteoporosis using DPR images, in most previous studies [13–23], although the measurement algorithm was different. This indicates that most osteoporosis patients have image feature characteristics, on DPR images, at the lower border of the cortical bone in the mandible. However, image quality issues, such as blurring, low contrast, and ghost images of adjacent objects can cause incorrect predictions. When the image features were specified in the center region of the mandible, or when the ghost images of the hyoid bone were in the ROI region, the accuracy was reduced. Therefore, to improve the deep CNN-based screening performance of osteoporosis in DPR images, it is suggested that the ROI setting be limited to the area around the middle of the left and right side of the lower border of the mandible.

5. Conclusions

This study presents the usefulness of transfer learning and fine tuning with a deep CNN for the screening of osteoporosis in DPR images, in cases with a limited training dataset. We have applied various transfer learning techniques on pre-trained networks VGG16 for the discrimination of osteoporosis using a DPR image dataset, labeled based on T-score. The experimental results showed that transfer learning with pre-trained weights and fine-tuning techniques achieved the highest overall accuracy of 84%. The presented results suggest that the combination of the appropriate deep CNN architectures and transfer learning techniques has effectively resolved the issue of a small training set of images and that DPR images have the potential for osteoporosis prescreening. In addition, using the Grad-CAM technique, this study performed a deep learning-based visual explanation for the area where the image feature difference occurred. Therefore, this study confirmed the previous osteoporosis screening studies using DPR images that set the ROI at the middle of the left and right side of the lower border of the mandible. Given the increasing burden of osteoporosis on the global healthcare system, as our population ages, and the proliferation of dental panoramic image devices, the results presented in this study suggest that deep learning-based image analysis of DPRs could serve an important role in cost-effective prescreening for patients unaware of osteoporosis. To further improve screening performance, future research is needed, using different deep CNN architectures and deep learning techniques, more validated and qualified labeled image dataset, the appropriate number of datasets, and automated configuration techniques for more limited range of ROI.

Author Contributions: Conceptualization, K.-S.L., J.-J.R. and S.-W.S.; Data curation, K.-S.L. and S.-K.J.; Formal analysis, K.-S.L.; Funding acquisition, K.-S.L.; Investigation, K.-S.L.; Methodology, K.-S.L.; Project administration, K.-S.L. and J.C.; Software, K.-S.L.; Supervision, J.C.; Validation, S.-K.J. and J.C.; Visualization, K.-S.L.; Writing—original draft, K.-S.L.; Writing—review & editing, K.-S.L. and J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2019R1I1A1A01062961).

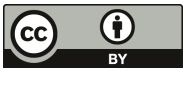
Conflicts of Interest: The authors declare no conflict of interest.

References

1. NIH Consensus Development Panel on Osteoporosis Prevention, Diagnosis, and Therapy, March 7–29, 2000: Highlights of the conference. *South. Med. J.* **2001**, *94*, 569–573.
2. Cauley, J.A. Public health impact of osteoporosis. *J. Gerontol. A Biol. Sci. Med. Sci.* **2013**, *68*, 1243–1251. [CrossRef]
3. Bliuc, D.; Nguyen, N.D.; Nguyen, T.V.; Eisman, J.A.; Center, J.R. Compound risk of high mortality following osteoporotic fracture and refracture in elderly women and men. *J. Bone Miner. Res.* **2013**, *28*, 2317–2324. [CrossRef] [PubMed]

4. Sozen, T.; Ozisik, L.; Basaran, N.C. An overview and management of osteoporosis. *Eur. J. Rheumatol.* **2017**, *4*, 46–56. [CrossRef] [PubMed]
5. Melton, L.J., 3rd; Chrischilles, E.A.; Cooper, C.; Lane, A.W.; Riggs, B.L. Perspective. How many women have osteoporosis? *J. Bone Miner. Res.* **1992**, *7*, 1005–1010. [CrossRef]
6. Melton, L.J., 3rd; Atkinson, E.J.; O'Connor, M.K.; O'Fallon, W.M.; Riggs, B.L. Bone density and fracture risk in men. *J. Bone Miner. Res.* **1998**, *13*, 1915–1923. [CrossRef]
7. Kanis, J.A.; Johnell, O.; Oden, A.; Sembo, I.; Redlund-Johnell, I.; Dawson, A.; De Laet, C.; Jonsson, B. Long-term risk of osteoporotic fracture in Malmo. *Osteoporos. Int.* **2000**, *11*, 669–674. [CrossRef]
8. Kalinowski, P.; Rozylo-Kalinowska, I.; Piskorz, M.; Bojakowska-Komsta, U. Correlations between periodontal disease, mandibular inferior cortex index and the osteoporotic fracture probability assessed by means of the fracture risk assessment body mass index tool. *BMC Med. Imaging* **2019**, *19*, 41. [CrossRef]
9. Marcucci, G.; Brandi, M.L. Rare causes of osteoporosis. *Clin. Cases Miner. Bone Metab.* **2015**, *12*, 151–156. [CrossRef]
10. Kanis, J.A.; Johnell, O. Requirements for DXA for the management of osteoporosis in Europe. *Osteoporos. Int.* **2005**, *16*, 229–238. [CrossRef]
11. Kanis, J.A. Diagnosis of osteoporosis and assessment of fracture risk. *Lancet* **2002**, *359*, 1929–1936. [CrossRef]
12. Mithal, A.; Bansal, B.; Kyer, C.S.; Ebeling, P. The Asia-Pacific Regional Audit-Epidemiology, Costs, and Burden of Osteoporosis in India 2013: A report of International Osteoporosis Foundation. *Indian J. Endocrinol. Metab.* **2014**, *18*, 449–454. [CrossRef] [PubMed]
13. Taguchi, A.; Suei, Y.; Ohtsuka, M.; Otani, K.; Tanimoto, K.; Ohtaki, M. Usefulness of panoramic radiography in the diagnosis of postmenopausal osteoporosis in women. Width and morphology of inferior cortex of the mandible. *Dentomaxillofac. Radiol.* **1996**, *25*, 263–267. [CrossRef] [PubMed]
14. Ledgerton, D.; Horner, K.; Devlin, H.; Worthington, H. Radiomorphometric indices of the mandible in a British female population. *Dentomaxillofac. Radiol.* **1999**, *28*, 173–181. [CrossRef]
15. White, S.C.; Taguchi, A.; Kao, D.; Wu, S.; Service, S.K.; Yoon, D.; Suei, Y.; Nakamoto, T.; Tanimoto, K. Clinical and panoramic predictors of femur bone mineral density. *Osteoporos. Int.* **2005**, *16*, 339–346. [CrossRef]
16. Yasar, F.; Akgunlu, F. The differences in panoramic mandibular indices and fractal dimension between patients with and without spinal osteoporosis. *Dentomaxillofac. Radiol.* **2006**, *35*, 1–9. [CrossRef]
17. Taguchi, A.; Ohtsuka, M.; Tsuda, M.; Nakamoto, T.; Kodama, I.; Inagaki, K.; Noguchi, T.; Kudo, Y.; Suei, Y.; Tanimoto, K. Risk of vertebral osteoporosis in post-menopausal women with alterations of the mandible. *Dentomaxillofac. Radiol.* **2007**, *36*, 143–148. [CrossRef]
18. Devlin, H.; Karayianni, K.; Mitsea, A.; Jacobs, R.; Lindh, C.; van der Stelt, P.; Marjanovic, E.; Adams, J.; Pavitt, S.; Horner, K. Diagnosing osteoporosis by using dental panoramic radiographs: The OSTEODENT project. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod.* **2007**, *104*, 821–828. [CrossRef]
19. Okabe, S.; Morimoto, Y.; Ansai, T.; Yoshioka, I.; Tanaka, T.; Taguchi, A.; Kito, S.; Wakasugi-Sato, N.; Oda, M.; Kuroiwa, H.; et al. Assessment of the relationship between the mandibular cortex on panoramic radiographs and the risk of bone fracture and vascular disease in 80-year-olds. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod.* **2008**, *106*, 433–442. [CrossRef]
20. Taguchi, A. Triage screening for osteoporosis in dental clinics using panoramic radiographs. *Oral Dis.* **2010**, *16*, 316–327. [CrossRef]
21. Al-Dam, A.; Blake, F.; Atac, A.; Amling, M.; Blessmann, M.; Assaf, A.; Hanken, H.; Smeets, R.; Heiland, M. Mandibular cortical shape index in non-standardised panoramic radiographs for identifying patients with osteoporosis as defined by the German Osteology Organization. *J. Craniomaxillofac. Surg.* **2013**, *41*, e165–e169. [CrossRef] [PubMed]
22. Kavitha, M.S.; Asano, A.; Taguchi, A.; Kurita, T.; Sanada, M. Diagnosis of osteoporosis from dental panoramic radiographs using the support vector machine method in a computer-aided system. *BMC Med. Imaging* **2012**, *12*, 1. [CrossRef] [PubMed]
23. Kavitha, M.S.; Ganesh Kumar, P.; Park, S.Y.; Huh, K.H.; Heo, M.S.; Kurita, T.; Asano, A.; An, S.Y.; Chien, S.I. Automatic detection of osteoporosis based on hybrid genetic swarm fuzzy classifier approaches. *Dentomaxillofac. Radiol.* **2016**, *45*, 20160076. [CrossRef]
24. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.; van Ginneken, B.; Sanchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60–88. [CrossRef] [PubMed]

25. Park, C.; Took, C.C.; Seong, J.K. Machine learning in biomedical engineering. *Biomed. Eng. Lett.* **2018**, *8*, 1–3. [CrossRef] [PubMed]
26. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
27. Baker, B.; Gupta, O.; Naik, N.; Raskar, R. Designing neural network architectures using reinforcement learning. *arXiv* **2016**, arXiv:1611.02167.
28. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? *arXiv* **2014**, arXiv:1411.1792.
29. Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **2009**, *22*, 1345–1359. [CrossRef]
30. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. *arXiv* **2015**, arXiv:1512.03385.
31. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
32. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the 25th International Conference on Neural Information Processing Systems, Lake Tahoe, Nevada, 3–8 December 2012; Curran Associates Inc.: Red Hook, NY, USA, 2012; Volume 1, pp. 1097–1105.
33. Han, Z.; Wei, B.; Zheng, Y.; Yin, Y.; Li, K.; Li, S. Breast cancer multi-classification from histopathological images with structured deep learning model. *Sci. Rep.* **2017**, *7*, 4172. [CrossRef] [PubMed]
34. Christopher, M.; Belghith, A.; Bowd, C.; Proudfoot, J.A.; Goldbaum, M.H.; Weinreb, R.N.; Girkin, C.A.; Liebmann, J.M.; Zangwill, L.M. Performance of Deep Learning Architectures and Transfer Learning for Detecting Glaucomatous Optic Neuropathy in Fundus Photographs. *Sci. Rep.* **2018**, *8*, 16685. [CrossRef] [PubMed]
35. Shin, H.-C.; Roth, H.R.; Gao, M.; Lu, L.; Xu, Z.; Nogues, I.; Yao, J.; Mollura, D.; Summers, R.M. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans. Med. Imaging* **2016**, *35*, 1285–1298. [CrossRef] [PubMed]
36. Ravishankar, H.; Sudhakar, P.; Venkataramani, R.; Thiruvenkadam, S.; Annangi, P.; Babu, N.; Vaidya, V. Understanding the mechanisms of deep transfer learning for medical images. *arXiv* **2017**, arXiv:1704.06040.
37. Kanis, J.A. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: Synopsis of a WHO report. WHO Study Group. *Osteoporos. Int.* **1994**, *4*, 368–381. [CrossRef]
38. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M. Imagenet large scale visual recognition challenge. *Int. J. Compute. Vis.* **2015**, *115*, 211–252. [CrossRef]
39. Stone, M. Cross-validatory choice and assessment of statistical predictions. *J. R. Stat. Soc. Ser. B Methodol.* **1974**, *36*, 111–133. [CrossRef]
40. Chollet, F. Keras: Deep Learning Library for Theano and Tensorflow. 2015, 7, p. T1. Available online: https://keras.io (accessed on 30 January 2020).
41. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv* **2016**, arXiv:1603.04467.
42. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual explanations from deep networks via gradient-based localization. *arXiv* **2016**, arXiv:1610.02391.
43. Sawagashira, T.; Hayashi, T.; Hara, T.; Katsumata, A.; Muramatsu, C.; Zhou, X.; Iida, Y.; Katagi, K.; Fujita, H. An automatic detection method for carotid artery calcifications using top-hat filter on dental panoramic radiographs. *IEICE Trans. Inf. Syst.* **2013**, *96*, 1878–1881. [CrossRef]
44. Lee, J.-S.; Adhikari, S.; Liu, L.; Jeong, H.-G.; Kim, H.; Yoon, S.-J. Osteoporosis detection in panoramic radiographs using a deep convolutional neural network-based computer-assisted diagnosis system: A preliminary study. *Dentomaxillofac. Radiol.* **2019**, *48*, 20170344. [CrossRef] [PubMed]
45. Nogueira, K.; Penatti, O.A.; Dos Santos, J.A. Towards better exploiting convolutional neural networks for remote sensing scene classification. *Pattern Recognit.* **2017**, *61*, 539–556. [CrossRef]



Journal of
Clinical Medicine

MDPI

Article

Effect of an Electronic Medical Record-Based Screening System on a Rapid Response System: 8-Years' Experience of a Single Center Cohort

Se Hee Lee, Chae-Man Lim, Younsuck Koh, Sang-Bum Hong and Jin Won Huh *

Department of Pulmonary and Critical Care Medicine, Asan Medical Center, University of Ulsan College of Medicine, 88, Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Korea; celestia7@gmail.com (S.H.L.); cmlim@amc.seoul.kr (C.-M.L.); yskoh@amc.seoul.kr (Y.K.); sbhong@amc.seoul.kr (S.-B.H.)
* Correspondence: jwhuh@amc.seoul.kr; Tel.: +82-2-3010-3985

Received: 30 December 2019; Accepted: 21 January 2020; Published: 1 February 2020

Abstract: An electronic medical record (EMR)-based screening system has been developed as a trigger system for a rapid response team (RRT) that traditionally used direct calling. We compared event characteristics, intensive care unit (ICU) admission, and 28-day mortality following RRT activation of the two trigger systems. A total of 10,026 events were classified into four groups according to the activation time (i.e., daytime or on-call time) and the triggering type (i.e., calling or screening). Among surgical patients, the ICU admission was lowest for the on-call screening group (26.2%). Compared to the on-call screening group, the on-call calling group and daytime calling group showed higher ICU admission with an odds ratio (OR) of 2.07 (95% CI 1.50–2.84, $p < 0.001$) and OR of 2.68 (95% CI 1.91–3.77, $p < 0.001$), respectively. The 28-day mortality was lowest for the on-call screening group (8.7%). Compared to the on-call screening group, on-call calling (OR 1.88, 95% CI 1.20–2.95, $p = 0.006$) and daytime calling (OR 1.89, 95% CI 1.17–3.05, $p < 0.001$) showed higher 28-day mortality. The EMR-based screening system might be useful in detecting at-risk surgical patients, particularly during on-call time. The clinical usefulness of an EMR-based screening system can vary depending on patients' characteristics.

Keywords: clinical deterioration; early medical intervention; electronic health records; hospital rapid response team; intensive care units; medical records system; computerized

1. Introduction

Rapid response teams (RRT) were widely deployed in the early 2000s to promptly detect deteriorating patients outside critical care and to provide appropriate advanced critical care early on [1]. RRTs are activated by calls by medical staff based on the calling criteria and the clinical concern. Increasing the RRT dose could improve patient outcomes [2,3]. However, previous research indicates that only 30% of at-risk patients who satisfied the calling criteria received critical care from RRTs [1]. In addition, diurnal variation affects RRT activation and clinical outcomes. RRT calls frequently occur during the day. Diurnal variation in RRT utilization influences hospital mortality dependent upon the time of the call [4,5]. Delayed RRT activation occurs more frequently between midnight and 8:00 am and is associated with increased hospital mortality [6]. Infrequent activation during early morning hours is followed by a spike in mortality at 7:00 am [7]. These findings suggest a delay in recognition of at-risk patients and suboptimal RRT utilization by caregivers at night results in poor patient outcomes.

Abundant clinical data and conclusions derived from electronic medical records (EMR) can be utilized to improve not only health care quality but also point-of-care management by detecting clinical deterioration early. The 24-h accessibility of EMR is beneficial in that automatic EMR monitoring by RRTs tends to reinforce the screening of at-risk patients. Currently, vital signs and certain laboratory

data in EMR are used as criteria parameters for detecting deteriorating patients working as an additional limb of RRT [8,9]. However, results of EMR-based RRT systems have been mixed [9–11]. In our hospital, an RRT with dual-triggering afferent limbs, which utilizes both direct calling from bedside doctors or nurses and 24-h based EMR screening criteria, was introduced in 2008. We adopted a single-parameter EMR screening system. Previous research that assessed clinical outcomes in the first two-year period after dual triggering system deployment reported that EMR screening resulted in lower intensive care unit (ICU) admission rates but only surgically ill patients had reduced 28-day mortality rates [12]. This study aimed to analyze the event characteristics and clinical outcomes of RRT activations according to the trigger-type and activation time using the 8-year-period RRT cohort.

2. Methods

2.1. Study Populations

The study protocol was approved by the Institutional Review Board (2016-0857) of Asan Medical Center. Due to the retrospective nature, informed consent was not required, and patients' data were used anonymously. This study was conducted at an academic tertiary care hospital with approximately 2400 adult beds. All adult patients in general wards who received treatment from the RRT were eligible. RRT operated for 24 h a day, 7 days a week during the study period. As the purpose of the study was to compare clinical characteristics and effectiveness of two triggering systems in early detection and management of at-risk patients, RRT events which were categorized as cardiopulmonary cerebral resuscitation (CPCR), post-CPCR care, educational purpose, procedure assistance, and counseling for end-of-life were excluded. Patients who requested to be listed as "do not resuscitate" (DNR) were also excluded. If the patient had more than one event in the same admission period, only the first event was included for analysis to eliminate redundancy.

Based on the duty hours of training residents, daytime for weekdays was defined as 7:00 am to 5:59 pm while on-call time was defined as 6:00 pm to 6:59 am on the following day. Daytime for weekends or holidays was defined as from 7:00 am to 11:59 am while on-call time was defined as from 12:00 pm to 6:59 am on the following day. Nursing staffs work in three shifts, day shift (6:30 am to 2:30 pm), evening shift (2:30 pm to 10:30 pm), and night shift (10:30 pm to 6:30 am on the following day), regardless of weekend or weekdays. We divided the type of activation events into four groups: calling vs. screening, based on trigger type, and daytime vs. on-call time, based on activation time.

At this hospital, the RRT not only provides advanced critical care but is also actively involved in monitoring and assessing at-risk patients throughout the day. Direct calling from ward physicians and nurses activates the RRT. Additionally, the EMR-based screening system is utilized, which automatically activates the RRT when the pre-defined criteria based on the vital signs and laboratory measurements of the patients' medical records are met. Details of the criteria are shown in Table S1.

2.2. Study Variables and Outcomes

Data were routinely collected for patients' demographics, illness-type (medical or surgical), RRT activation time and date (weekdays or weekend), trigger parameter for activation, therapeutic intervention during the event (e.g., intubation, ventilator, high flow nasal cannular (HFNC), bilevel positive airway pressure (BiPAP), advanced cardiovascular life support (ACLS), etc.), the outcome of the RRT intervention (e.g., ICU transfer vs. ward stay), and the 28-day mortality following the event. If the alarm was triggered by both calling and screening, the first trigger was recorded. Patients' vital signs (e.g., systolic/diastolic blood pressure (BP), pulse rate (PR), respiratory rate (RR), body temperature (BT), and mental status) at the time of the event were also collected to risk-stratify patients. We calculated a modified early warning score (MEWS), which was validated for both medical in-patients and surgical in-patients [13,14] and used this score for adjustment. MEWS is the sum of scores for five parameters: systolic BP, PR, RR, and mental status. Each parameter and detailed pre-assigned score are described in Table S2.

The number of RRT events per year and the number of RRT activations per clock hour were analyzed to identify the RRT activation pattern. RRT events per each clock hour were classified as per screening and calling which were further divided into doctor-calling, and nurse-calling. All events were categorized into four different groups: daytime calling, daytime screening, on-call calling, and on-call screening. The primary outcome considered was ICU admission after RRT activation. Twenty-eight day mortality following RRT activation was also assessed for the four groups.

2.3. Statistical Methods

Annual RRT activations and the number of RRT events for each clock hour are presented graphically. Differences between two groups (i.e., calling vs. screening in daytime, and calling vs. screening in on-call time) were tested using a Chi-square test for categorical variables, and the independent t-test for continuous variables. ICU admission and 28-day mortality were compared between the on-call screening group and the other three individual groups and presented as an odds ratio (OR) with a 95% confidence interval using a multivariate logistic regression model adjusting for age, sex, MEWS, weekend, and activation coding after univariate analysis. The risk factors for ICU admission and 28-day mortality were identified following univariate logistic analysis and statistically significant variables were further applied for the multivariate logistic regression. All statistical analyses were performed using SPSS software (version 24.0; IBM Corp., Armonk, NY).

3. Results

From 1 January 2009 to 31 December 2016, 15,641 RRT events were identified (Figure 1); of these, 10,026 events were included for analysis. All events were classified according to activation time and trigger type. A total of 6293 events occurred during on-call time and 54.3% (n = 3419) were activated by screening rather than calling. Figure 2 represents the number of RRT activation events per year over the 8-year period. The number of RRT triggers by calling did not vary significantly between each year but activation by screening increased steadily from 2009. This increase was more prominent during on-call time.

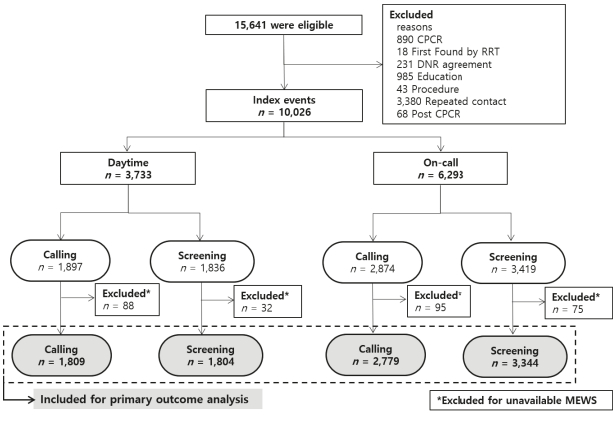


Figure 1. Schematic flow chart of the study Given that our rapid response team (RRT) performs a multifunctional role, only events related to early detection and management of at-risk patients were included for the study. Among 15,641 eligible events, 10,026 events were analyzed to describe the pattern of RRT activations. For clinical outcome analysis, 9736 events were included after excluding 290 events due to unavailable MEWS. Cardiopulmonary cerebral resuscitation (CPCR); Do not resuscitate (DNR); Rapid response team (RRT); Modified early weaning score (MEWS).

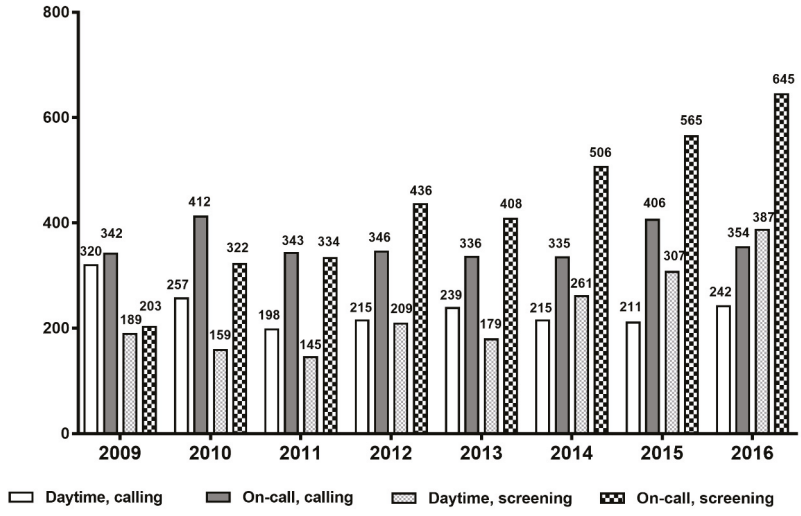


Figure 2. The number of RRT events per year since 2009. Overall RRT events increased from 1055 in 2009 to 1627 in 2016. The total number of RRT activations by screening in 2016 was 2.63-fold higher than that in 2009. Data are presented as number of events.

The number of events for each clock hour are illustrated in Figure 3. In total, 4771 events (47.6%) were call-triggered (doctor calling and nurse calling). The number of activations by nurse calling was relatively stable in each clock hour compared with the number of doctor calling. RRT contacts were most frequent from midnight to 00:59 am (n = 952, 9.5%). The proportions of activation by screening were higher during on-call time, particularly from midnight to 0:59 at which time the vital check is conducted by night duty nurses.

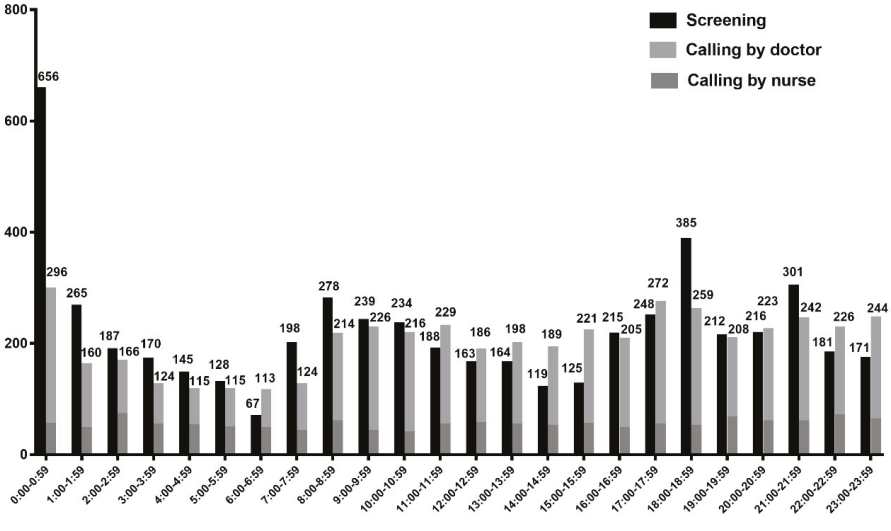


Figure 3. The RRT frequency according to each clock hour. Among 10,026 events, 4771 (47.6%) were triggered by calling and 5255 (52.4%) were triggered by screening. RRT contacts are most frequent at midnight to 00:59 am (n = 952, 9.5%). The total frequency was higher in order of 18:00 pm, 21:00 pm, and 8:00 am. Data are presented as number of events.

Patients' baseline characteristics, illness type, and MEWS are presented in Table 1. Approximately half of the patients in the screening group had solid malignancy (daytime 50.3% vs. 35.3%, $p < 0.001$; on-call time 50.2% vs. 41.5%, $p < 0.001$). Moreover, a greater number of patients had hematologic malignancy in the screening group than in the calling group (daytime 19.3% vs. 13.4%, $p < 0.001$; on-call time 16.7% vs. 11.6%, $p < 0.001$). In contrast, the proportion of patients with chronic lung disease, cardiovascular disease, or neurologic disease was higher in the calling group. Among surgical patients, a higher number of at-risk patients were activated by calling rather than screening (daytime 18.9% vs. 10.7%, $p < 0.001$: on-call time 19.6% vs. 12.9%, $p < 0.001$). MEWS was available in 9,736 events. MEWS was significantly higher in the calling group than in the screening group (daytime 4.54 vs. 4.30, $p = 0.031$; on-call 4.57 vs. 4.37, $p = 0.0017$). Activation coding and type of intervention are presented in Table S3.

Table 1. Baseline characteristics of included events.

	Daytime		On-Call	
	Calling	**Screening**	**Calling**	**Screening**
	***N* = 1897**	***N* = 1836**	***N* = 2874**	***N* = 3419**
Age	64 (52–73)	64 (53–72)	64 (53–73)	64 (53–72)
Sex				
Male—No. (%)	1165 (61.7)	1120 (61.0)	1768 (61.5)	2097 (61.3)
Underlying disease				
Solid malignancy	669 (35.3)	924 (50.3) ‡	1194 (41.5)	1717 (50.2) ‡
Hematologic malignancy	254 (13.4)	355 (19.3) ‡	334 (11.6)	571 (16.7) ‡
Chronic lung disease	277 (14.6) *	224 (12.2)	405 (14.1) *	375 (11.0)
Cardiovascular disease	839 (44.2) †	720 (39.2)	1327 (46.2) *	1483 (43.4)
Chronic liver disease	273 (14.4)	267 (14.5)	424 (14.8)	479 (14.0)
Gastrointestinal disease	7 (0.4)	6 (0.3)	16 (0.6)	14 (0.4)
Neurologic disease	262 (13.8) ‡	166 (9.0)	402 (14.0) ‡	322 (9.4)
Chronic kidney disease	158 (8.3) ‡	102 (5.6)	217 (7.6) *	203 (5.9)
Thyroid disease	95 (5.0) *	61 (3.3)	113 (3.9)	140 (4.1)
Diabetes mellitus	440 (23.2)	429 (23.4)	703 (24.5)	825 (24.1)
Solid organ transplant	70 (3.7)	58 (3.2)	86 (3.0)	93 (2.7)
Illness type				
Medical	1450 (78.7)	1627 (88.6) ‡	2251 (79.9)	2987 (87.4) ‡
Surgical	392 (21.3) ‡	209 (11.4)	567 (20.1) ‡	432 (12.6)
MEWS	4.54 ± 2.23 ‡	4.30 ± 2.02	4.57 ± 2.24 ‡	4.37 ± 2.01
Weekend	330 (17.4)	377 (20.5) *	1103 (38.4)	1347 (40.2)

Among continuous variables, age is presented as median (interquartile range) and MEWS are presented as mean ± SD. Categorical variables are presented as No. (%). MEWS was available in 9736 patients. * p-value < 0.05, † p-value < 0.01, ‡ p-value < 0.001. Chi-square test was done for the comparison between daytime calling and daytime screening. The same analytic technique was used for the comparison between on-call calling and on-call screening. MEWS = modified early weaning score.

The overall ICU admission was 28.9% and 28-day mortality was 30% among 9736 patients. As more patients in the screening group had a malignancy (Table 1), a subgroup analysis was conducted to compare the clinical outcomes between patients with cancer and those without cancer (Table 2). Among patients with an underlying malignancy, the overall ICU admission rate was 21.5% and the on-call screening group displayed the lowest ICU admission (14.9%). The overall 28-day mortality was 40.6% and mortality was lower for the on-call calling group compared to the on-call screening group (OR 0.84, 95% CI 0.72–0.98). Among patients without cancer, overall ICU admission and 28-day mortality were 36.1% and 21.4%, respectively. ICU admission was lowest for the on-call screening group (22.3%) as was 28-day mortality (18.6%). Similar to patients with cancer, the daytime screening group had higher mortality compared to on-call screening group (OR 1.48, 95% CI 1.14–1.93).

Table 2. Clinical outcomes among medical patients without cancer and with cancer.

		With Cancer (N = 4980)			Without Cancer (N = 3256)		
		OR	**95% CI**	***p*-Value**	**OR**	**95% CI**	***p*-Value**
ICU admission	On-call, screening	1			1		
	Daytime, screening	1.01	0.81–1.26	0.92	1.26	0.97–1.64	0.086
	On-call, calling	2.30	1.91–2.77	<0.001	2.42	1.98–2.97	<0.001
	Daytime, calling	3.85	3.11–4.77	<0.001	3.65	2.93–4.54	<0.001
28-day mortality	On-call, screening	1			1		
	Daytime, screening	1.16	0.99–1.35	0.063	1.48	1.14–1.93	0.004
	On-call, calling	0.84	0.72–0.98	0.026	1.10	0.88–1.38	0.417
	Daytime, calling	0.87	0.72–1.05	0.136	1.17	0.92–1.50	0.204

Data are presented as odds ratio (OR) with 95% confidence interval (CI). ICU admission and 28-day mortality were analyzed using a multivariate logistic regression model adjusting for age, sex, MEWS, weekend, and activation code. MEWS, weekend, and activation code were variables finally selected for the regression model for ICU admission in cancer patients. For 28-day mortality in cancer patients, sex, MEWS, weekend, and activation code were adopted variables for the regression model. Among patients without cancer, age, MEWS, and activation code were variables adopted for ICU admission and 28-day mortality. For patients with cancer: overall (n = 4980); on-call screening (n = 1971); daytime screening (n = 1131); on-call calling (n = 1188); daytime calling (n = 690). For patients without cancer: overall (n = 3256); On-call screening (n = 961); Daytime screening (n = 472); On-call calling (n = 1063); Daytime calling (n = 760).

Among patients with surgical illnesses (Table 3) overall ICU admission and 28-day mortality were 37.6% and 13.4%, respectively. The on-call screening group was significantly associated with lower ICU admission (26.2%); daytime screening had an ICU admission of 28.4% (OR 1.06, 95% CI 0.70–1.59, p = 0.794), on-call calling 42.4% (OR 2.07, 95% CI 1.50–2.84, $p < 0.001$), and daytime calling 46.1% (OR 2.68, 95% CI 1.91–3.77, $p < 0.001$). The on-call screening group was also associated with lower 28-day mortality (8.7%); the 28-day mortality for daytime screening was 12.9% (OR 1.44, 95% CI 0.82–2.51, p = 0.203), on-call calling 15.5% (OR 1.88, 95% CI 1.20–2.95, p = 0.006), and daytime calling 15.9% (OR 1.89, 95% CI 1.17–3.05, p = 0.0009).

Table 3. Clinical outcomes among patients with surgical illness.

		OR	**95% CI**	***p*-Value**
ICU admission	On-call, screening	1		
	Daytime, screening	1.06	0.70–1.59	0.794
	On-call, calling	2.07	1.50–2.84	<0.0001
	Daytime, calling	2.68	1.91–3.77	<0.0001
28-day mortality	On-call, screening	1		
	Daytime, screening	1.44	0.82–2.51	0.203
	On-call, calling	1.88	1.20–2.95	0.006
	Daytime, calling	1.89	1.17–3.05	0.009

Data are presented as odds ratio with 95% confidence interval (CI). ICU admission and 28-day mortality were analyzed using a multivariate logistic regression model adjusting for age, sex, MEWS, weekend, and activation code. Sex, MEWS, weekend, and activation code were variables finally selected for the regression model for ICU admission. For 28-day mortality, sex, MEWS, and activation code were adopted variables for the regression model. Overall (n = 1500); on-call screening (n = 412); daytime screening (n = 201); on-call calling (n = 528); daytime calling (n = 359).

The risk factors associated with ICU transfer and 28-day mortality following RRT activation are shown in Table S4. On-call time patients were less likely to be transferred to the ICU (OR 0.76, 95% CI: 0.68–0.84, $p < 0.001$) and had a lower mortality rate than daytime patients (OR 0.85, 95% CI: 0.78–0.94, $p < 0.001$). Compared to calling, the RRT activation by screening was associated with a lower ICU transfer rate, (OR 0.51, 95% CI: 0.46–0.57, $p < 0.001$) but a higher 28-day mortality rate (OR 1.19, 95% CI: 1.08–1.33).

4. Discussion

We evaluated the effects of an EMR-based screening system using the RRT cohort over an 8-year period following the employment of a dual triggering system. In addition to the calling system, the EMR-based screening system was implemented to aid in increasing the detection sensitivity of at-risk patients. Between 1 January 2009 and 31 December 2016, a total of 10,026 events were included. The number of RRT activations by screening system continually increased over the course of data collection. The higher proportion of screening group (especially on-call time) compared to the other study groups can be explained by two factors: (1) the experiences gained by the RRT over the study years might have led to an increase in the sensitivity of the screening system in detecting at-risk patients; (2) a small number of doctors on duty compared to daytime might result in a decrease in on-call calling, thus eventually increasing the burden of RRT work during on-call time.

Among the patients without cancer and surgical patients, the on-call screening group had lower ICU admission and lower 28-day mortality than other groups, possibly due to early detection using EMR screening during on-call time. However, this positive effect of the screening system was not observed among at-risk patients with cancer. Although more patients in the on-call calling group transferred to the ICU and had higher MEWS than the on-call screening group, 28-day mortality was significantly lower in the on-call calling group. The higher ICU admission rate in the calling group can be explained by the fact that activation by calling is (1) more likely to be associated with acute medical events and (2) more likely to reflect greater motivation on behalf of attending physicians to treat the patient. Alternatively, the daytime screening group had lower MEWS and lower ICU admission, but higher 28-day mortality. This result suggests that various factors affect mortality in medically ill at-risk patients. DNR agreement following RRT activation or consideration for end-of-life care might be closely related to 28-day mortality.

In many reports, the dose-response effect of the RRT was well described [15–17]. Increasing RRT dose was associated with dose-related reduction of cardiac arrest and cardiac arrests were most common overnight when RRT dose was the lowest. Because the trigger threshold by traditional calling criteria may vary depending on the experience or concern of ward physicians and nurses, the achievement of optimal RRT dose is important. Real-time monitoring by experienced RRT could increase the sensitivity of detecting clinical deterioration and improve clinical outcome. As our results indicate, triggering frequency itself is largely dependent on the interval of vital sign measurements. As vital sign check-ups at the ward are typically recorded by nurses at intervals of 8-h or 4-h, the activation frequency was higher during the regular vital sign check-up hours. Therefore, unless clinicians order frequent vital sign check-ups or laboratory tests for possible at-risk patients, the possibility of missing indicators of deteriorating patients will persist. Employment of automatic continuous monitoring could overcome this limitation [18].

We used single parameters including laboratory data, such as lactate levels and arterial blood gas analysis, as the triggers for the EMR-based system. Previous research indicates various degrees of sensitivity and the accuracy of multiple aggregate weighted scoring systems (AWSS) for predicting ICU transfer, cardiac arrest, and mortality [19,20]. However, the current AWSS depends mainly on vital signs, without assessing other characteristics of at-risk patients. Therefore, the development of a modified scoring system which includes vital signs, laboratory data, and characteristics of the patient population is necessary to increase the sensitivity and specificity of the screening system.

There are several limitations in this study. First, this is a single-center study and data were analyzed retrospectively. As our center is an academic tertiary hospital, the patients' disease severity is generally higher than in non-tertiary hospitals; thus, our results may not generalize well to non-teaching hospitals or hospitals with different patient populations. Nevertheless, the use of a screening system during on-call time seems to be associated with an improved 28-day mortality rate, particularly among surgically ill patients and patients without malignancy. As our hospital adopted a dual-triggering system at the time of the launch of RRT, for ethical and patient safety concerns, a prospective randomized trial comparing calling system and screening system based on the duty hours was not

possible. A prospective multicenter study is required to evaluate the efficacy of the screening system more accurately. Second, the proportion of patients who agreed on a plan for end-of-life care following RRT activation was not considered. There are other factors affecting primary outcomes aside from the management of RRT, such as end-of-life care, spontaneous decisions by attending physicians, will of family caregivers, and the availability of ICU beds at the time of the RRT visit. Therefore, these results should be interpreted with careful consideration of multiple clinical factors, not by the RRT intervention alone.

5. Conclusions

Deployment of an EMR-based screening system offers additional improvement in detecting and managing at-risk patients, particularly during on-call time. However, the clinical effectiveness of this system can vary depending on patients' characteristics. The deployment of a modified screening system reflecting the physiologic parameters, laboratory measurements, and underlying diseases of the patient population at each hospital would maximize the beneficial role of the RRT in point of care management.

Supplementary Materials: The following are available online at http://www.mdpi.com/2077-0383/9/2/383/s1, Table S1: Criteria for medical alert team activation, Table S2: Modified early weaning score (MEWS) Table S3: Baseline patients and event characteristics, Table S4: The multivariate analysis for the risk factors associated with ICU admission and 28-day mortality after RRT activation.

Author Contributions: S.H.L. contributed to analyzing data and writing of the manuscript. C.-M.L., Y.K., and S.-B.H. contributed in planning the study, clinical work, and approval of the submitted article. J.W.H. conceived the study; contributed in clinical work, integrity, and accuracy of data; preparation and approval of the submitted article. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI15C1106).

Acknowledgments: The authors thank the nurses of rapid response team for their contribution to this work.

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

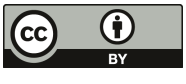
Abbreviations

ACLS: Advanced cardiovascular life support; AWSS: Aggregate weighted scoring systems; BiPAP: Bilevel positive airway pressure; CI: Confidence interval; CPCR: Cardiopulmonary cerebral resuscitation; DNR: Do not resuscitate; EMR: electronic medical record; HFNC: High flow nasal cannular; ICU: Intensive care unit; MEWS: Modified early warning score; RRT: Rapid response team

References

1. Hillman, K.; Chen, J.; Cretikos, M.; Bellomo, R.; Brown, D.; Doig, G.; Finfer, S.; Flabouris, A. Introduction of the medical emergency team (MET) system: A cluster-randomised controlled trial. *Lancet* **2005**, *365*, 2091–2097. [CrossRef] [PubMed]
2. Maharaj, R.; Raffaele, I.; Wendon, J. Rapid response systems: A systematic review and meta-analysis. *Crit. Care* **2015**, *19*, 254. [CrossRef] [PubMed]
3. Jones, D.; Bellomo, R.; DeVita, M.A. Effectiveness of the medical emergency team: The importance of dose. *Crit. Care* **2009**, *13*, 313. [CrossRef] [PubMed]
4. Galhotra, S.; DeVita, M.A.; Simmons, R.L.; Schmid, A. Impact of patient monitoring on the diurnal pattern of medical emergency team activation. *Crit. Care Med.* **2006**, *34*, 1700–1706. [CrossRef] [PubMed]
5. Sundararajan, K.; Flabouris, A.; Thompson, C. Diurnal variation in the performance of rapid response systems: The role of critical care services-a review article. *J. Intensive Care* **2016**, *4*, 15. [CrossRef] [PubMed]
6. Barwise, A.; Thongprayoon, C.; Gajic, O.; Jensen, J.; Herasevich, V.; Pickering, B.W. Delayed rapid response team activation is associated with increased hospital mortality, morbidity, and length of stay in a tertiary care institution. *Crit. Care Med.* **2016**, *44*, 54–63. [CrossRef] [PubMed]
7. Churpek, M.M.; Edelson, D.P.; Lee, J.Y.; Carey, K.; Snyder, A. Association between survival and time of day for rapid response team calls in a national registry. *Crit. Care Med.* **2017**, *45*, 1677–1682. [CrossRef] [PubMed]

8. Wong Lama, K.M.; DeVita, M.A. Trigger criteria: Big data. *Crit. Care Clin.* **2018**, *34*, 199–207. [CrossRef] [PubMed]
9. Kang, M.A.; Churpek, M.M.; Zadravecz, F.J.; Adhikari, R.; Twu, N.M.; Edelson, D.P. Real-time risk prediction on the wards: A feasibility study. *Crit. Care Med.* **2016**, *44*, 1468–1473. [CrossRef] [PubMed]
10. Green, M.; Lander, H.; Snyder, A.; Hudson, P.; Churpek, M.; Edelson, D. Comparison of the between the flags calling criteria to the MEWS, NEWS and the electronic cardiac arrest risk triage (eCART) score for the identification of deteriorating ward patients. *Resuscitation* **2018**, *123*, 86–91. [CrossRef] [PubMed]
11. Kollef, M.H.; Chen, Y.; Heard, K.; LaRossa, G.N.; Lu, C.; Martin, N.R.; Martin, N.; Micek, S.T.; Bailey, T. A randomized trial of real-time automated clinical deterioration alerts sent to a rapid response team. *J. Hosp. Med.* **2014**, *9*, 424–429. [CrossRef] [PubMed]
12. Huh, J.W.; Lim, C.M.; Koh, Y.; Lee, J.; Jung, Y.K.; Seo, H.S.; Hong, S.B. Activation of a medical emergency team using an electronic medical recording-based screening system. *Crit. Care Med.* **2014**, *42*, 801–808. [CrossRef] [PubMed]
13. Subbe, C.P.; Kruger, M.; Rutherford, P.; Gemmel, L. Validation of a modified Early Warning Score in medical admissions. *Q J. Med.* **2001**, *94*, 521–526. [CrossRef] [PubMed]
14. Gardner-Thorpe, J.; Love, N.; Wrightson, J.; Walsh, S.; Keeling, N. The value of Modified Early Warning Score (MEWS) in surgical in-patients: A prospective observational study. *Ann. R Coll. Surg. Engl.* **2006**, *88*, 571–575. [CrossRef] [PubMed]
15. Foraida, M.I.; DeVita, M.A.; Braithwaite, R.S.; Stuart, S.A.; Brooks, M.M.; Simmons, R.L. Improving the utilization of medical crisis teams (Condition C) at an urban tertiary care hospital. *J. Crit. Care* **2003**, *18*, 87–94. [CrossRef] [PubMed]
16. Jones, D.; Bellomo, R.; Bates, S.; Warrillow, S.; Goldsmith, D.; Hart, G.; Opdam, H. Patient monitoring and the timing of cardiac arrests and medical emergency team calls in a teaching hospital. *Intensive Care Med.* **2006**, *32*, 1352–1356. [CrossRef] [PubMed]
17. Jones, D.; Bellomo, R.; Bates, S.; Warrillow, S.; Goldsmith, D.; Hart, G.; Opdam, H.; Gutteridge, G. Long term effect of a medical emergency team on cardiac arrests in a teaching hospital. *Crit. Care* **2005**, *9*, R808–R815. [CrossRef] [PubMed]
18. Bellomo, R.; Ackerman, M.; Bailey, M.; Beale, R.; Clancy, G.; Danesh, V.; Hvarfner, A.; Jimenez, E.; Konrad, D.; Lecardo, M.; et al. A controlled trial of electronic automated advisory vital signs monitoring in general hospital wards. *Crit. Care Med.* **2012**, *40*, 2349–2361. [CrossRef] [PubMed]
19. Jansen, J.O.; Cuthbertson, B.H. Detecting critical illness outside the ICU: The role of track and trigger systems. *Curr. Opin. Crit. Care* **2010**, *16*, 184–190. [CrossRef] [PubMed]
20. Churpek, M.M.; Yuen, T.C.; Edelson, D.P. Risk stratification of hospitalized patients on the wards. *Chest* **2013**, *143*, 1758–1765. [CrossRef] [PubMed]



Journal of
Clinical Medicine

MDPI

Article

Ensemble Learning to Improve the Prediction of Fetal Macrosomia and Large-for-Gestational Age

Shangyuan Ye [1], Hui Zhang [2], Fuyan Shi [3], Jing Guo [4], Suzhen Wang [3,*] and Bo Zhang [5,*]

1 Department of Population Medicine, Harvard Pilgrim Health Care Institute and Harvard Medical School, Boston, MA 02115, USA; shangyuan_ye@harvardpilgrim.org
2 Division of Biostatistics, Department of Prevention Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA.; hzhang@northwestern.edu
3 School of Public Health and Management, Weifang Medical University, Weifang, Shandong 261053, China; shifuyan@wfmc.edu.cn
4 School of Public Health, Peking University, Beijing 100191, China; jing624218@bjmu.edu.cn
5 Department of Neurology and ICCTR Biostatistics and Research Design Center, Boston Children's Hospital and Harvard Medical School, Boston, MA 02115, USA
* Correspondence: wangsz@wfmc.edu.cn (S.W.); bo.zhang@childrens.harvard.edu (B.Z.)

Received: 28 December 2019; Accepted: 28 January 2020; Published: 31 January 2020

Abstract: Background: The objective of this study was to investigate the use of ensemble methods to improve the prediction of fetal macrosomia and large for gestational age from prenatal ultrasound imaging measurements. Methods: We evaluated and compared the prediction accuracies of nonlinear and quadratic mixed-effects models coupled with 26 different empirical formulas for estimating fetal weights in predicting large fetuses at birth. The data for the investigation were taken from the Successive Small-for-Gestational-Age-Births study. Ensemble methods, a class of machine learning techniques, were used to improve the prediction accuracies by combining the individual models and empirical formulas. Results: The prediction accuracy of individual statistical models and empirical formulas varied considerably in predicting macrosomia but varied less in predicting large for gestational age. Two ensemble methods, voting and stacking, with model selection, can combine the strengths of individual models and formulas and can improve the prediction accuracy. Conclusions: Ensemble learning can improve the prediction of fetal macrosomia and large for gestational age and have the potential to assist obstetricians in clinical decisions.

Keywords: macrosomia; large for gestational age; machine learning; ensemble methods; prediction; sensitivity; specificity

1. Introduction

Excessive fetal growth poses risks to maternal and infant well-being [1]. The term "macrosomia" is used to describe the condition of a fetus with a birth weight of more than 4000 g, regardless of gestational age [1,2]. Macrosomia is sometimes confused with "large for gestational age" (LGA), which describes an infant with a 90th percentile or higher birthweight for gestational age [1]. It has been shown that the infants with either macrosomia or LGA pose a large risk of perinatal morbidity and mortality to their mothers [2,3]. Therefore, accurately predicting and diagnosing these conditions has been a major goal of obstetric investigators, for purposes of conducting early intervention or targeted perinatal medical care to reduce the risks.

Measurements taken from prenatal ultrasound imaging are the primary quantitative resources for predicting birth weights and diagnosing macrosomia or LGA. Zhang et al. [4] took the empirical formula given by Hadlock et al. [5] for estimating fetal weights and implemented a joint mixed-effects model to predict macrosomia and LGA. This procedure of predicting macrosomia or LGA from

prenatal ultrasound measurements was a two-step supervised learning process. In the first step, an empirical formula was chosen to derive the estimated fetal weights (EFWs) from sonographic ultrasound measurements at each of the gestational time points when the ultrasound measurements were taken and recorded. In the second step, a joint mixed-effects model with latent subject-specific random effects was fitted. One component of the joint model was a quadratic mixed-effects model to derive the predicted birth weights (PBWs). Another component was a probit mixed-effects model, from which the classification of macrosomia or LGA was determined from the PBWs by comparing the PBWs with a pre-specified threshold [6]. However, previous literature has noted that there were 26 candidate empirical formulas for estimating fetal weights [7–9]. In addition, some literature also argued that nonlinear mixed-effects models should be more appropriate in modeling growth curves than linear or quadratic mixed-effects models [10]. Selecting a particular EFW empirical formula and a quadratic mixed-effects model, as in Zhang et al. [4], increases the uncertainties in predicting macrosomia or LGA, because these empirical formulas and the statistical models function diversely.

In this article, we investigate the use of ensemble methods to aggregate prediction results given by different EFW empirical formulas and the statistical models. The goal of this practice is to improve the prediction in macrosomia or LGA. With ensemble methods, it is not required to select any specific statistical model or empirical formulas. Instead, the prediction capability of each combination of the empirical formulas and statistical models is combined and aggregated to generate a learning procedure that gives the best prediction performance.

2. Methods

2.1. Data

The Successive Small-for-Gestational-Age Births study (SGA study) was funded by the National Institute of Child Health and Human Development (NICHD), in the National Institutes of Health of the USA in 1983. The study was conducted concurrently by the University of Bergen in Norway, the University of Uppsala in Sweden, and the University of Alabama, USA, from 1984 to 1985 [11]. The initial goal of the SGA study was to characterize the different types of intra-uterine growth restriction and to assess the associated risk factors.

To demonstrate the strengths of ensemble methods in predicting fetal macrosomia or LGA, we took the Scandinavian data in the SGA study that were collected in Norway and Sweden. The Scandinavia SGA data were collected from January 1st, 1986, through March 31st, 1988 from nulliparous (parity 1) and primiparous (parity 2) Caucasian pregnant women prior to the 20th gestational week who had a singleton pregnancy and spoke one of the Scandinavian languages. A total of 6354 women were recruited to the study, and 632 of them were excluded from the study, according to the exclusion criteria (n = 432) or due to absence of first prenatal visit (n = 200). The remaining 5722 patients were split into three subgroups. A random sample of 561 patients was first selected. Then, a "high-risk" group of 1384 patients was identified out of the random sample of 561 patients on the basis of five small-for-gestational-age (SGA) risk factors: giving birth to an infant with birthweight below 2750 g in the past, maternal cigarette smoking at conception, pre-pregnancy weight lower than 50 kg, a previous perinatal death, and the presence of chronic maternal disease (chronic renal disease, hypertension, or heart disease). The remaining 3777 patients were considered as the "low-risk" group. The patients from the random sample and "high-risk" group were eligible to participate in a detailed follow-up study, during which the women were examined at approximately 17th, 25th, 33rd, and 37th weeks of gestation. For each visit, ultrasound examination was performed, and demographic and medical information were collected. At the end of the Scandinavia SGA study, only 1945 were able to complete the follow-up study.

In our study, we restricted the study data to the 1115 women in the Scandinavia SGA study who had all four ultrasound examinations and complete covariate information (maternal age, pre-pregnancy body weight and height, previous diseases history, and smoking history) in the Scandinavia SGA

study. In the ultrasound examination records in the Scandinavia SGA study, there were three fetal measurements: biparietal diameter (BPD), middle abdominal diameter (MAD), and femur length (FL).

2.2. An Ensemble Learning Procedure of Predicting Macrosomia and LGA

Here, we developed a four-step ensemble learning procedure to predict macrosomia or LGA from prenatal ultrasound measurements. The output of the learning procedure is the binary classification of either macrosomia or LGA. The prediction of macrosomia and LGA is conducted based upon a set of input features. The primary input features are the sonographic measurements BPDs, MADs, and FLs collected from 17th, 25th, 33rd, and 37th weeks of gestation, as well as gestational age at delivery, in the Scandinavia SGA study. Other features also include maternal age, pre-pregnancy body mass index, parity, smoking status, existing diabetes, and gestational diabetes. Figure 1 shows a diagram that delineates the four steps in the ensemble learning procedure, to predict macrosomia or LGA. Step 1 is to take the sonographic measurements for each of the gestational weeks and each of the empirical formulas in Table A2, to obtain EFWs at each gestational time point. Step 2 is to fit either a nonlinear mixed-effects model or a quadratic mixed-effects model to predict the corresponding PBWs at birth from the EFWs obtained in Step 1. Step 3 is to use the PBWs to derive the classification of macrosomia or LGA by using a specified threshold. Step 4 is the ensemble learning step, in which prediction output from various empirical formulas and the nonlinear and quadratic mixed-effects models are combined to generate the ensemble learning prediction results.

2.3. Estimated Fetal Weights with 26 Empirical Formulas

In the ultrasound examination records in the Scandinavia SGA study, there were three fetal measurements BPD, MAD, and FL. Melamed et al. [7] summarized 26 different empirical sonographic formulas (see Appendix A Table A1) that can be taken to estimate fetal weights from sonographic ultrasound measures. In our ensemble learning procedure, we considered all the 26 models in our analysis. Abdominal circumference (AC) in the empirical sonographic formulas can be calculated by $3.1416 \times$ MAD, and head circumference (HC) can be derived by the formula introduced in [12].

2.4. Mixed-Effects Models for Predicting PBWs and Deriving the Classification of Macrosomia or LGA

We built a nonlinear three-parameter mixed-effects logistic model with latent random effects to predict PBWs and derive the classification of macrosomia or LGA [10,13]. The three-parameter mixed-effects logistic model for the ith fetus ($i = 1, \cdots, n$ indexing the study subjects) at gestational time t_{ij} ($j = 1, 2, 3, 4$ indexing the time when the gestational ultrasound measurements were recorded, $j = 5$ indexing the time of birth) is as follows:

$$y_{ij} = \frac{\phi_{1i}}{1 + \exp\left[-\left(t_{ij} - \phi_{2i}\right)/\phi_{3i}\right]} + \epsilon_{ij},$$

where y_{ij} is the EFW obtained from one of the 26 formulas in Table A1 at gestational time t_{ij}, $j = 1, 2, 3,$ and 4, and y_{ij} is the birth weight when $j = 5$, $\phi_i = (\phi_{1i}, \phi_{2i}, \phi_{3i})$ are model parameters in which ϕ_{1i} indicates amplitude, ϕ_{2i} indicates the smoothness, ϕ_{3i} indicates stretch, and ϵ_{ij} are within-subject random errors. For the nonlinear mixed-effects model, each parameter ϕ_{ki} ($k = 1, 2, 3$ indexing the parameter), can be further modeled by a linear representation:

$$\phi_{ki} = X'_{ij}\beta_k + b_{ki},$$

where X_{ij} is a vector of fixed-effects the covariates of maternal age, pre-pregnancy body mass index, previous disease history, including diabetes, cardiac disease, high blood pressure, renal disorders, and other diseases, and smoking during pregnancy; and β_k denotes the corresponding regression parameters. The first element in X_{ij} equals 1 for the interception. In our learning procedure, the covariates were

included in the linear term $\phi_{1i} = X'_{ij}\beta_1 + b_{1i}$, and ϕ_{2i} and ϕ_{3i} were specified as $\phi_{2i} = \beta_2 + b_{2i}$ and $\phi_{3i} = \beta_3 + b_{3i}$. We also considered the nonlinear three-parameter mixed-effects logistic model without any covariates in ϕ_{1i} such that $\phi_{1i} = \beta_1 + b_{1i}$. The random effects $b_i = (b_{1i}, b_{2i}, b_{3i})'$ represent the latent individual variations that are not explained by the covariates. We assumed the random effects, b_i, independently follow a multivariate normal distribution, with a vector of mean 0 and a variance-covariance matrix Σ, where Σ was assumed to be positive, definite, and unstructured. Further, we assumed the following heteroscedastic model [10] for the within-subject random errors ϵ_{ij}:

$$\epsilon_{ij} \sim N\left(0, \left|\frac{\phi_{1i}}{1+\exp\left[-\left(t_{ij}-\phi_{2i}\right)/\phi_{3i}\right]}\right|^{2\delta}\right).$$

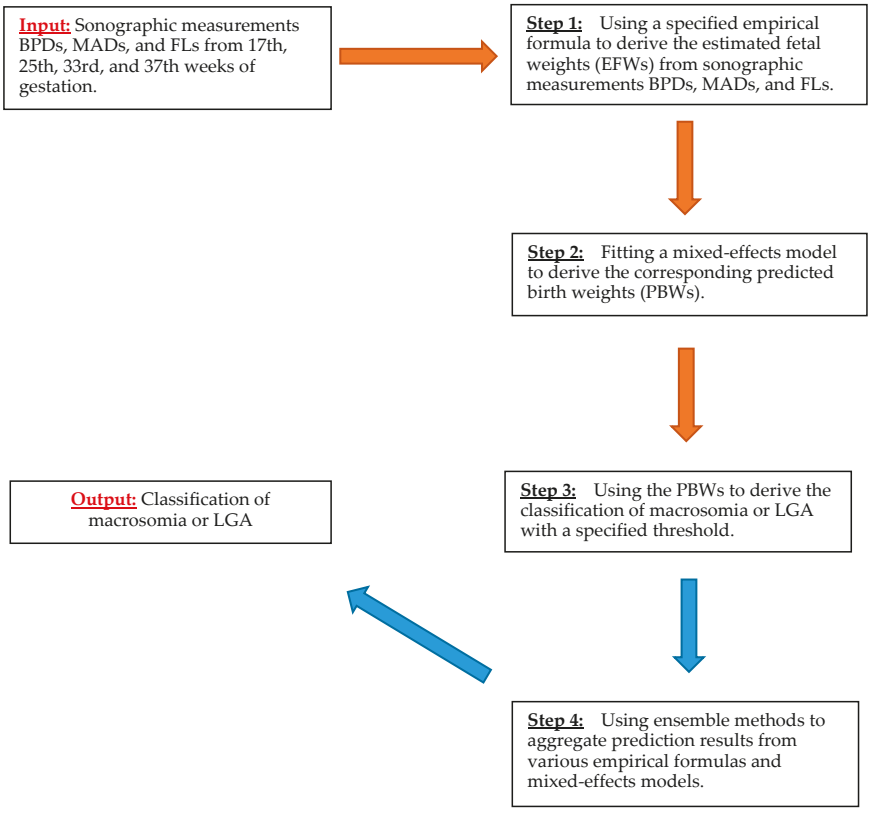


Figure 1. An ensemble learning procedure to predict macrosomia or LGA from prenatal ultrasound measurements.

We compared the above nonlinear mixed-effects models with the following quadratic mixed-effects model implemented in Zhang et al. [4].

$$y_{ij} = X\prime_{ij}\beta + \theta_1 t_{ij} + \theta_2 t_{ij}^2 + b_{i0} + b_{i1}t_{ij} + b_{i2}t_{ij}^2 + \epsilon_{ij}.$$

The configuration of X_{ij}, β, b_i, and ϵ_{ij} is identical to those in the nonlinear mixed-effects model, and $\theta = (\theta_1, \theta_2)$ are the parameters of time fixed-effects t_{ij} and t_{ij}^2. For the quadratic mixed-effects

model, we also considered the one without any covariates for predicting the birth weights and deriving the classification of macrosomia or LGA.

2.5. Ensemble Learning Methods

Here, we propose to apply ensemble methods [14], to combine the prediction results generated from 26 individual EFW empirical formulas and from nonlinear and linear mixed-effects models. One appealing property for ensemble methods is that they combine the classification strengths of individual models but do not overfit the data. Two types of ensemble algorithms, majority voting and stacking, are considered. Majority voting is one of the most fundamental ensemble methods for classification [14]. For a binary classification problem, the final classification of majority voting is the class that receives more than half of the votes from the individual learning models. Because the individual learning models can be correlated, it is necessary to select among individual learning models that are combined in majority voting [15–17]. Here, we implemented least absolute shrinkage and selection operator (LASSO) [18], smoothed clipped absolute deviation (SCAD) [19], and minimax penalized likelihood (MCP) [20] to select individual learning models.

The stacking method [21,22] is one of the most known meta-learning methods. It combines the prediction results from several different individual learning models, called "first-level learners", by another learning model, named "second-level learner" or "meta-learner". Van der Laan et al. [23] proposed to train the meta-learner by a unified cross-validation algorithm or super learner and proved its oracle properties. The unified cross-validation algorithm is summarized as follows. For a K-fold cross-validation procedure ($K = 10$ is set here), the training dataset was split into K equal-sized groups, stratified by the response variable. Then, let the k-th group be the validation data, take the remaining data to train the first-level learners, and collect the prediction values from the validation data as the covariates of the meta-leaner. By repeating the above procedure on every fold of data, along with the original response variable, a complete dataset, called the "leave-one data", is generated for training the meta-learner. Several reports [21,22,24] proposed using the logistic regression with positive constraint on the regression coefficients as the meta-learner for classification, whereas Van der Laan et al. [23] used linear regression as the meta-learner for regression models. Debray et al. [24] suggested performing model selection on the meta-learner, so we applied LASSO, SCAD, and MCP to conduct model selection on the meta-learner.

2.6. Evaluation of Prediction Performance

In our investigation, the original dataset was randomly divided into a training dataset (70%, n = 781) and a testing dataset (30%, n = 334), stratified by the presence of macrosomia or LGA. For an individual mixed-effects model with a particular EFW empirical formula, the entire training dataset, along with the ultrasound measures and demographic information in the testing dataset, was used to fit the model. The birth time, birthweight, and the true macrosomia or LGA status in the testing dataset were used to evaluate the prediction performance. For ensemble methods, the ensemble learner was trained by the training dataset and tested by the testing dataset.

The prediction accuracy of each individual mixed-effects model with a particular EFW empirical formula, as well as the ensemble learner, was assessed by the areas under the receiver operating characteristic curve (AUC), sensitivity, specificity, positive predictive value (+PV), negative predictive value (−PV), positive likelihood ratio (+LR), negative likelihood ratio (−LR), and Youden's index (sensitivity + specificity − 1) for predicting both macrosomia and LGA.

3. Results

Table 1 shows the baseline characteristics of our study subjects (n = 1115). The mean maternal age of the study subjects was 28 years (standard deviation or SD: 4 years), and the average height and weight before pregnancy were 166 centimeters (SD: 6 centimeters) and 59 kg (SD: 10 kg), respectively. Twenty-one percent of women had a history of SGA births, and about half of them smoked at enrollment.

Only a few had high blood pressure, cardiac disease, diabetes, or renal disorder, but 15% had other types of diseases. The mean gestational age at birth was 280 days (SD: 8 days), and the mean birthweight was 3562 g (SD: 478 g). Seventeen percent of infants had macrosomia, 11% had LGA at birth, and, among the LGA infants, 9 of them (7%) were not macrosomia infants. The current study sample was used as the reference of LGA [11].

Table 1. Baseline characteristics of study subjects.

Demographic Characteristic	Mean (Standard Deviation) n = 1115
Maternal age (years)	28.32 (4.12)
Birthweight (grams)	3562 (478)
Gestational age (days)	279.99 (8.34)
Maternal height (centimeters)	165.96 (5.99)
Maternal weight (kilograms)	59.20 (10.00)
Health History	**Number of Subjects (Percentage)**
High blood pressure	20 (1.8)
Cardiac disease	10 (0.9)
Diabetes	3 (0.3)
Renal disorders	11 (1.0)
Other diseases	172 (15.4)
Small for gestational age	235 (21.1)
Large for gestational age	124 (11.1)
Macrosomia	195 (17.5)
Smoking at enrollment (cigarettes/day)	
0	436 (49.1)
1–9	180 (16.1)
10–19	397 (35.6)
20+	102 (9.2)

3.1. Prediction Performance of Individual Models and Empirical Formulas in Macrosomia

We predicted macrosomia or LGA with the nonlinear and quadratic mixed-effects models described above, and also considered those models with or without the covariates. Thus, four mixed-effects models combined with 26 empirical formulas for EFWs, totally 104 learning models, were fitted. The prediction performance of individual learning models in predicting macrosomia is reported in Appendix A Tables A2–A5. For the three-parameter mixed-effects logistic models, their prediction performance varied with different EFW empirical formulas, and adding covariates into the models did not improve their prediction performance (see Appendix A Tables A2 and A3). The three-parameter mixed-effects logistic models, either with or without covariates, predicted all birth weights to be under 4000 g when combined with empirical formula 3 in Table A1. The two nonlinear mixed-effects models combined with empirical formula 11 in Table A1 gave the highest value of Youden's index of 0.670, and the lowest value of Youden's index of 0.050 was obtained by empirical formula 18 in Table A1 without covariates. The best sensitivities of 0.857 and 0.839 were obtained from the two models with or without covariates, respectively, when coupled with empirical formula 7 in Table A1. The AUCs for these models ranged from 0.871 to 0.910. The prediction performance of the quadratic mixed-effects models also varied among different empirical formulas, and adding covariates did not improve their prediction performance either (see Appendix A Tables A4 and A5). The quadratic mixed-effects models combined with empirical formula 12 in Table A1 gave the highest value of Youden's index of 0.663, whereas the lowest value of Youden's index of 0.310 was obtained by empirical formula 3, in Table A1, without covariates. The best sensitivity of 0.982 was obtained from the two models with or without covariates when coupled with empirical formulas 7 and 11 in Table A1. The AUCs for these models ranged from 0.857 to 0.905.

3.2. Prediction Performance of Individual Models and Empirical Formulas in LGA

LGA is defined as a newborn with a birthweight greater than the 90th percentile for gestational age. Here, given the input of sonographic ultrasound measures, both the nonlinear and quadratic mixed-effects models can generate PBWs for each of fetuses at any time point, regardless of its actual birth time. A newborn was classified as LGA if his or her PBW was above the 90th percentile of PBWs of all other fetuses, when the fetuses were assumed to be born at the identical birth time as that newborn. The prediction performance of the nonlinear and quadratic mixed-effects models is reported in Appendix A Tables A6–A9. The prediction performance showed small variation among different EFWs empirical formulas and among these models. The Youden's indexes ranged from 0.354 to 0.526, the sensitivities ranged from 0.424 to 0.576, the specificities ranged from 0.930 to 0.950, and the AUCs ranged from 0.863 to 0.894.

3.3. Prediction Performance of Ensemble Methods

Two ensemble methods, voting and stacking methods, were applied to combine the 104 learning models for the prediction of macrosomia. The voting method was implemented by using two approaches: voting from all learning models without selection and voting from the selected learning models. To select among the learning models, a penalized logistic regression was run with either a LASSO, SCAD, or MCP penalty. When the stacking method was implemented, we fit a linear regression model as our meta-learner, either without variable selection or with three variable selection methods, LASSO, SCAD, and MCP. The prediction performance of the ensemble methods for the prediction of macrosomia is summarized in Table 2. The voting and stacking methods with the SCAD selection yielded the best Youden's indexes of 0.681 and 0.688, respectively, which were higher than the Youden's indexes generated by any other individual learning models in Tables A2–A5. The voting method with the SCAD and MCP selection and the stacking method with the LASSO, SCAD, and MCP selection outperformed most of the individual learning models listed in Tables A2–A5. The voting method with the SCAD or MCP model selection generated an AUC of 0.932 and 0.924, respectively, higher than the AUCs generated by any other individual learning models in Tables A2–A5.

We also applied the voting and stacking methods for the prediction of LGA, to combine the 104 learning models. The voting method was implemented as described in the prediction of macrosomia. When the stacking method was implemented, we fit a logistic regression with positive constraints [22], without further selection on first-level learners, and a logistic regression with the LASSO, SCAD, and MCP selection on first-level learners as our meta-learner. The results are summarized in Table 3. These results showed that the best prediction results were supplied by the voting method with the MCP selection with a Youden's index of 0.537 and a sensitivity of 0.636, both higher than any of the individual learning models in Tables A6–A9.

Table 2. Prediction performance of macrosomia by the ensemble methods.

Method	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	+LR (95% CI)	−LR (95% CI)	+PV (95% CI)	−PV (95% CI)	Youden's Index
Stacking: no selection	0.852 (0.805, 0.898)	0.607 (0.468, 0.735)	0.867 (0.821, 0.905)	4.562 (3.162, 6.582)	0.453 (0.326, 0.630)	0.479 (0.359, 0.601)	0.916 (0.876, 0.900)	0.474
Stacking: LASSO	0.907 (0.891, 0.943)	0.768 (0.636, 0.870)	0.874 (0.829, 0.911)	6.099 (4.334, 8.582)	0.266 (0.165, 0.429)	0.551 (0.434, 0.664)	0.949 (0.915, 0.973)	0.642
Stacking: SCAD	0.909 (0.873, 0.944)	0.839 (0.717, 0.924)	0.849 (0.801, 0.889)	5.555 (4.110, 7.509)	0.189 (0.104, 0.345)	0.528 (0.419, 0.635)	0.963 (0.931, 0.983)	0.688
Stacking: MCP	0.907 (0.872, 0.943)	0.786 (0.656, 0.844)	0.863 (0.817, 0.901)	5.748 (4.151, 7.960)	0.258 (0.150, 0.411)	0.537 (0.423, 0.647)	0.952 (0.918, 0.975)	0.649
Voting: no selection	0.903 (0.887, 0.919)	0.613 (0.553, 0.670)	0.912 (0.900, 0.927)	6.991 (5.726, 8.536)	0.425 (0.366, 0.492)	0.608 (0.549, 0.665)	0.914 (0.897, 0.929)	0.525
Voting: LASSO	0.891 (0.871, 0.911)	0.542 (0.482, 0.601)	0.959 (0.947, 0.969)	8.885 (6.988, 11.296)	0.809 (0.552, 0.673)	0.664 (0.599, 0.724)	0.902 (0.885, 0.918)	0.481
Voting: SCAD	0.932 (0.918, 0.945)	0.845 (0.798, 0.885)	0.836 (0.814, 0.855)	5.143 (4.501, 5.876)	0.185 (0.141, 0.244)	0.681 (0.612, 0.741)	0.533 (0.486, 0.580)	0.681
Voting: MCP	0.924 (0.910, 0.938)	0.810 (0.759, 0.854)	0.854 (0.834, 0.873)	5.565 (4.817, 6.428)	0.223 (0.175, 0.283)	0.553 (0.504, 0.601)	0.953 (0.939, 0.964)	0.664

Areas under the receiver operating characteristic curve (AUC), confident interval (CI), positive likelihood ratio (+LR), negative likelihood ratios (−LR), positive predictive values (+PV), negative predictive values (−PV).

Table 3. Prediction performance of large for gestational age by the ensemble methods.

Method	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	+LR (95% CI)	−LR (95% CI)	+PV (95% CI)	−PV (95% CI)	Youden's Index
Stacking: no selection	0.785 (0.697, 0.874)	0.485 (0.308, 0.665)	0.947 (0.915, 0.969)	9.121 (5.044, 16.495)	0.544 (0.390, 0.758)	0.500 (0.319, 0.681)	0.944 (0.911, 0.967)	0.432
Stacking: LASSO	0.811 (0.728, 0.894)	0.455 (0.281, 0.636)	0.944 (0.911, 0.967)	8.048 (4.443, 14.578)	0.578 (0.423, 0.790)	0.469 (0.291, 0.653)	0.940 (0.907, 0.964)	0.398
Stacking: SCAD	0.785 (0.696, 0.873)	0.455 (0.281, 0.636)	0.944 (0.911, 0.967)	8.048 (4.443, 14.578)	0.578 (0.423, 0.790)	0.469 (0.291, 0.653)	0.940 (0.907, 0.964)	0.398
Stacking: MCP	0.775 (0.688, 0.863)	0.485 (0.308, 0.665)	0.937 (0.903, 0.962)	7.681 (4.389, 13.441)	0.550 (0.394, 0.767)	0.457 (0.288, 0.634)	0.943 (0.911, 0.967)	0.421
Voting: no selection	0.823 (0.740, 0.906)	0.485 (0.308, 0.665)	0.950 (0.919, 0.972)	9.729 (5.309, 17,831)	0.542 (0.389, 0.756)	0.516 (0.331, 0.698)	0.944 (0.912, 0.967)	0.435
Voting: LASSO	0.811 (0.728, 0.895)	0.515 (0.335, 0.692)	0.940 (0.907, 0.964)	8.614 (4.936, 15.035)	0.516 (0.362, 0.734)	0.486 (0.314, 0.660)	0.946 (0.915, 0.969)	0.455
Voting: SCAD	0.799 (0.713, 0.885)	0.515 (0.335, 0.692)	0.944 (0.911, 0.967)	9.121 (5.168, 16.099)	0.514 (0.361, 0.731)	0.500 (0.319, 0.681)	0.947 (0.915, 0.969)	0.459
Voting: MCP	0.774 (0.687, 0.861)	0.636 (0.451, 0.796)	0.900 (0.861, 0.932)	6.385 (4.168, 9.780)	0.404 (0.257, 0.635)	0.412 (0.276, 0.558)	0.958 (0.927, 0.978)	0.537

Areas under the receiver operating characteristic curve (AUC), confident interval (CI), positive likelihood ratio (+LR), negative likely ratios (−LR), positive predictive values (+PV), negative predictive values (−PV).

4. Discussion and Conclusions

We proposed using ensemble methods to combine the strengths from nonlinear and quadratic mixed-effects models and 26 empirical formulas of EFWs to predict macrosomia and LGA of newborns from sonographic ultrasound measurements. The prediction performance of the ensemble methods was studied with the data from the Scandinavia SGA study. We showed that the prediction performance varied among the empirical formulas and mixed-effects models. The three-parameter mixed-effects logistic model combined with empirical formula 11 in Table A1 gave the best prediction results in predicting macrosomia. In predicting LGA, the prediction performance also varied. The best prediction results were obtained from the quadratic mixed-effects model combined with empirical formula 6 in Table A1. These results showed that it was difficult to select any individual statistical learning model combined with only one empirical formula of EFWs, to predict macrosomia or LGA from sonographic ultrasound measurements.

We subsequently proposed applying ensemble methods to aggregate the prediction results from mixed-effects models and empirical formulas of EFWs, to achieve better prediction performance. Our investigation showed that, with the aid of either the SCAD or MCP for model selection, both stacking and voting methods improved the prediction accuracies in predicting macrosomia, as opposed to those from individual statistical models and empirical formulas. The voting method with the MCP for model selection predicted LGA more accurately than the individual statistical models and empirical formulas. In this study, the ensemble prediction algorithms were created from all the prenatal ultrasound measures and birth weights. However, the algorithms can be used to make prediction on macrosomia or LGA with the ultrasound measures only from the first or second trimester, although the ultrasound measures collected in the third trimester or before birth can substantially improve the prediction accuracy. Our current study is a feasibility study that demonstrates the ensemble methods can be integrated with ultrasound examinations to assist obstetricians in clinical diagnosis on whether a pregnant woman will give birth to a large infant, and to further guide clinical interventions for the condition. However, measurable clinical benefits are unclear until they are demonstrated in prospective clinical studies.

Our current study has several limitations. The proportion of macrosomia or LGA infants in the Scandinavia SGA dataset is small (15% for macrosomia and 11% LGA of n = 1115 infants). This may largely influence the prediction accuracy given by the individual models and empirical formulas. However, using machine learning methods specifically ensemble methods, can accommodate such imbalanced data, and improve prediction accuracy [23]. In addition, the initial objective of the SGA study was to characterize the intra-uterine growth restriction and assess the associated risk factors of SGA, but the study was not designed for studying macrosomia or LGA. As a consequence, the risk factors associated with macrosomia or LGA were not thoroughly collected. This may be the reason why, in our study, we did not receive benefits from adding covariates into the mixed-effects models. Lastly, The SGA study was conducted in 1980s, with out-of-date sonographic ultrasound examination technologies, so the prediction models developed in this study may not be directly applied to predict macrosomia or LGA with the ultrasound measures from most recent state-of-the-art ultrasound technologies. However, the ensemble methods can still be applied to aggregate any available ultrasound prediction models. Also, the subjects in this study were the Caucasians from Europe that were mostly not obese (11.6% overweight rate and 2.1% obese rate in the study). New studies and data on a diverse population should be able to substantially improve the prediction of macrosomia and LGA among both the whites and minorities, as well as the overweight and obese population.

Author Contributions: S.Y., H.Z., F.S., and B.Z. conceived of the presented methodology. S.Y., F.S., and B.Z. verified the analytical methods and conducted data analysis. S.Y., F.S., and B.Z. took the lead in writing the manuscript. H.Z., S.W., and J.G. provided critical feedback, helped shape the research and analysis, and helped in manuscript writing. All authors proved the manuscript contents. All authors have read and agreed to the published version of the manuscript.

Funding: Dr. Suzhen Wang's research was partially supported by the National Natural Science Foundation of China (No. 81872719), the National Bureau of Statistics Foundation Project (No. 2018LY79), the Natural Science Foundation of Shandong Province (No. ZR201807090257), and the Poverty Alleviation Fund Project of Weifang Medical University (No. FP1801001). Dr. Fuyan Shi's research was partially supported by the National Natural Science Foundation of China (No. 81803337), the Shandong Provincial Government Fund for Overseas Study (No. 27, 2019, Lu-Jiao), the Shandong Science and Technology Development Plan Project (No. 2015 WS0067), and the Weifang Medical University Doctoral Foundation Project (No. 2017BSQD51). Dr. Hui Zhang receives support from the Robert H. Lurie Comprehensive Cancer Center of Northwestern University in Chicago, IL, which is supported in part by a NCI Cancer Center Support Grant #P30CA060553, and Mesulam Center for Cognitive Neurology and Alzheimer's Disease of Northwestern University, which is supported in part by a NIA Center Support Grant #P30AG13854.

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. Empirical formulas used for estimating fetal weight from sonographic ultrasound measurements (Melamed et al. [5]).

Formula	Reference	Equation
1	Campbell and Wilkin [25]	$\ln EFW = -4.564 + 0.383 \cdot AC - 0.00331 \cdot AC^2$
2	Hadlock et al. [5]	$\ln EFW = 2.695 + 0.253 \cdot AC - 0.00275 \cdot AC^2$
3	Jordaan [26]	$\log_{10} EFW = 0.6328 + 0.1881 \cdot AC - 0.0043 \cdot AC^2 + 0.000036239 \cdot AC^3$
4	Warsof et al. [27]	$\log_{10} EFW = 1.1633 + 0.092 \cdot AC - 0.000019 \cdot AC$
5	Higginbottom et al. [28]	$EFW = 0.0816 \cdot AC^3$
6	Hadlock et al. [5]	$\log_{10} EFW = 1.304 + 0.05281 \cdot AC + 0.1938 \cdot FL - 0.004 \cdot AC \cdot FL$
7	Woo et al. [29]	$\log_{10} EFW = 0.59 + 0.08 \cdot AC + 0.28 \cdot FL - 0.00716 \cdot AC \cdot FL$
8	Warsof et al. [30]	$\ln EFW = 2.792 + 0.108 \cdot FL + 0.0036 \cdot AC^2 - 0.0027 \cdot AC \cdot FL$
9	Vintzileos et al. [31]	$\log_{10} EFW = 1.879 + 0.084 \cdot BPD + 0.026 \cdot AC$
10	Warsof et al. [27]	$\log_{10} EFW = 1.401 + 0.144 \cdot BPD + 0.032 \cdot AC - 0.000111 \cdot AC \cdot BPD^2$
11	Shepard et al. [32]	$\log_{10} EFW = 1.2508 + 0.166 \cdot BPD + 0.046 \cdot AC - 0.002546 \cdot BPD \cdot AC$
12	Jordaan [26]	$\log_{10} EFW = 1.9317 + 0.0377 \cdot AC + 0.0950 \cdot BPD - 0.0015 \cdot BPD \cdot AC$
13	Hadlock et al. [5]	$\log_{10} EFW = 1.1134 + 0.05845 \cdot AC - 0.000604 \cdot AC^2 - 0.007365 \cdot BPD^2 + 0.000595 \cdot BPD \cdot AC + 0.1694 \cdot BPD$
14	Woo et al. [29]	$\log_{10} EFW = 1.63 + 0.16 \cdot BPD + 0.00111 \cdot AC^2 - 0.0000859 \cdot BPD \cdot AC^2$
15	Mirghani et al. [33]	$\log_{10} EFW = 2.1315 + 0.0056541 \cdot AC \cdot BPD - 0.00015515 \cdot BPD \cdot AC^2 + 0.000019782 \cdot AC^3 + 0.052594 \cdot BPD$
16	Hadlock et al. [5]	$\log_{10} EFW = 1.182 + 0.0273 \cdot HC + 0.07057 \cdot AC - 0.00063 \cdot AC^2 - 0.0002184 \cdot HC \cdot AC$
17	Jordaan [26]	$\log_{10} EFW = 0.9119 + 0.0488 \cdot HC + 0.0824 \cdot AC - 0.001599 \cdot HC \cdot AC$
18	Jordaan [26]	$\log_{10} EFW = 2.3231 + 0.02904 \cdot AC + 0.0079 \cdot HC - 0.0058 \cdot BPD$
19	Hadlock et al. [5]	$\log_{10} EFW = 1.335 - 0.0034 \cdot AC \cdot FL + 0.0316 \cdot BPD + 0.0457 \cdot AC + 0.1623 \cdot FL$
20	Woo et al. [29]	$\log_{10} EFW = 1.54 + 0.15 \cdot BPD + 0.00111 \cdot AC^2 - 0.0000764 \cdot BPD \cdot AC^2 + 0.05 \cdot FL - 0.000992 \cdot FL \cdot AC$
21	Shinozuka et al. [34]	$EFW = 0.23966 \cdot AC^2 \cdot FL + 1.6230 \cdot BPD^3$
22	Mirghani et al. [33]	$\log_{10} EFW = 2.7193 + 0.0094962 \cdot AC \cdot BPD - 0.1432 \cdot FL - 0.00076742 \cdot AC \cdot BPD^2 + 0.001745 \cdot FL \cdot BPD^2$
23	Hadlock et al. [5]	$\log_{10} EFW = 1.326 - 0.00326 \cdot AC \cdot FL + 0.0107 \cdot HC + 0.0438 \cdot AC + 0.158 \cdot FL$
24	Combs et al. [35]	$EFW = 0.23718 \cdot AC^2 \cdot FL + 0.03312 \cdot HC^3$
25	Ott et al. [36]	$\log_{10} EFW = 0.9339 + 0.04355 \cdot HC + 0.05392 \cdot AC - 0.0008582 \cdot HC \cdot AC + 1.2594 \cdot (FL/AC)$
26	Hadlock et al. [5]	$\log_{10} EFW = 1.3596 + 0.0064 \cdot HC + 0.0424 \cdot AC + 0.174 \cdot FL + 0.00061 \cdot BPD \cdot AC - 0.00386 \cdot AC \cdot FL$

Abdominal circumference (AC), femur length (FL), biparietal diameter (BPD), and head circumference (HC) are expressed in centimeters, and EFW is expressed in grams, unless stated otherwise. This is an identical table to the Table 1 in Melamed et al. [5].

Table A2. Prediction performance of the three-parameter mixed-effects logistic model without covariance when combined with the 26 estimated fetal weight empirical formulas * in predicting macrosomia.

Formulas	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	+LR (95% CI)	−LR (95% CI)	+PV (95% CI)	−PV (95% CI)	Youden's Index
1	0.884 (0.839, 0.930)	0.250 (0.144, 0.384)	0.982 (0.958, 0.994)	13.900 (5.217, 37.033)	0.764 (0.656, 0.889)	0.737 (0.488, 0.910)	0.867 (0.824, 0.902)	0.232
2	0.885 (0.840, 0.930)	0.464 (0.330, 0.603)	0.946 (0.913, 0.969)	8.605 (4.881, 15.169)	0.566 (0.443, 0.724)	0.634 (0.469, 0.779)	0.898 (0.857, 0.930)	0.410
3	All predicted birth weights were under 4000 g.							
4	0.886 (0.842, 0.931)	0.571 (0.432, 0.703)	0.914 (0.874, 0.944)	6.619 (4.243, 10.325)	0.469 (0.346, 0.636)	0.571 (0.432, 0.703)	0.914 (0.874, 0.944)	0.485
5	0.883 (0.838, 0.928)	0.536 (0.397, 0.670)	0.932 (0.895, 0.958)	7.838 (4.765, 12.895)	0.498 (0.375, 0.661)	0.612 (0.462, 0.748)	0.909 (0.869, 0.940)	0.467
6	0.882 (0.835, 0.928)	0.464 (0.330, 0.603)	0.924 (0.887, 0.953)	6.146 (3.734, 10.116)	0.579 (0.453, 0.741)	0.553 (0.401, 0.698)	0.895 (0.854, 0.928)	0.389
7	0.882 (0.837, 0.926)	0.857 (0.738, 0.936)	0.799 (0.747, 0.844)	4.255 (3.290, 5.504)	0.179 (0.094, 0.341)	0.462 (0.363, 0.562)	0.965 (0.933, 0.985)	0.656
8	0.872 (0.824, 0.920)	0.429 (0.297, 0.568)	0.932 (0.895, 0.958)	6.271 (3.695, 10.643)	0.613 (0.488, 0.771)	0.558 (0.399, 0.709)	0.890 (0.848, 0.924)	0.360
9	0.907 (0.871, 0.942)	0.821 (0.696, 0.911)	0.845 (0.797, 0.886)	5.311 (3.931, 7.174)	0.211 (0.120, 0.371)	0.517 (0.408, 0.624)	0.959 (0.926, 0.980)	0.667
10	0.906 (0.867, 0.945)	0.536 (0.397, 0.670)	0.953 (0.921, 0.975)	11.456 (6.388, 20.544)	0.487 (0.367, 0.646)	0.698 (0.539, 0.828)	0.911 (0.872, 0.941)	0.489
11	0.904 (0.867, 0.942)	0.821 (0.696, 0.911)	0.849 (0.801, 0.889)	5.437 (4.011, 7.370)	0.210 (0.120, 0.370)	0.523 (0.414, 0.630)	0.959 (0.927, 0.980)	0.670
12	0.910 (0.873, 0.946)	0.643 (0.504, 0.766)	0.921 (0.883, 0.950)	8.123 (5.201, 12.689)	0.388 (0.272, 0.552)	0.621 (0.484, 0.745)	0.928 (0.890, 0.955)	0.564
13	0.908 (0.869, 0.946)	0.696 (0.559, 0.812)	0.910 (0.870, 0.941)	7.744 (5.129, 11.693)	0.334 (0.224, 0.497)	0.609 (0.479, 0.729)	0.937 (0.901, 0.963)	0.607
14	0.908 (0.870, 0.946)	0.214 (0.116, 0.344)	0.989 (0.969, 0.998)	19.857 (5.792, 68.082)	0.792 (0.692, 0.911)	0.800 (0.519, 0.898)	0.862 (0.819, 0.898)	0.203
15	0.910 (0.874, 0.947)	0.732 (0.597, 0.842)	0.910 (0.870, 0.941)	9.141 (5.424, 12.220)	0.294 (0.191, 0.455)	0.621 (0.493, 0.738)	0.944 (0.909, 0.968)	0.642
16	0.905 (0.865, 0.969)	0.429 (0.297, 0.568)	0.968 (0.939, 0.985)	13.238 (6.507, 26.933)	0.591 (0.470, 0.742)	0.727 (0.545, 0.867)	0.894 (0.853, 0.926)	0.396
17	0.871 (0.824, 0.917)	0.339 (0.218, 0.478)	0.946 (0.913, 0.969)	6.288 (3.406, 11.608)	0.698 (0.578, 0.844)	0.559 (0.379, 0.728)	0.877 (0.834, 0.912)	0.285
18	0.888 (0.848, 0.928)	0.054 (0.011, 0.149)	0.996 (0.980, 1.000)	14.893 (1.578, 140.579)	0.950 (0.892, 1.011)	0.750 (0.194, 0.994)	0.839 (0.795, 0.877)	0.050
19	0.901 (0.860, 0.943)	0.679 (0.540, 0.797)	0.924 (0.887, 0.953)	8.983 (5.734, 12.074)	0.348 (0.237, 0.509)	0.644 (0.509, 0.764)	0.935 (0.899, 0.961)	0.603
20	0.910 (0.871, 0.948)	0.607 (0.468, 0.735)	0.935 (0.900, 0.961)	9.377 (5.722, 15.367)	0.420 (0.303, 0.583)	0.654 (0.509, 0.780)	0.922 (0.884, 0.950)	0.542
21	0.905 (0.865, 0.945)	0.429 (0.297, 0.568)	0.950 (0.917, 0.972)	8.510 (4.702, 15.403)	0.602 (0.479, 0.756)	0.632 (0.460, 0.782)	0.892 (0.851, 0.925)	0.378
22	0.901 (0.865, 0.937)	0.179 (0.089, 0.304)	0.982 (0.959, 0.994)	9.929 (3.529, 27.934)	0.836 (0.740, 0.946)	0.667 (0.384, 0.882)	0.856 (0.812, 0.892)	0.161
23	0.902 (0.861, 0.943)	0.464 (0.330, 0.603)	0.939 (0.904, 0.964)	7.592 (4.426, 13.025)	0.571 (0.446, 0.730)	0.605 (0.444, 0.750)	0.897 (0.856, 0.929)	0.403
24	0.906 (0.886, 0.945)	0.232 (0.130, 0.364)	0.993 (0.974, 0.999)	32.268 (7.488, 133.049)	0.773 (0.669, 0.894)	0.867 (0.595, 0.983)	0.865 (0.823, 0.901)	0.225
25	0.906 (0.866, 0.946)	0.393 (0.265, 0.532)	0.960 (0.930, 0.980)	9.929 (5.109, 19.294)	0.632 (0.511, 0.781)	0.667 (0.482, 0.820)	0.887 (0.846, 0.920)	0.353
26	0.903 (0.863, 0.944)	0.661 (0.522, 0.782)	0.932 (0.895, 0.958)	9.667 (6.025, 15.512)	0.364 (0.252, 0.526)	0.661 (0.522, 0.782)	0.932 (0.895, 0.958)	0.592

* See Table A1 for specifications of empirical formulas. Areas under the receiver operating characteristic curve (AUC), confident interval (CI), positive likelihood ratio (+LR), negative likelihood ratios (−LR), positive predictive values (+PV), negative predictive values (−PV).

Table A3. Prediction performance of the three-parameter mixed-effects logistic model with covariance when combined with the 26 estimated fetal weight empirical formulas* in predicting macrosomia.

Formulas	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	+LR (95% CI)	−LR (95% CI)	+PV (95% CI)	−PV (95% CI)	Youden's Index
1	0.885 (0.840, 0.931)	0.232 (0.130, 0.364)	0.982 (0.958, 0.994)	12.907 (4.793, 34.758)	0.782 (0.676, 0.904)	0.722 (0.465, 0.903)	0.864 (0.821, 0.900)	0.215
2	0.885 (0.830, 0.930)	0.500 (0.363, 0.637)	0.942 (0.908, 0.967)	8.688 (5.047, 14.953)	0.531 (0.408, 0.690)	0.636 (0.478, 0.776)	0.903 (0.863, 0.935)	0.442
3	All predicted birth weights were under 4000 g.							
4	0.885 (0.840, 0.930)	0.607 (0.468, 0.735)	0.917 (0.878, 0.947)	7.339 (4.705, 11.446)	0.428 (0.309, 0.594)	0.596 (0.458, 0.724)	0.921 (0.882, 0.950)	0.524
5	0.883 (0.837, 0.928)	0.518 (0.380, 0.653)	0.931 (0.895, 0.958)	7.577 (4.586, 12.520)	0.518 (0.394, 0.680)	0.604 (0.453, 0.742)	0.906 (0.866, 0.937)	0.450
6	0.880 (0.834, 0.927)	0.464 (0.330, 0.603)	0.928 (0.891, 0.956)	6.454 (3.886, 10.717)	0.577 (0.451, 0.738)	0.565 (0.411, 0.711)	0.896 (0.855, 0.929)	0.392
7	0.882 (0.838, 0.926)	0.839 (0.717, 0.924)	0.799 (0.747, 0.844)	4.166 (3.211, 5.407)	0.201 (0.110, 0.367)	0.456 (0.358, 0.557)	0.961 (0.927, 0.982)	0.638
8	0.880 (0.834, 0.927)	0.464 (0.330, 0.603)	0.928 (0.891, 0.956)	6.454 (3.886, 10.717)	0.577 (0.451, 0.738)	0.565 (0.411, 0.711)	0.896 (0.855, 0.929)	0.392
9	0.906 (0.870, 0.942)	0.804 (0.676, 0.898)	0.845 (0.797, 0.886)	5.195 (3.534, 7.039)	0.232 (0.136, 0.396)	0.511 (0.402, 0.619)	0.955 (0.921, 0.977)	0.649
10	0.907 (0.867, 0.946)	0.500 (0.363, 0.637)	0.950 (0.917, 0.972)	9.929 (5.594, 17.622)	0.527 (0.405, 0.685)	0.667 (0.505, 0.804)	0.904 (0.864, 0.935)	0.500
11	0.904 (0.867, 0.942)	0.821 (0.696, 0.911)	0.849 (0.801, 0.889)	5.437 (4.011, 7.370)	0.210 (0.120, 0.370)	0.523 (0.414, 0.630)	0.959 (0.926, 0.980)	0.670
12	0.910 (0.873, 0.946)	0.643 (0.504, 0.766)	0.921 (0.553, 0.950)	8.123 (5.201, 12.689)	0.388 (0.272, 0.552)	0.621 (0.484, 0.745)	0.928 (0.890, 0.955)	0.564
13	0.907 (0.868, 0.945)	0.679 (0.540, 0.797)	0.914 (0.874, 0.944)	7.860 (5.150, 11.996)	0.352 (0.240, 0.516)	0.613 (0.481, 0.734)	0.934 (0.897, 0.960)	0.592
14	0.909 (0.871, 0.947)	0.214 (0.116, 0.344)	0.986 (0.964, 0.996)	14.893 (4.984, 44.498)	0.797 (0.695, 0.915)	0.750 (0.476, 0.927)	0.862 (0.819, 0.898)	0.200
15	0.910 (0.874, 0.947)	0.714 (0.578, 0.827)	0.910 (0.870, 0.941)	7.943 (5.277, 11.956)	0.314 (0.207, 0.476)	0.615 (0.486, 0.733)	0.941 (0.905, 0.966)	0.624
16	0.905 (0.865, 0.944)	0.411 (0.281, 0.550)	0.968 (0.939, 0.985)	12.687 (6.205, 25.937)	0.609 (0.489, 0.759)	0.719 (0.533, 0.863)	0.891 (0.850, 0.924)	0.378
17	0.871 (0.825, 0.917)	0.339 (0.218, 0.478)	0.946 (0.913, 0.969)	6.288 (3.406, 11.608)	0.698 (0.578, 0.844)	0.559 (0.379, 0.728)	0.877 (0.834, 0.912)	0.285
18	0.885 (0.845, 0.924)	0.232 (0.130, 0.364)	0.975 (0.949, 0.990)	9.219 (3.851, 22.069)	0.788 (0.681, 0.911)	0.650 (0.408, 0.846)	0.863 (0.820, 0.899)	0.207
19	0.901 (0.860, 0.943)	0.679 (0.540, 0.797)	0.928 (0.891, 0.956)	9.432 (5.960, 14.927)	0.346 (0.236, 0.507)	0.655 (0.519, 0.775)	0.935 (0.899, 0.961)	0.607
20	0.909 (0.871, 0.948)	0.589 (0.450, 0.719)	0.932 (0.895, 0.958)	8.622 (5.303, 14.018)	0.441 (0.322, 0.604)	0.635 (0.490, 0.764)	0.918 (0.880, 0.948)	0.521
21	0.905 (0.865, 0.944)	0.429 (0.297, 0.568)	0.950 (0.917, 0.972)	8.510 (4.702, 15.403)	0.602 (0.479, 0.756)	0.632 (0.460, 0.782)	0.892 (0.851, 0.925)	0.378
22	0.903 (0.867, 0.939)	0.179 (0.089, 0.304)	0.982 (0.959, 0.994)	9.929 (3.529, 27.934)	0.836 (0.740, 0.946)	0.667 (0.384, 0.882)	0.856 (0.812, 0.892)	0.161
23	0.902 (0.861, 0.943)	0.482 (0.347, 0.620)	0.935 (0.900, 0.961)	7.446 (4.415, 12.560)	0.554 (0.429, 0.714)	0.600 (0.443, 0.743)	0.900 (0.859, 0.932)	0.417
24	0.906 (0.867, 0.946)	0.232 (0.130, 0.364)	0.993 (0.974, 0.999)	32.269 (7.488, 39.049)	0.773 (0.669, 0.894)	0.867 (0.595, 0.983)	0.865 (0.823, 0.901)	0.225
25	0.906 (0.866, 0.946)	0.393 (0.265, 0.532)	0.964 (0.935, 0.983)	10.921 (5.477, 21.778)	0.630 (0.510, 0.778)	0.688 (0.500, 0.839)	0.887 (0.846, 0.921)	0.357
26	0.903 (0.863, 0.944)	0.643 (0.504, 0.766)	0.932 (0.895, 0.958)	9.406 (5.844, 15.139)	0.383 (0.269, 0.546)	0.655 (0.514, 0.778)	0.928 (0.891, 0.956)	0.575

* See Table A1 for specifications of empirical formulas. Areas under the receiver operating characteristic curve (AUC), confident interval (CI), positive likelihood ratio (+LR), negative likelihood ratios (−LR), positive predictive values (+PV), negative predictive values (−PV).

Table A4. Prediction performance of the quadratic mixed-effects model without covariance when combined with the 26 estimated fetal weight empirical formulas* in predicting macrosomia.

Formulas	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	+LR (95% CI)	−LR (95% CI)	+PV (95% CI)	−PV (95% CI)	Youden's Index
1	0.861 (0.816, 0.906)	0.679 (0.540, 0.797)	0.845 (0.797, 0.886)	4.387 (3.158, 6.094)	0.380 (0.259, 0.558)	0.469 (0.357, 0.583)	0.929 (0.890, 0.957)	0.524
2	0.882 (0.841, 0.923)	0.839 (0.717, 0.924)	0.777 (0.723, 0.825)	3.763 (2.938, 4.820)	0.207 (0.113, 0.378)	0.431 (0.337, 0.530)	0.960 (0.925, 0.982)	0.616
3	0.877 (0.835, 0.919)	0.357 (0.234, 0.496)	0.953 (0.921, 0.975)	7.637 (4.041, 14.454)	0.674 (0.554, 0.821)	0.606 (0.421, 0.771)	0.880 (0.838, 0.915)	0.310
4	0.858 (0.813, 0.903)	0.786 (0.656, 0.884)	0.727 (0.670, 0.778)	2.874 (2.271, 3.637)	0.295 (0.178, 0.489)	0.367 (0.281, 0.459)	0.944 (0.904, 0.971)	0.512
5	0.889 (0.849, 0.930)	0.768 (0.636, 0.870)	0.827 (0.778, 0.870)	4.447 (3.311, 5.980)	0.281 (0.174, 0.453)	0.473 (0.367, 0.580)	0.947 (0.910, 0.971)	0.595
6	0.882 (0.840, 0.923)	0.839 (0.717, 0.924)	0.770 (0.716, 0.818)	3.646 (2.857, 4.651)	0.209 (0.114, 0.381)	0.423 (0.330, 0.521)	0.960 (0.925, 0.981)	0.609
7	0.874 (0.932, 0.917)	0.982 (0.904, 0.999)	0.482 (0.422, 0.542)	1.896 (1.684, 2.135)	0.037 (0.005, 0.259)	0.276 (0.215, 0.344)	0.993 (0.959, 0.999)	0.464
8	0.869 (0.825, 0.913)	0.911 (0.804, 0.970)	0.716 (0.659, 0.768)	3.205 (2.614, 3.929)	0.975 (0.944, 0.992)	0.392 (0.308, 0.482)	0.975 (0.944, 0.992)	0.627
9	0.903 (0.868, 0.938)	0.964 (0.877, 0.996)	0.662 (0.603, 0.717)	2.852 (2.401, 3.387)	0.054 (0.014, 0.211)	0.365 (0.287, 0.448)	0.989 (0.962, 0.999)	0.626
10	0.902 (0.867, 0.937)	0.839 (0.717, 0.924)	0.805 (0.754, 0.851)	4.321 (3.313, 5.634)	0.199 (0.109, 0.364)	0.465 (0.365, 0.567)	0.961 (0.928, 0.982)	0.645
11	0.895 (0.858, 0.932)	0.982 (0.904, 0.999)	0.579 (0.519, 0.638)	2.334 (2.024, 2.691)	0.031 (0.004, 0.216)	0.320 (0.251, 0.395)	0.994 (0.966, 0.999)	0.561
12	0.905 (0.870, 0.939)	0.875 (0.759, 0.948)	0.788 (0.735, 0.834)	4.123 (3.220, 5.279)	0.159 (0.079, 0.318)	0.454 (0.358, 0.552)	0.969 (0.937, 0.987)	0.663
13	0.899 (0.863, 0.935)	0.893 (0.781, 0.960)	0.745 (0.689, 0.795)	3.496 (2.805, 4.357)	0.144 (0.067, 0.307)	0.413 (0.324, 0.506)	0.972 (0.940, 0.990)	0.637
14	0.898 (0.863, 0.934)	0.643 (0.504, 0.766)	0.892 (0.850, 0.926)	5.957 (4.032, 8.801)	0.400 (0.281, 0.570)	0.545 (0.418, 0.669)	0.925 (0.887, 0.954)	0.535
15	0.904 (0.869, 0.939)	0.875 (0.759, 0.948)	0.763 (0.708, 0.811)	3.686 (2.920, 4.652)	0.164 (0.082, 0.329)	0.426 (0.334, 0.522)	0.968 (0.935, 0.987)	0.637
16	0.898 (0.863, 0.934)	0.857 (0.738, 0.936)	0.799 (0.747, 0.844)	4.255 (3.290, 5.504)	0.179 (0.094, 0.341)	0.462 (0.363, 0.562)	0.965 (0.933, 0.985)	0.656
17	0.879 (0.836, 0.921)	0.804 (0.676, 0.898)	0.788 (0.735, 0.834)	3.786 (2.917, 4.915)	0.249 (0.146, 0.425)	0.433 (0.336, 0.533)	0.952 (0.916, 0.976)	0.591
18	0.901 (0.865, 0.938)	0.804 (0.676, 0.898)	0.838 (0.789, 0.879)	4.964 (3.688, 6.682)	0.234 (0.138, 0.399)	0.500 (0.393, 0.607)	0.955 (0.921, 0.977)	0.642
19	0.895 (0.856, 0.933)	0.875 (0.759, 0.948)	0.748 (0.693, 0.798)	3.475 (2.773, 4.354)	0.167 (0.083, 0.335)	0.412 (0.322, 0.506)	0.967 (0.934, 0.987)	0.623
20	0.900 (0.864, 0.936)	0.875 (0.759, 0.948)	0.755 (0.700, 0.805)	3.577 (2.845, 4.498)	0.165 (0.082, 0.332)	0.419 (0.328, 0.514)	0.968 (0.935, 0.987)	0.630
21	0.898 (0.861, 0.935)	0.857 (0.738, 0.936)	0.752 (0.697, 0.801)	3.453 (2.741, 4.350)	0.190 (0.100, 0.362)	0.410 (0.320, 0.505)	0.963 (0.929, 0.984)	0.609
22	0.903 (0.868, 0.938)	0.875 (0.759, 0.948)	0.758 (0.704, 0.808)	3.631 (2.882, 4.574)	0.165 (0.082, 0.330)	0.422 (0.331, 0.518)	0.968 (0.935, 0.987)	0.634
23	0.896 (0.858, 0.934)	0.839 (0.717, 0.924)	0.791 (0.739, 0.838)	4.023 (3.114, 5.197)	0.203 (0.111, 0.371)	0.448 (0.350, 0.548)	0.961 (0.927, 0.982)	0.631
24	0.898 (0.861, 0.935)	0.786 (0.656, 0.884)	0.831 (0.782, 0.873)	4.647 (3.463, 6.238)	0.258 (0.156, 0.427)	0.484 (0.377, 0.591)	0.951 (0.915, 0.974)	0.617
25	0.900 (0.864, 0.936)	0.821 (0.696, 0.911)	0.795 (0.743, 0.841)	4.006 (3.084, 5.205)	0.225 (0.128, 0.395)	0.447 (0.349, 0.548)	0.957 (0.922, 0.979)	0.616
26	0.897 (0.859, 0.934)	0.857 (0.738, 0.936)	0.758 (0.707, 0.808)	3.557 (2.813, 4.496)	0.188 (0.099, 0.359)	0.417 (0.326, 0.513)	0.963 (0.929, 0.984)	0.616

* See Table A1 for specifications of empirical formulas. Areas under the receiver operating characteristic curve (AUC), confident interval (CI), positive likelihood ratio (+LR), negative likelihood ratios (−LR), positive predictive values (+PV), negative predictive values (−PV).

Table A5. Prediction performance of the quadratic mixed-effects model with covariance when combined with the 26 estimated fetal weight empirical formulas* in predicting macrosomia.

Formulas	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	+LR (95% CI)	−LR (95% CI)	+PV (95% CI)	−PV (95% CI)	Youden's Index
1	0.860 (0.815, 0.905)	0.661 (0.522, 0.782)	0.845 (0.797, 0.886)	4.272 (3.062, 5.958)	0.401 (0.278, 0.580)	0.463 (0.350, 0.578)	0.925 (0.886, 0.954)	0.506
2	0.882 (0.841, 0.923)	0.839 (0.717, 0.924)	0.781 (0.727, 0.828)	3.825 (2.980, 4.909)	0.206 (0.113, 0.376)	0.435 (0.340, 0.534)	0.960 (0.926, 0.982)	0.620
3	0.877 (0.835, 0.919)	0.357 (0.234, 0.496)	0.953 (0.921, 0.975)	7.637 (4.041, 14.434)	0.674 (0.554, 0.821)	0.606 (0.421, 0.771)	0.880 (0.838, 0.915)	0.310
4	0.857 (0.812, 0.903)	0.786 (0.656, 0.884)	0.730 (0.674, 0.781)	2.912 (2.298, 3.691)	0.293 (0.177, 0.487)	0.370 (0.283, 0.463)	0.944 (0.905, 0.971)	0.516
5	0.889 (0.849, 0.930)	0.768 (0.636, 0.870)	0.827 (0.778, 0.870)	4.447 (3.311, 5.972)	0.281 (0.174, 0.453)	0.473 (0.367, 0.580)	0.947 (0.910, 0.971)	0.595
6	0.881 (0.839, 0.923)	0.839 (0.717, 0.924)	0.770 (0.716, 0.820)	3.646 (2.857, 4.651)	0.209 (0.114, 0.381)	0.423 (0.330, 0.521)	0.960 (0.925, 0.981)	0.609
7	0.873 (0.830, 0.916)	0.982 (0.904, 0.999)	0.478 (0.418, 0.539)	1.883 (1.673, 2.119)	0.037 (0.005, 0.261)	0.275 (0.214, 0.342)	0.993 (0.959, 1.000)	0.461
8	0.868 (0.824, 0.912)	0.911 (0.804, 0.970)	0.712 (0.655, 0.765)	3.165 (2.585, 3.874)	0.125 (0.054, 0.290)	0.389 (0.305, 0.478)	0.975 (0.943, 0.992)	0.623
9	0.902 (0.867, 0.937)	0.964 (0.877, 0.996)	0.669 (0.610, 0.724)	2.914 (2.447, 3.470)	0.053 (0.014, 0.209)	0.370 (0.292, 0.454)	0.989 (0.962, 0.999)	0.633
10	0.902 (0.867, 0.937)	0.839 (0.717, 0.924)	0.806 (0.754, 0.851)	4.321 (3.313, 5.634)	0.199 (0.109, 0.364)	0.465 (0.365, 0.567)	0.961 (0.928, 0.982)	0.645
11	0.893 (0.856, 0.931)	0.982 (0.904, 0.999)	0.579 (0.519, 0.638)	2.334 (2.024, 2.691)	0.031 (0.004, 0.216)	0.320 (0.251, 0.395)	0.994 (0.966, 1.000)	0.561
12	0.905 (0.870, 0.939)	0.875 (0.759, 0.948)	0.784 (0.731, 0.831)	4.054 (3.173, 5.179)	0.159 (0.079, 0.320)	0.500 (0.354, 0.548)	0.969 (0.937, 0.987)	0.659
13	0.898 (0.862, 0.934)	0.893 (0.781, 0.960)	0.748 (0.693, 0.798)	3.546 (2.840, 4.427)	0.143 (0.067, 0.306)	0.417 (0.327, 0.510)	0.972 (0.940, 0.990)	0.641
14	0.898 (0.875, 0.948)	0.643 (0.504, 0.766)	0.892 (0.850, 0.926)	5.957 (4.032, 8.801)	0.400 (0.281, 0.570)	0.545 (0.418, 0.669)	0.925 (0.887, 0.954)	0.535
15	0.903 (0.869, 0.938)	0.875 (0.759, 0.948)	0.759 (0.704, 0.808)	3.631 (2.882, 4.574)	0.165 (0.082, 0.330)	0.422 (0.331, 0.518)	0.968 (0.935, 0.987)	0.634
16	0.898 (0.862, 0.934)	0.857 (0.738, 0.936)	0.799 (0.747, 0.844)	4.255 (3.290, 5.504)	0.179 (0.094, 0.341)	0.462 (0.363, 0.562)	0.965 (0.933, 0.985)	0.656
17	0.878 (0.836, 0.921)	0.786 (0.656, 0.884)	0.788 (0.735, 0.834)	3.702 (2.842, 4.823)	0.272 (0.164, 0.451)	0.427 (0.330, 0.528)	0.948 (0.911, 0.973)	0.573
18	0.901 (0.864, 0.938)	0.804 (0.676, 0.898)	0.842 (0.793, 0.883)	5.077 (3.760, 6.856)	0.233 (0.137, 0.397)	0.506 (0.398, 0.613)	0.955 (0.921, 0.977)	0.645
19	0.894 (0.856, 0.932)	0.875 (0.759, 0.948)	0.745 (0.689, 0.795)	3.426 (2.739, 4.285)	0.168 (0.084, 0.337)	0.408 (0.320, 0.502)	0.967 (0.934, 0.987)	0.620
20	0.900 (0.865, 0.936)	0.875 (0.759, 0.948)	0.755 (0.700, 0.805)	3.577 (2.845, 4.498)	0.165 (0.082, 0.332)	0.419 (0.328, 0.514)	0.968 (0.935, 0.987)	0.630
21	0.897 (0.860, 0.934)	0.857 (0.738, 0.936)	0.755 (0.700, 0.805)	3.504 (2.777, 4.422)	0.189 (0.099, 0.360)	0.414 (0.323, 0.509)	0.963 (0.929, 0.984)	0.613
22	0.903 (0.868, 0.938)	0.875 (0.759, 0.948)	0.755 (0.700, 0.805)	3.577 (2.845, 4.498)	0.165 (0.082, 0.332)	0.419 (0.328, 0.514)	0.968 (0.935, 0.987)	0.630
23	0.895 (0.857, 0.933)	0.821 (0.696, 0.911)	0.791 (0.739, 0.838)	3.937 (3.037, 5.104)	0.226 (0.128, 0.397)	0.442 (0.345, 0.543)	0.957 (0.921, 0.979)	0.613
24	0.898 (0.861, 0.935)	0.786 (0.656, 0.884)	0.831 (0.782, 0.873)	4.647 (3.463, 6.238)	0.258 (0.156, 0.427)	0.484 (0.377, 0.591)	0.951 (0.915, 0.974)	0.617
25	0.899 (0.863, 0.936)	0.821 (0.696, 0.911)	0.799 (0.747, 0.844)	4.078 (3.132, 5.310)	0.957 (0.922, 0.979)	0.451 (0.352, 0.553)	0.957 (0.922, 0.979)	0.620
26	0.897 (0.859, 0.934)	0.857 (0.738, 0.936)	0.755 (0.700, 0.805)	3.504 (2.777, 4.422)	0.189 (0.099, 0.360)	0.414 (0.323, 0.509)	0.963 (0.929, 0.984)	0.613

* See Table A1 for specifications of empirical formulas. Areas under the receiver operating characteristic curve (AUC), confident interval (CI), positive likelihood ratio (+LR), negative likelihood ratios (−LR), positive predictive values (+PV), negative predictive values (−PV).

Table A6. Prediction performance of the three-parameter mixed-effects logistic model without covariance when combined with the 26 estimated fetal weight empirical formulas* in predicting large for gestational age.

Formulas	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	+LR (95% CI)	−LR (95% CI)	+PV (95% CI)	−PV (95% CI)	Youden's Index
1	0.865 (0.810, 0.919)	0.424 (0.255, 0.608)	0.930 (0.895, 0.956)	6.081 (3.429, 10.783)	0.612 (0.461, 0.831)	0.400 (0.239, 0.579)	0.936 (0.903, 0.961)	0.354
2	0.867 (0.812, 0.922)	0.455 (0.281, 0.636)	0.940 (0.907, 0.964)	7.601 (4.241, 13.622)	0.580 (0.424, 0.793)	0.454 (0.281, 0.636)	0.940 (0.907, 0.964)	0.395
3	0.867 (0.811, 0.923)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
4	0.868 (0.814, 0.922)	0.455 (0.281, 0.636)	0.940 (0.907, 0.964)	7.601 (4.241, 13.622)	0.580 (0.424, 0.793)	0.454 (0.281, 0.636)	0.940 (0.907, 0.964)	0.395
5	0.868 (0.812, 0.923)	0.455 (0.281, 0.636)	0.944 (0.911, 0.967)	8.048 (4.443, 14.578)	0.578 (0.423, 0.790)	0.469 (0.291, 0.653)	0.940 (0.907, 0.964)	0.398
6	0.873 (0.818, 0.927)	0.515 (0.335, 0.692)	0.937 (0.903, 0.962)	8.161 (4.724, 14.100)	0.518 (0.364, 0.737)	0.472 (0.304, 0.645)	0.946 (0.914, 0.969)	0.452
7	0.891 (0.839, 0.944)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
8	0.876 (0.820, 0.932)	0.485 (0.308, 0.665)	0.944 (0.911, 0.967)	8.585 (4.805, 15.339)	0.546 (0.392, 0.761)	0.485 (0.308, 0.665)	0.944 (0.911, 0.967)	0.428
9	0.894 (0.854, 0.934)	0.455 (0.281, 0.636)	0.934 (0.899, 0.959)	6.841 (3.889, 12.034)	0.584 (0.427, 0.780)	0.429 (0.263, 0.606)	0.940 (0.907, 0.964)	0.388
10	0.886 (0.846, 0.927)	0.485 (0.308, 0.665)	0.930 (0.895, 0.956)	6.949 (4.041, 11.950)	0.554 (0.397, 0.772)	0.432 (0.271, 0.605)	0.943 (0.910, 0.966)	0.415
11	0.895 (0.856, 0.935)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
12	0.891 (0.849, 0.933)	0.515 (0.335, 0.692)	0.950 (0.919, 0.972)	10.337 (5.707, 18.724)	0.510 (0.359, 0.726)	0.531 (0.347, 0.709)	0.947 (0.915, 0.969)	0.465
13	0.890 (0.849, 0.932)	0.485 (0.308, 0.665)	0.944 (0.911, 0.967)	8.585 (4.805, 15.339)	0.546 (0.392, 0.761)	0.485 (0.308, 0.665)	0.944 (0.911, 0.967)	0.428
14	0.886 (0.845, 0.926)	0.545 (0.364, 0.719)	0.944 (0.911, 0.967)	9.658 (5.533, 16.856)	0.482 (0.331, 0.701)	0.514 (0.340, 0.686)	0.950 (0.919, 0.972)	0.489
15	0.892 (0.852, 0.931)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
16	0.883 (0.842, 0.924)	0.485 (0.308, 0.665)	0.937 (0.903, 0.962)	7.681 (4.389, 13.441)	0.550 (0.394, 0.767)	0.457 (0.288, 0.634)	0.943 (0.911, 0.967)	0.422
17	0.861 (0.803, 0.919)	0.455 (0.281, 0.636)	0.934 (0.899, 0.959)	6.841 (3.889, 12.034)	0.584 (0.427, 0.780)	0.429 (0.263, 0.606)	0.940 (0.907, 0.964)	0.388
18	0.871 (0.818, 0.924)	0.515 (0.335, 0.692)	0.937 (0.903, 0.962)	8.161 (4.724, 14.100)	0.518 (0.364, 0.737)	0.472 (0.304, 0.645)	0.946 (0.914, 0.969)	0.452
19	0.887 (0.841, 0.933)	0.515 (0.335, 0.692)	0.937 (0.903, 0.962)	8.161 (4.724, 14.100)	0.518 (0.364, 0.737)	0.472 (0.304, 0.645)	0.946 (0.914, 0.969)	0.452
20	0.892 (0.853, 0.932)	0.545 (0.364, 0.719)	0.944 (0.911, 0.967)	9.658 (5.533, 16.856)	0.482 (0.331, 0.701)	0.514 (0.340, 0.686)	0.950 (0.919, 0.972)	0.489
21	0.890 (0.850, 0.931)	0.424 (0.255, 0.608)	0.937 (0.903, 0.962)	6.721 (3.728, 12.117)	0.615 (0.458, 0.825)	0.424 (0.255, 0.608)	0.937 (0.903, 0.962)	0.361
22	0.884 (0.840, 0.929)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
23	0.882 (0.836, 0.929)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
24	0.887 (0.846, 0.929)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
25	0.889 (0.847, 0.931)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
26	0.885 (0.840, 0.930)	0.515 (0.335, 0.692)	0.940 (0.907, 0.964)	8.614 (4.936, 15.035)	0.516 (0.362, 0.734)	0.456 (0.243, 0.656)	0.946 (0.915, 0.969)	0.455

* See Table A1 for specifications of empirical formulas. Areas under the receiver operating characteristic curve (AUC), confident interval (CI), positive likelihood ratio (+LR), negative likelihood ratios (−LR), positive predictive values (+PV), negative predictive values (−PV).

Table A7. Prediction performance of the three-parameter mixed-effects logistic model with covariance when combined with the 26 estimated fetal weight empirical formulas* in predicting large for gestational age.

Formulas	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	+LR (95% CI)	−LR (95% CI)	+PV (95% CI)	−PV (95% CI)	Youden's index
1	0.868 (0.814, 0.922)	0.455 (0.281, 0.636)	0.934 (0.899, 0.959)	6.841 (3.889, 12.034)	0.584 (0.427, 0.799)	0.429 (0.263, 0.606)	0.940 (0.907, 0.964)	0.388
2	0.867 (0.812, 0.922)	0.455 (0.281, 0.636)	0.940 (0.907, 0.964)	7.601 (4.241, 13.622)	0.580 (0.424, 0.793)	0.454 (0.281, 0.636)	0.940 (0.907, 0.964)	0.395
3	0.867 (0.811, 0.923)	0.515 (0.335, 0.692)	0.940 (0.907, 0.964)	8.614 (4.936, 15.035)	0.516 (0.362, 0.734)	0.456 (0.243, 0.656)	0.946 (0.915, 0.969)	0.455
4	0.868 (0.814, 0.922)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
5	0.868 (0.812, 0.923)	0.455 (0.281, 0.636)	0.940 (0.907, 0.964)	7.601 (4.241, 13.622)	0.580 (0.424, 0.793)	0.454 (0.281, 0.636)	0.940 (0.907, 0.964)	0.395
6	0.873 (0.818, 0.927)	0.545 (0.364, 0.719)	0.944 (0.911, 0.967)	9.658 (5.533, 16.856)	0.482 (0.331, 0.701)	0.514 (0.340, 0.686)	0.950 (0.919, 0.972)	0.489
7	0.891 (0.839, 0.944)	0.455 (0.281, 0.636)	0.940 (0.907, 0.964)	7.601 (4.241, 13.622)	0.580 (0.424, 0.793)	0.454 (0.281, 0.636)	0.940 (0.907, 0.964)	0.395
8	0.876 (0.820, 0.932)	0.485 (0.308, 0.665)	0.947 (0.915, 0.969)	9.121 (5.044, 16.495)	0.544 (0.390, 0.758)	0.500 (0.319, 0.681)	0.944 (0.911, 0.967)	0.432
9	0.894 (0.854, 0.934)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
10	0.886 (0.846, 0.927)	0.455 (0.281, 0.636)	0.934 (0.899, 0.959)	6.841 (3.889, 12.034)	0.584 (0.427, 0.799)	0.429 (0.263, 0.606)	0.940 (0.907, 0.964)	0.388
11	0.895 (0.856, 0.935)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
12	0.891 (0.849, 0.933)	0.515 (0.335, 0.692)	0.950 (0.919, 0.972)	10.337 (5.707, 18.724)	0.510 (0.359, 0.726)	0.465 (0.255, 0.664)	0.947 (0.915, 0.969)	0.465
13	0.890 (0.849, 0.932)	0.455 (0.281, 0.636)	0.940 (0.907, 0.964)	7.601 (4.241, 13.622)	0.580 (0.424, 0.793)	0.454 (0.281, 0.636)	0.940 (0.907, 0.964)	0.395
14	0.886 (0.845, 0.926)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
15	0.892 (0.852, 0.931)	0.455 (0.281, 0.636)	0.944 (0.911, 0.967)	8.048 (4.443, 14.578)	0.578 (0.423, 0.790)	0.469 (0.291, 0.653)	0.940 (0.907, 0.964)	0.398
16	0.883 (0.842, 0.924)	0.485 (0.308, 0.665)	0.937 (0.903, 0.962)	7.681 (4.389, 13.441)	0.550 (0.394, 0.767)	0.457 (0.288, 0.634)	0.943 (0.911, 0.967)	0.422
17	0.861 (0.803, 0.919)	0.424 (0.255, 0.608)	0.937 (0.903, 0.962)	6.721 (4.728, 12.117)	0.615 (0.458, 0.825)	0.424 (0.255, 0.608)	0.937 (0.903, 0.962)	0.361
18	0.871 (0.818, 0.924)	0.455 (0.281, 0.636)	0.934 (0.899, 0.959)	6.841 (3.889, 12.034)	0.584 (0.427, 0.780)	0.429 (0.263, 0.606)	0.940 (0.907, 0.964)	0.388
19	0.887 (0.841, 0.933)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
20	0.892 (0.853, 0.932)	0.515 (0.335, 0.692)	0.944 (0.911, 0.967)	9.121 (5.168, 16.099)	0.514 (0.361, 0.731)	0.500 (0.247, 0.659)	0.947 (0.915, 0.969)	0.459
21	0.890 (0.850, 0.931)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
22	0.884 (0.840, 0.929)	0.545 (0.364, 0.719)	0.944 (0.911, 0.967)	9.658 (5.533, 16.846)	0.482 (0.331, 0.701)	0.514 (0.340, 0.686)	0.950 (0.919, 0.972)	0.489
23	0.882 (0.836, 0.929)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
24	0.887 (0.846, 0.929)	0.424 (0.255, 0.608)	0.934 (0.899, 0.959)	6.385 (3.572, 11.412)	0.617 (0.459, 0.828)	0.412 (0.246, 0.593)	0.937 (0.903, 0.961)	0.358
25	0.889 (0.847, 0.931)	0.485 (0.308, 0.665)	0.937 (0.903, 0.962)	7.681 (4.389, 13.441)	0.550 (0.394, 0.767)	0.457 (0.288, 0.634)	0.943 (0.911, 0.967)	0.422
26	0.885 (0.840, 0.930)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425

* See Table A1 for specifications of empirical formulas. Areas under the receiver operating characteristic curve (AUC), confident interval (CI), positive likelihood ratio (+LR), negative likelihood ratios (−LR), positive predictive values (+PV), negative predictive values (−PV).

Table A8. Prediction performance of the quadratic mixed-effects model without covariance when combined with the 26 estimated fetal weight empirical formulas* in predicting large for gestational age.

Formulas	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	+LR (95% CI)	−LR (95% CI)	+PV (95% CI)	−PV (95% CI)	Youden's Index
1	0.864 (0.810, 0.919)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
2	0.867 (0.812, 0.922)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
3	0.867 (0.811, 0.923)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
4	0.868 (0.814, 0.922)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
5	0.868 (0.812, 0.923)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
6	0.873 (0.818, 0.927)	0.576 (0.392, 0.745)	0.950 (0.919, 0.972)	11.554 (6.510, 20.505)	0.446 (0.300, 0.665)	0.559 (0.379, 0.728)	0.953 (0.923, 0.974)	0.526
7	0.891 (0.839, 0.944)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
8	0.876 (0.820, 0.932)	0.545 (0.364, 0.719)	0.947 (0.915, 0.969)	10.261 (5.806, 18.136)	0.480 (0.330, 0.698)	0.529 (0.351, 0.702)	0.950 (0.919, 0.972)	0.492
9	0.894 (0.854, 0.934)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
10	0.886 (0.846, 0.927)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
11	0.895 (0.856, 0.935)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
12	0.891 (0.849, 0.933)	0.515 (0.335, 0.692)	0.944 (0.911, 0.967)	9.121 (5.168, 16.099)	0.514 (0.361, 0.731)	0.500 (0.247, 0.659)	0.947 (0.915, 0.969)	0.459
13	0.890 (0.849, 0.932)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
14	0.886 (0.845, 0.926)	0.424 (0.255, 0.608)	0.930 (0.895, 0.956)	6.081 (3.429, 10.783)	0.619 (0.461, 0.831)	0.400 (0.239, 0.579)	0.936 (0.903, 0.961)	0.354
15	0.892 (0.852, 0.931)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
16	0.883 (0.842, 0.924)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
17	0.861 (0.803, 0.919)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
18	0.871 (0.818, 0.924)	0.485 (0.308, 0.665)	0.943 (0.911, 0.967)	8.585 (4.805, 15.339)	0.546 (0.392, 0.761)	0.485 (0.308, 0.665)	0.943 (0.911, 0.967)	0.428
19	0.887 (0.841, 0.933)	0.515 (0.335, 0.692)	0.944 (0.911, 0.967)	9.121 (5.168, 16.099)	0.514 (0.361, 0.731)	0.500 (0.247, 0.659)	0.947 (0.915, 0.969)	0.459
20	0.892 (0.853, 0.932)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
21	0.890 (0.850, 0.931)	0.485 (0.308, 0.665)	0.937 (0.903, 0.962)	7.681 (4.389, 13.441)	0.550 (0.394, 0.767)	0.457 (0.288, 0.634)	0.943 (0.911, 0.967)	0.422
22	0.884 (0.840, 0.929)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
23	0.882 (0.836, 0.929)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
24	0.887 (0.846, 0.929)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
25	0.889 (0.847, 0.931)	0.515 (0.335, 0.692)	0.940 (0.907, 0.964)	8.614 (4.936, 15.035)	0.516 (0.362, 0.734)	0.456 (0.243, 0.656)	0.946 (0.915, 0.969)	0.455
26	0.885 (0.840, 0.930)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425

* See Table A1 for specifications of empirical formulas. Areas under the receiver operating characteristic curve (AUC), confident interval (CI), positive likelihood ratio (+LR), negative likelihood ratios (−LR), positive predictive values (+PV), negative predictive values (−PV).

Table A9. Prediction performance of the quadratic mixed-effects model with covariance when combined with the 26 estimated fetal weight empirical formulas* in predicting large for gestational age.

Formulas	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	+LR (95% CI)	−LR (95% CI)	+PV (95% CI)	−PV (95% CI)	Youden's index
1	0.865 (0.810, 0.919)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
2	0.867 (0.812, 0.922)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
3	0.867 (0.811, 0.923)	0.455 (0.281, 0.636)	0.934 (0.899, 0.959)	6.841 (3.889, 12.034)	0.584 (0.427, 0.799)	0.429 (0.263, 0.606)	0.940 (0.907, 0.964)	0.388
4	0.868 (0.814, 0.922)	0.485 (0.308, 0.665)	0.937 (0.903, 0.962)	7.681 (4.389, 13.441)	0.550 (0.394, 0.767)	0.457 (0.288, 0.634)	0.943 (0.911, 0.967)	0.422
5	0.868 (0.812, 0.923)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
6	0.873 (0.818, 0.927)	0.576 (0.392, 0.745)	0.950 (0.919, 0.972)	11.554 (6.510, 20.505)	0.446 (0.300, 0.665)	0.559 (0.379, 0.728)	0.953 (0.923, 0.974)	0.526
7	0.891 (0.839, 0.944)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
8	0.876 (0.820, 0.932)	0.515 (0.335, 0.692)	0.947 (0.915, 0.969)	9.691 (5.424, 17.316)	0.512 (0.360, 0.727)	0.515 (0.335, 0.692)	0.947 (0.915, 0.969)	0.462
9	0.894 (0.854, 0.934)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
10	0.886 (0.846, 0.927)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
11	0.895 (0.856, 0.935)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
12	0.891 (0.849, 0.933)	0.515 (0.335, 0.692)	0.944 (0.911, 0.967)	9.121 (5.168, 16.099)	0.514 (0.361, 0.731)	0.500 (0.247, 0.659)	0.947 (0.915, 0.969)	0.459
13	0.890 (0.849, 0.932)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
14	0.886 (0.845, 0.926)	0.424 (0.255, 0.608)	0.930 (0.895, 0.956)	6.081 (3.429, 10.783)	0.619 (0.461, 0.831)	0.400 (0.239, 0.579)	0.936 (0.903, 0.961)	0.354
15	0.892 (0.852, 0.931)	0.455 (0.281, 0.636)	0.940 (0.907, 0.964)	7.601 (4.241, 13.622)	0.580 (0.424, 0.793)	0.454 (0.281, 0.636)	0.940 (0.907, 0.964)	0.395
16	0.883 (0.842, 0.924)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
17	0.861 (0.803, 0.919)	0.455 (0.281, 0.636)	0.937 (0.903, 0.962)	7.201 (4.057, 12.780)	0.582 (0.426, 0.796)	0.441 (0.272, 0.621)	0.940 (0.907, 0.964)	0.391
18	0.871 (0.818, 0.924)	0.485 (0.308, 0.665)	0.944 (0.911, 0.967)	8.585 (4.805, 15.339)	0.546 (0.392, 0.761)	0.485 (0.308, 0.665)	0.944 (0.911, 0.967)	0.428
19	0.887 (0.841, 0.933)	0.515 (0.335, 0.692)	0.944 (0.911, 0.967)	9.121 (5.168, 16.099)	0.514 (0.361, 0.731)	0.500 (0.247, 0.659)	0.947 (0.915, 0.969)	0.459
20	0.892 (0.853, 0.932)	0.455 (0.281, 0.636)	0.940 (0.907, 0.964)	7.601 (4.241, 13.622)	0.580 (0.424, 0.793)	0.454 (0.281, 0.636)	0.940 (0.907, 0.964)	0.395
21	0.890 (0.850, 0.931)	0.485 (0.308, 0.665)	0.937 (0.903, 0.962)	7.681 (4.389, 13.441)	0.550 (0.394, 0.767)	0.457 (0.288, 0.634)	0.943 (0.911, 0.967)	0.422
22	0.884 (0.840, 0.929)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
23	0.882 (0.836, 0.929)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
24	0.887 (0.846, 0.929)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425
25	0.889 (0.847, 0.931)	0.515 (0.335, 0.692)	0.944 (0.911, 0.967)	9.121 (5.168, 16.099)	0.514 (0.361, 0.731)	0.500 (0.247, 0.659)	0.947 (0.915, 0.969)	0.459
26	0.885 (0.840, 0.930)	0.485 (0.308, 0.665)	0.940 (0.907, 0.964)	8.108 (4.587, 14.330)	0.548 (0.393, 0.764)	0.471 (0.298, 0.649)	0.943 (0.911, 0.967)	0.425

* See Table A1 for specifications of empirical formulas. Areas under the receiver operating characteristic curve (AUC), confident interval (CI), positive likelihood ratio (+LR), negative likelihood ratios (−LR), positive.

References

1. Esakoff, T.F.; Cheng, Y.W.; Sparks, T.N.; Caughey, A.B. The association between birthweight 4000 g or greater and perinatal outcomes in patients with and without gestational diabetes mellitus. *Am. J. Obstet. Gynecol.* **2009**, *200*, e1–e4. [CrossRef] [PubMed]
2. Boulet, S.L.; Alexander, G.R.; Salihu, H.M.; Pass, M. Macrosomic births in the United States: Determinants, outcomes, and proposed grades of risk. *Am. J. Obstet. Gynecol.* **2003**, *188*, 1372–1378. [CrossRef] [PubMed]
3. Meshari, A.A.; De Silva, S.; Rahman, I. Fetal macrosomia—maternal risks and fetal outcome. *Int. J. Gynaecol. Obstet.* **1990**, *32*, 215–222. [CrossRef]
4. Zhang, J.; Kim, S.; Grewal, J.; Albert, P.S. Predicting large fetuses at birth: Do multiple ultrasound examinations and longitudinal statistical modelling improve prediction? *Paediatr. Perinat. Epidemiol.* **2012**, *26*, 199–207. [CrossRef]
5. Hadlock, F.P.; Harrist, R.B.; Carpenter, R.J.; Deter, R.L.; Park, S.K. Sonographic estimation of fetal weight. The value of femur length in addition to head and abdomen measurements. *Radiology* **1984**, *150*, 535–540. [CrossRef]
6. Liang, K.Y.; Zeger, S.L. Longitudinal data analysis using generalized linear models. *Biometrika* **1986**, *73*, 13–22. [CrossRef]
7. Melamed, N.; Yogev, Y.; Meizner, I.; Mashiach, R.; Bardin, R.; Ben-Haroush, A. Sonographic fetal weight estimation: Which model should be used? *J. Ultrasound Med.* **2009**, *28*, 617–629. [CrossRef]
8. Siemer, J.; Egger, N.; Hart, N.; Meurer, B.; Müller, A.; Dathe, O.; Goecke, T.; Schild, R.L. Fetal weight estimation by ultrasound: Comparison of 11 different formulae and examiners with differing skill levels. *Ultraschall Med.* **2008**, *29*, 159–164. [CrossRef]
9. Aviram, A.; Yogev, Y.; Ashwal, E.; Hiersch, L.; Hadar, E.; Gabbay-Benziv, R. Prediction of large for gestational age by various sonographic fetal weight estimation formulas - Which should we use? *J. Perinatol.* **2017**, *37*, 513–517. [CrossRef]
10. Pinheiro, J.C.; Bates, D.M. *Mixed Effects Models in S and S-Plus*, 1st ed.; Springer: New York, NY, USA, 2000; pp. 217–304.
11. Bakketeig, L.S.; Jacobsen, G.; Hoffman, H.J.; Lindmark, G.; Bergsjø, P.; Molne, K.; Rødsten, J. Pre-pregnancy risk factors for small-for-gestational age births among parous women in Scandinavia. *Acta. Obstet. Gynecol. Scand.* **1993**, *72*, 273–279. [CrossRef]
12. Loughna, P.; Chitty, L.; Evans, T.; Chudleigh, T. Fetal Size and Dating: Charts Recommended for Clinical Obstetric Practice. *Ultrasound* **2009**, *17*, 161–167. [CrossRef]
13. Davidian, M.; Giltinan, D.M. Nonlinear models for repeated measurement data: An overview and update. *J. Agric. Biol. Environ. Stat.* **2003**, *8*, 387. [CrossRef]
14. Zhou, Z.H. *Ensemble Methods: Foundations and Algorithms*, 1st ed.; Chapman & Hall: Boca Raton, FL, USA, 2012; pp. 67–98.
15. Ruta, D.; Gabrys, B. Classifier selection for majority voting. *Inf. Fusion* **2005**, *6*, 63–81. [CrossRef]
16. Ruta, D.; Gabrys, B. Analysis of the Correlation Between Majority Voting Error and the Diversity Measures in Multiple Classifier Systems. In Proceedings of the Soft Computing and Intelligent Systems for Industry, Proceedings of Fourth International ICSC Symposium, Paisley, Scotland, 26–29 June 2001; p. 2001.
17. Windeatt, T. Diversity measures for multiple classifier system analysis and design. *Inf. Fusion* **2005**, *6*, 21–36. [CrossRef]
18. Tibshirani, R. Regression Selection and Shrinkage via the Lasso. *J. R. Stat. Soc. B* **1996**, *58*, 267–288. [CrossRef]
19. Fan, J.; Li, R. Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Am. Stat. Assoc.* **2001**, *96*, 1348–1360. [CrossRef]
20. Zhang, C.H. Nearly unbiased variable selection under minimax concave penalty. *Ann. Stat.* **2010**, *38*, 894–942. [CrossRef]
21. Wolpert, D.H. Stacked generalization. *Neural Netw.* **1992**, *5*, 241–259. [CrossRef]
22. Breiman, L. Stacked regressions. *Mach. Learn.* **1996**, *24*, 49–64. [CrossRef]
23. van der Laan, M.J.; Polley, E.C.; Hubbard, A.E. Super Learner. *Stat. Appl. Genet. Mol. Biol.* **2007**, *6*, e25. [CrossRef]
24. Debray, T.P.A.; Koffijberg, H.; Nieboer, D.; Vergouwe, Y.; Steyerberg, E.W.; Moons, K.G.M. Meta-analysis and aggregation of multiple published prediction models. *Stat. Med.* **2014**, *33*, 2341–2362. [CrossRef] [PubMed]

25. Campbel, S.; Wilkin, D. Ultrasonic measurement of fetal abdomen circumference in the estimation of fetal weight. *Br. J. Obstet. Gynaecol.* **1975**, *82*, 689–697. [CrossRef] [PubMed]
26. Jordaan, H. Estimation of fetal weight by ultrasound. *J. Clin. Ultrasound.* **1983**, *11*, 59–66. [CrossRef] [PubMed]
27. Warsof, S.L.; Gohari, P.; Berkowitz, R.L.; Hobbins, J.C. The estimation of fetal weight by computer-assisted analysis. *Am. J. Obstet. Gynecol.* **1977**, *128*, 881–892. [CrossRef]
28. Higginbottom, J.; Slater, J.; Porter, G.; Whitfield, C.R. Estimation of fetal weight from ultrasound measurement trunk circumference. *Br. J. Obstet. Gynaecol.* **1975**, *82*, 698–701. [CrossRef] [PubMed]
29. Woo, J.S.; Wan, C.W.; Cho, K.M. Computer-assisted evaluation of ultrasonic fetal weight prediction using multiple regression equations with and without the fetal femur length. *J. Ultrasound Med.* **1985**, *4*, 65–67. [CrossRef]
30. Warsof, S.L.; Wolf, P.; Coulehan, J.; Queenan, J.T. Comparison of fetal weight estimation formulas with and without head measurements. *Obstet. Gynecol.* **1986**, *67*, 569–573.
31. Vintzileos, A.M.; Campbell, W.A.; Rodis, J.F.; Bors-Koefoed, R.; Nochimson, D.J. Fetal weight estimation formulas with head, abdominal, femur, and thigh circumference measurements. *Am. J. Obstet. Gynecol.* **1987**, *157*, 410–414. [CrossRef]
32. Shepard, M.J.; Richards, V.A.; Berkowitz, R.L.; Warsof, S.L.; Hobbins, J.C. An evaluation of two equations for predicting fetal weight by ultrasound. *Am. J. Obstet. Gynecol.* **1982**, *142*, 47–54. [CrossRef]
33. Mirghani, H.M.; Weerasinghe, S.; Ezimokhai, M.; Smith, J.R. Ultrasonic estimation of fetal weight at term: An evaluation of eight formulae. *J. Obstet. Gynaecol. Res.* **2005**, *31*, 409–413. [CrossRef]
34. Shinozuka, N.; Okai, T.; Kohzuma, S.; Mukubo, M.; Shih, C.T.; Maeda, T.; Kuwabara, Y.; Mizuno, M. Formulas for fetal weight estimation by ultrasound measurements based on neonatal specific gravities and volumes. *Am. J. Obstet. Gynecol.* **1987**, *157*, 1140–1145. [CrossRef]
35. Combs, C.A.; Jaekle, R.K.; Rosenn, B.; Pope, M.; Miodovnik, M.; Siddiqi, T.A. Sonographic estimation of fetal weight based on a model of fetal volume. *Obstet. Gynecol.* **1993**, *82*, 365–370. [PubMed]
36. Ott, W.J.; Doyle, S.; Flamm, S.; Wittman, J. Accurate Ultrasonic Estimation of Fetal Weight: Prospective Analysis of New Ultrasonic Formulas. *Am. J. Perinatol.* **1986**, *3*, 307–310. [CrossRef] [PubMed]

Journal of
Clinical Medicine

MDPI

Article

Seasonal Variation in Physical Activity among Preoperative Patients with Lung Cancer Determined Using a Wearable Device

Sunga Kong [1,2,†], Hye Yun Park [3,†], Danbee Kang [1,4], Jae Kyung Lee [2], Genehee Lee [2], O Jung Kwon [3], Young Mog Shim [5], Jae Ill Zo [5,*] and Juhee Cho [1,4,6,*]

1 Department of Clinical Research Design and Evaluation, SAIHST, Sungkyunkwan University, Seoul 06351, Korea; sunga00kong@gmail.com (S.K.); dbee.kang@gmail.com (D.K.)
2 Patient-Centered Outcomes Research Institute, Samsung Medical Center, Seoul 06351, Korea; roemgirls1230@gmail.com (J.K.L.); genehee@gmail.com (G.L.)
3 Division of Pulmonary and Critical Care Medicine, Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul 06351, Korea; hyeyunpark@skku.edu (H.Y.P.); ojung.kwon@samsung.com (O.J.K.)
4 Center for Clinical Epidemiology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul 06351, Korea
5 Department of Thoracic and Cardiovascular Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul 06351, Korea; youngmog.shim@samsung.com
6 Departments of Epidemiology and Health, Behavior and Society, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA
* Correspondence: jayl.zo@samsung.com (J.I.Z.); jcho@skku.edu (J.C.); Tel.: +82-2-3410-3485 (J.I.Z.); +82-2-3410-1448 (J.C.); Fax: +82-2-2226-3296 (J.I.Z.); +82-2-3410-6639 (J.C.)
† S.K. and H.Y.P. contributed equally as joint first authors.

Received: 22 December 2019; Accepted: 22 January 2020; Published: 27 January 2020

Abstract: We aim to examine how season and temperature levels affect physical activity using a wearable device among patients scheduled to undergo surgical resection of lung cancer. Physical activity (PA) data from the wearable device were analyzed by seasons for 555 preoperative lung cancer patients from the CATCH-LUNG cohort study. The seasons were divided into spring, summer, autumn, and winter using the study enrollment date before surgery. The overall mean (SD) age was 61.1 (8.9) years, and the mean (SD) daily steps at each season were 11,438 (5922), 11,147 (5065), 10,404 (4403), and 8548 (4293), respectively. In the fully-adjusted models, patients in the winter season had 27.04% fewer daily steps (95% CI = −36.68%, −15.93%) and 35.22% less time spent performing moderate to vigorous physical activity (MVPA) compared to patients in the spring. The proportion of participants with over 8000 steps and duration of MVPA were significantly lower in the winter than the spring. In particular, daily steps had a negative linear association with wind chill temperature in patients who lived in Seoul. In conclusion, PA was significantly lower in the winter and it was more robust in patients who had a low cardiorespiratory function.

Keywords: lung cancer; physical activity; season; preoperative; wearable

1. Introduction

Lung cancer is the leading cause of cancer-related death worldwide, contributing to 1.6 million deaths annually [1]. Surgical resection remains the best curative treatment option in patients with early-stage non-small cell lung cancer (NSCLC), and a patient's preoperative status is important to assess the feasibility of undergoing surgical lung resection under general anesthesia. In particular, patients with poor pulmonary function and cardiorespiratory fitness are considered inoperable due

to increased morbidity and mortality after surgical resection [2]. Both cardiopulmonary fitness and functional capacity are widely recognized as strong predictors of postoperative complications, specifically mortality and long-term survival, in NSCLC [3].

Cardiorespiratory fitness (CRF) and functional capacity are affected by physical activity (PA) [4]. Numerous studies have conducted PA or exercise programs to improve physical fitness and functions before thoracic surgery [2,5]. Adopting and sustaining a more physically active lifestyle has been shown to reduce the risk of complications and mortality and improve health-related quality of life in patients who underwent thoracic surgery [6]. However, promoting long-term PA has been challenging due to various factors such as lack of motivation, access to facilities for physical activities, and inclement weather [7]. Studies reported that seasonal weather conditions could promote or deter PA [8]. Furthermore, most studies were conducted with a small number of participants (<50) and only included limited populations such as children [9] or the elderly [10]. In addition, few quantitative assessments have been focused on seasonal variation in PA among preoperative lung cancer patients. Studies have attempted to measure daily PA with quantitative assessments using simple and non-expensive devices during the perioperative periods of lung cancer surgery [11–13]. They found that that daily walk distance predicted maximum oxygen consumption per minute in patients undergoing lung resection [11], and the time and the quality of the daily ambulatory activity of the patients decreased during the first postoperative month [12]. Thus, our study aims to use a wearable device (Fitbit) to examine how the season and temperature level affect PA among patients who are scheduled to undergo surgical resection for lung cancer.

2. Methods

2.1. Subjects and Data Sources

Patients with lung cancer in this study were selected from the Coordinate Approach to Cancer patients' Health for Lung Cancer (CATCH-LUNG) cohort of preoperative lung cancer patients between March 2016 and October 2018 at the Samsung Medical Center in Seoul, Korea. Inclusion criteria for CATCH-LUNG cohort were (1) patients who were expected to undergo curative lung cancer surgery for suspected or histologically confirmed NSCLC, (2) patients who were able to walk and keep a normal life with the Eastern Cooperative Oncology Group Performance Status (ECOG PS ≤1), and (3) patients understood the purpose of this study and agreed to participate in the study. Exclusion criteria were (1) patients who had undergone neoadjuvant treatment before surgery, (2) patients free of NSCLC after pathological exams, (3) patients whose surgery was canceled, or (4) patients who withdrew consent before baseline data collection. Among patients who met these criteria, we furthermore excluded patients who had either pathologically confirmed stage IV cancer after surgery (n = 2) and 63 patients who were excluded due to lack of Fitbit data (28 patients wore their Fitbit less than one day and 35 patients had either hardware or software failures of Fitbit). The final study sample included 555 patients. The study protocol was approved by the Institutional Review Board of Samsung Medical Center (no. 2015-11-025). Written informed consent was obtained from all participants.

2.2. Grouping and Weather Data Collection

The main exposure variable was the season, which was divided into spring (March to May), summer (June to August), autumn (September to November), and winter (December to February) using the study enrollment date before surgery. For preoperative patients living in the metropolitan Seoul area, the weather-related factor of wind chill temperature was obtained from the Korea Meteorological Administration (https://data.kma.go.kr).

2.3. Physical Activity and other Variable

The main outcome was PA, which was assessed using a reliable wearable activity tracker [14]. The Flex tracker (Fitbit, San Francisco, CA, USA) was used to quantify the PA intensity, time, activity

type, and steps per day. We asked participants to wear a Fitbit activity tracker 24 h per day for 7 consecutive days. We determined that patients did not wear the Fitbit if there was no movement (0 steps) for more than 4 consecutive hours during daytime (9:00 a.m. to 4:00 p.m.). Physical activity time, frequency, and intensity data were automatically measured and saved by an internal sensor. We then calculated activity level by combining the Fitbit data with age, gender, height, and weight data recorded at registration. Activities were classified into four categories: (1) sleeping or sedentary activity (1 MET), (2) light physical activity (1~2.9 METs), (3) moderate physical activity (3~5.9 METs), and (4) vigorous physical activity (more than 6 METs). The calculated values were averaged to define daily activity level.

The tracker was worn on the wrist. To reduce bias, the device had no screen so that patients could not see their recorded level of PA. In addition, there was no specific education or guidelines for patients regarding PA and patients were recommended to maintain PA prior to surgery, as usual.

CRF was measured using the 6-min walk test (6MWT), which was performed according to ATS guidelines [15]. Each participant was asked to walk (not run) back and forth along the corridor as far as possible for 6 min and was given standardized verbal encouragement every minute. The test has been widely used for preoperative and postoperative evaluations of CRF. In some clinical situations, the 6MWT provides a better index of the patient's ability to perform daily activities compared to peak oxygen uptake [16]. In addition, the 6MWT is a sub-maximal test of CRF, as opposed to cardiopulmonary exercise testing (CPET). While it is well known that the incremental shuttle walk test (ISWT) has a higher correlation with CPET than 6MWT, we could not use the ISWT because it was not available in Korea. In fact, 6MWT has been widely used in real-world clinics for cardiopulmonary evaluation in patients with chronic obstructive pulmonary disease (COPD) [17]. To obtain quality data, trained researchers provided study participants detailed instructions about 6MWT, and asked patients to do a pilot walk (for 15~20 s) before the actual test.

Spirometry and DLco measurements were performed using a Vmax 22 respiratory analyzer (SensorMedics, OH, USA) according to the American Thoracic Society/European Respiratory Society criteria [18,19]. Absolute values of forced expiratory volume in 1 s (FEV_1), forced vital capacity (FVC), and DLco were obtained, and the percentage of predicted values (% predicted) for FEV_1, FVC, and DLco were calculated using a representative Korean sample [20,21] as a reference. Sociodemographic and behavioral information, including age, smoking status, and comorbidities, were recorded before surgery using a questionnaire. Treatment information regarding pathological stage and pulmonary function were collected after surgery.

2.4. Statistical Analysis

Continuous and categorical variables were compared among the seasons using analysis of variance (ANOVA) and the $\chi 2$ test, respectively.

For the main analyses, we used linear regression to compare the daily number of steps and the moderate-to-vigorous physical activity (MVPA) time by season. Since steps per day and MVPA minutes are markedly right skewed (p-values based on Shapiro–Wilk and Shapiro–Francia tests for normality were <0.001), we used log-transformed daily number of steps and MVPA time as the outcomes. The average difference (as a percent difference with 95% confidence interval (CI)) was estimated comparing patients enrolled in the summer, autumn, or winter to those patients enrolled in the spring.

Using daily number of steps and MVPA duration, we developed a dichotomized outcome to evaluate the proportion of participants who were physically active. Being physically active was defined as either taking more than 8000 steps per day or performing more than 60 min of MVPA. We chose 8000 steps per day based on a reference that adults usually take 5000 steps per day and perform daily activities such as house errands, walking, or shopping [22]. We then added 3000 steps per day to account for 30 min of MVPA (10 min of MVPA is around 1000 steps [23,24]). From this, we equated 8000 steps to 60 min of MVPA.

Logistic regression was conducted to compare the odds of being physically active by season with adjustments made for age, sex, smoking status, FEV_1% pred, any pulmonary comorbidities (COPD, asthma, or ILD), and any extra-pulmonary comorbidities (cardiovascular diseases or diabetes mellitus). In addition, we performed stratified analyses to evaluate differences in PA associated with the season in prespecified subgroups of age (<65 vs. ≥65 years) [25] and CRF (<500 vs. ≥500 m in 6MWD) [26].

To determine factors leading to the differences in PA according to season, we conducted an additional analysis for patients (n = 85) living in the Seoul metropolitan area because we considered the bias of area environment and topography in Korea. To find an association between weather and daily steps, we modeled wind chill temperature as a continuous variable using restricted cubic splines with knots at the 5th, 35th, 65th, and 95th percentiles of the sample distribution to provide a flexible estimate of the dose-response relationship between weather factors and daily steps.

For all analyses, a p-value of <0.05 was considered statistically significant. All analyses were performed using STATA software, version 14 (Stata Corp LP, College Station, TX, USA).

3. Results

The characteristics of 555 patients are described in Table 1. The mean (SD) age and daily steps of study participants were 61.1 (8.9) years and 10,603 (5200) (range, 425–32,143), respectively. Among the participants, Fitbit data from 30.8% (n = 171), 32.1% (n = 178), 17.8% (n = 99), and 19.3% (n = 107) of the study subjects were collected in the spring, summer, autumn, and winter seasons, respectively. Although patients in the winter season were likely to be younger and have better pulmonary function than patients in other seasons, the clinical characteristics, including sex, BMI, smoking status, comorbidities, pathologic stage, and cardiorespiratory fitness, were not different across the seasons (Table 1).

Table 1. Characteristics of study participants by season (n = 555).

Characteristics	Spring (n = 171)	Summer (n = 178)	Autumn (n = 99)	Winter (n = 107)	p
Mean age	60.6 (8.9)	62.6 (8.3)	60.6 (9.0)	59.9 (9.4)	0.05
Age categories					0.04
<65	118 (69.0)	104 (58.4)	66 (66.7)	79 (73.8)	
≥65	53 (31.0)	74 (41.6)	33 (33.3)	28 (26.2)	
Sex, male	97 (56.7)	101 (56.7)	57 (57.6)	57 (53.3)	0.92
Body mass index, kg/m^2	24.1 (2.9)	24.3 (2.7)	24.1 (2.7)	24.5 (3.2)	0.67
Smoking status					0.32
Never-smoker	78 (45.6)	89 (50.0)	48 (48.5)	57 (53.3)	
Ex-smoker	59 (34.5)	34 (19.1)	27 (27.3)	16 (15.0)	
Current smoker	34 (19.9)	55 (30.9)	24 (24.2)	34 (31.8)	
Marital status					0.64
Married	151 (88.3)	159 (89.3)	91 (91.9)	91 (85.1)	
Single/divorced/widowed	19 (11.1)	19 (10.7)	7 (7.1)	15 (14.0)	
Unknown	1 (0.6)	0	1 (1.0)	1 (0.9)	
Employment status					0.46
Current work	98 (57.3)	81 (45.5)	49 (49.5)	58 (54.2)	
No work	72 (42.1)	96 (53.9)	49 (49.5)	48 (44.9)	
Unknown	1 (0.6)	1 (0.6)	1 (1.0)	1 (0.9)	
Monthly family income					0.19
<$3.000	54 (31.6)	51 (28.7)	23 (23.2)	31 (29.0)	
≥$3.000	87 (50.9)	93 (52.3)	54 (54.6)	66 (61.7)	
Unknown	30 (17.5)	34 (19.1)	22 (22.2)	10 (9.4)	
Comorbidities					
Pulmonary comorbidities					
COPD	40 (23.4)	43 (24.2)	26 (26.3)	19 (17.8)	0.49
Asthma	3 (1.8)	8 (4.5)	1 (1.0)	3 (2.8)	0.34
ILD	1 (0.6)	2 (1.1)	1 (1.0)	2 (1.9)	0.86

Table 1. *Cont.*

Characteristics	Spring (n = 171)	Summer (n = 178)	Autumn (n = 99)	Winter (n = 107)	p
Extra-pulmonary comorbidities					
Hypertension	52 (30.4)	72 (40.5)	28 (28.3)	35 (32.7)	0.12
Diabetes mellitus	22 (12.9)	27 (15.2)	11 (11.1)	7 (6.5)	0.18
Cardiovascular disease	12 (7.0)	20 (11.2)	9 (9.1)	9 (8.4)	0.58
Pathologic stage					0.81
I	128 (74.9)	130 (73.0)	75 (75.8)	75 (70.1)	
II	25 (14.6)	31 (17.4)	13 (13.1)	16 (15.0)	
III	18 (10.5)	17 (9.6)	11 (11.1)	16 (15.0)	
Pulmonary function test					
FVC, L	3.6 (0.9)	3.4 (0.8)	3.7 (0.9)	3.7 (0.8)	0.02
FVC, % predicted	93.4 (11.5)	89.8 (12.9)	94.3 (12.7)	96.4 (12.1)	< 0.01
FEV_1, L	2.7 (0.7)	2.5 (0.6)	2.7 (0.6)	2.8 (0.6)	< 0.01
FEV_1, % predicted	90.0 (13.2)	87.5 (15.5)	90.2 (14.9)	95.0 (13.5)	< 0.01
FEV_1/FVC	73.8 (8.2)	73.6 (8.6)	72.6 (8.6)	75.1 (8.3)	0.21
DLco, %	91.6 (16.0)	89.2 (17.3)	89.7 (15.2)	91.3 (13.7)	0.47
Cardiorespiratory fitness					
6 min walk distance (m)	520.1(85.7)	506.9(89.1)	515.0(66.8)	508.1(80.7)	0.46
6 min walk distance					0.69
Short distance (<500 m)	68 (39.9)	77 (43.3)	37 (37.4)	49 (45.8)	
Long distance (≥500 m)	102 (59.7)	97 (54.5)	60 (60.1)	56 (52.3)	
Unknown	1 (0.6)	4 (2.3)	2 (2.1)	2 (1.9)	

Values are presented as either n (%) or mean (SD). COPD, chronic obstructive pulmonary disease; ILD, interstitial lung disease; FVC, forced vital capacity; FEV_1, forced expiratory volume in 1 s; DLco, diffusing capacity of carbon monoxide.

The mean (SD) daily steps were 11,438 (5922), 11,147 (5065), 10,404 (4403), and 8548 (4293) in patients who participated in this study in the spring, summer, autumn, and winter seasons, respectively. The proportion of patients who had 8000 or more steps per day was 71%, 72%, 71%, and 54% in participants who received surgery in the spring, summer, autumn, and winter seasons, respectively (Figure 1).

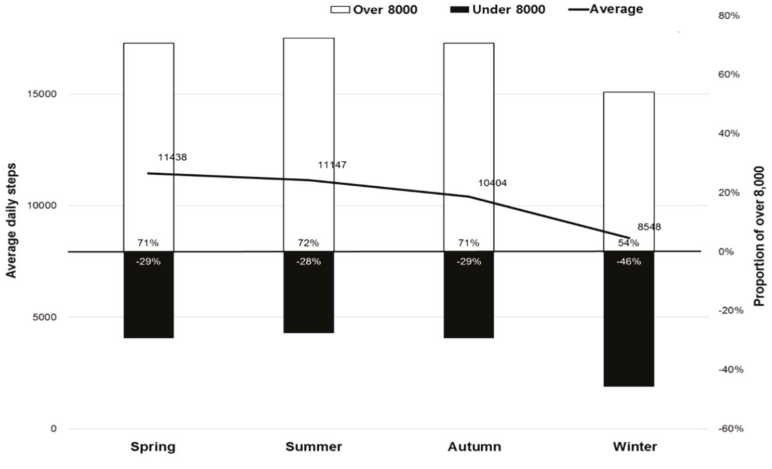


Figure 1. Mean daily steps and the proportion of participants who had more than 8000 steps in each season.

In the fully-adjusted models, patients in the winter season had a significantly lower number of daily steps, with a low of 27.04% (95% CI = −36.68%, −15.93%) compared to the daily steps of patients in the spring season. In comparing the mean (SD) MVPA time, patients in the winter season had the lowest MVPA among subjects (spring, 60.3 (57.2) min/d; summer, 58.5 (46.4) min/d; autumn, 51.0 (35.6) min/d; and winter, 35.0 (36.3) min/d). In the fully-adjusted models, compared to patients in the spring season, the MVPA time was significantly shorter by 35.22% (95% CI = −49.18%, −17.43%) for patients in the winter season. The number of steps and duration of MVPA were significantly lower in the winter season compared to spring season irrespective of age and 6MWD (Table 2).

Table 2. Differences in physical activity by season.

	Spring	Summer	Autumn	Winter
Difference in steps per day (%)				
Overall	Reference	−0.86 (−12.38, 12.18)	−6.24 (−18.86, 8.33)	−27.04 (−36.68, −15.93)
Age				
<65 years	Reference	−8.08 (−21.25, 7.30)	−10.27 (−24.74, 6.98)	−25.21 (−36.65, −11.70)
≥65 years	Reference	11.91 (−9.08, 37.74)	2.36 (−20.65, 32.02)	−32.16 (−48.31, −10.97)
p for interaction		0.14	0.40	0.55
Cardiorespiratory fitness (6MWD)				
<500 m	Reference	−1.21 (−18.1, 19.18)	7.07 (−14.97, 34.83)	−27.63 (−41.48, −10.50)
≥500 m	Reference	−0.75 (−15.53, 16.62)	−12.95 (−27.52, 4.55)	−20.97 (−34.54, −4.59)
p for interaction		0.97	0.17	0.54
Difference in MVPA minutes per day (%)				
Overall	Reference	5.06 (−14.6, 29.24)	−2.11(−23.13, 24.67)	−35.22 (−49.18, −17.43)
Age				
<65 years	Reference	5.13 (−19.26, 36.89)	−1.43 (−26.71, 32.57)	−33.03 (−49.63, −10.95)
≥65 years	Reference	4.05 (−26.37, 47.04)	−3.75 (−37.24, 47.60)	−40.72 (−62.76, −5.62)
p for interaction		0.96	0.93	0.66
Cardiorespiratory fitness (6MWD)				
<500 m	Reference	−4.25 (−30.56, 32.04)	20.57 (−18.5, 78.39)	−38.65 (−57.62, −11.18)
≥500 m	Reference	11.12 (−15.17, 45.55)	−10.57 (−34.3, 21.73)	−26.01 (−46.46, 2.25)
p for interaction		0.48	0.24	0.03

Models were adjusted for age, sex, smoking status, FEV_1% pred, any pulmonary comorbidities, and any extra-pulmonary comorbidities. Pulmonary comorbidities include chronic obstructive pulmonary disease, asthma, or interstitial lung disease, and extra-pulmonary comorbidities include cardiovascular diseases or diabetes mellitus. 6MWD, 6-min walk distance. MVPA, moderate-to-vigorous physical activity.

In the fully-adjusted models, the OR for 8000 or more steps per day was 0.46 (95% CI 0.28, 0.77) for patients in the winter compared to those in the spring. The proportion of patients who had 60 min or more MVPA per day was also similar (spring, 30.4%; summer, 42.7%; autumn, 32.3%; winter, 18.7%). In the fully-adjusted models, the OR for MVPA ≥ 60 min/day was 0.52 (95% CI = 0.29, 0.94) for patients in the winter compared to those in the spring. In particular, the OR for low CRF (<500 m) was 0.30 (95% CI 0.10, 0.87) for patients in the winter compared to those in the spring (Table 3).

Table 3. Odds ratios (95% CI) for physical activity by season.

	Spring	Summer	Autumn	Winter
Steps ≥8000/day				
Overall	Reference	1.11 (0.69, 1.79)	0.98 (0.57, 1.70)	0.46 (0.28, 0.77)
Age				
<65 years	Reference	0.79 (0.43, 1.44)	0.73 (0.37, 1.45)	0.38 (0.20, 0.71)
≥65 years	Reference	2.00 (0.93, 4.31)	1.75 (0.67, 4.56)	0.67 (0.26, 1.74)
p for interaction		0.06	0.15	0.32
Cardiorespiratory fitness (6MWD)				
<500 m	Reference	0.94 (0.48, 1.82)	1.75 (0.73, 4.20)	0.45 (0.21, 0.96)
≥500 m	Reference	1.35 (0.67, 2.74)	0.74 (0.36, 1.55)	0.58 (0.28, 1.22)
p for interaction		0.46	0.14	0.63
MVPA ≥60 min/day				
Overall	Reference	1.73 (1.10, 2.71)	1.07 (0.62, 1.83)	0.52 (0.29, 0.94)
Age				
<65 years	Reference	1.59 (0.91, 2.79)	1.19 (0.62, 2.27)	0.59 (0.30, 1.15)
≥65 years	Reference	2.11 (0.96, 4.61)	0.92 (0.34, 2.50)	0.29 (0.07, 1.13)
P for interaction		0.50	0.72	0.42
Cardiorespiratory fitness (6MWD)				
<500 m	Reference	1.26 (0.61, 2.58)	1.24 (0.52, 2.98)	0.30 (0.10, 0.87)
≥500 m	Reference	1.97 (1.10, 3.55)	1.03 (0.52, 2.05)	0.77 (0.37, 1.61)
p for interaction		0.34	0.75	0.15

Models were adjusted for age, sex, smoking status, FEV_1% pred, any pulmonary comorbidities, and any extra-pulmonary comorbidities. Pulmonary comorbidities include chronic obstructive pulmonary disease, asthma, or interstitial lung disease, and extra-pulmonary comorbidities include cardiovascular diseases or diabetes mellitus. 6MWD, 6-min walk distance. MVPA, moderate-to-vigorous physical activity.

In the spline regression models, the relationship between wind chill temperature and daily steps was linear below 25° wind chill temperature, with lower daily steps corresponding to colder temperatures (Figure 2). The association between winter and the decline in PA was consistent across all the subgroups.

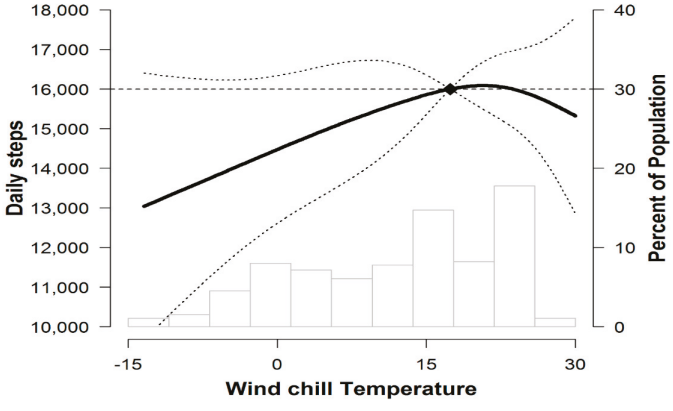


Figure 2. Mean daily steps (95% CI) by wind chill temperature. The curves represent daily step (solid line) and their 95% confidence intervals (dashed lines) based on restricted cubic splines for wind chill temperature with knots at the 5th, 35th, 65th, and 95th percentiles of their sample distributions. The reference value (diamond dot) was set at the 90th percentile.

4. Discussion

We found that preoperative PA was significantly lower in the winter season compared to other seasons, irrespective of age and CRF. In particular, the low MVPA in the winter season was more robust in patients who had low CRF, but the effect was not statistically significant. In terms of weather factors, wind chill temperature was inversely associated with PA. These results demonstrate that the PA of preoperative lung cancer patients is significantly affected by the season and temperature levels.

Our study showed that daily step was lower by 27% in the winter season compared with the spring season. These findings are consistent with previous studies that showed the association between daily PA and the season [8,27–30]. Those studies observed that PA decreased during the winter season in healthy adults [8] and elderly subjects [29]. This phenomenon was profoundly observed in patients with lung and heart diseases. In COPD patients, the daily step count reduced by 43.3 steps/day/°C in the winter season [27], and in heart failure patients, there was a significant seasonal variation in activity between summer and winter ranging from 13.78% to 20.69% [30]. In addition, the odds of walking more than 8000 steps or having more than 60 min MVPA per day in winter is around 0.5 compared to those in spring. It means that only about 50% of patients in the winter season had similar levels of PA as those of patients in spring time. It would be important for clinicians and researchers to evaluate PA in different seasons, especially in winter and provide more tailored educations for PA depending on season. In fact, the cold and snow of winter weather can significantly affect participation in outdoor activity. In particular, influenza during the winter season is an important contributor to the winter burden among older adults and lung-disease patients [31]. Furthermore, slippery roads and a high risk of falling can affect the decline in PA [32]. When outdoor activities are complicated by weather conditions, indoor activity equipment such as treadmills and stationary bicycles offer a helpful alternative. A recent study found that preoperative exercise was effective in reducing postoperative complications and length of hospital stay in patients with lung cancer [6], suggesting the importance of maintaining PA before lung cancer surgery. Another study found that that pulmonary rehabilitation programs using treadmills and bicycles were a valid preoperative strategy to improve physical performance in patients with both NSCLC and COPD [33]. It would be worthwhile to develop a PA intervention program for winter and evaluate its impact on postoperative pulmonary complications.

We found that a few patients achieving MVPA ≥ 60 min/day were more robust in patients with low CRF (≤500 m of 6MWD) compared to those with high CRF (>500 m of 6MWD). Lower exercise tolerance assessed through the 6MWD is associated with poorer postoperative outcomes after lung resection [34]. However, a 7-day intensive preoperative program of PA combined with inspiratory muscle training increased the 6MWD in lung cancer patients in a previous study [35]. Increasing PA has positive effects on CRF and muscle power, such as health-related physical function, leading to improved postoperative recovery in early-stage NSCLC patients. In a previous study, structured and planned preoperative PA reduced postoperative complications (from 45% to 67%) and the length of hospital stay (by 4–5 days) in patients with lung cancer [3,6], suggesting that preoperative PA is an effective strategy for lung cancer patients who will undergo surgical resection. Thus, a strategic PA program should be recommended to maintain and promote PA even in the winter season, especially for low-CRF patients and older adults, for whom it is feasible to receive surgery for lung cancer.

However, seasons are a crude measure when it comes to understanding the effects of weather on PA [36]. Previous studies have observed the effects of several factors related to the season on daily activities. In a healthy population, there was a 2.9% decrease in steps per day for every 10 °C drop in temperature [37]. In addition to temperature, day length and daylight hours account for 73% of the monthly differences in daily activity level, and these three parameters are independent predictors of daily activities [38]. To analyze the climate conditions affecting PA in our study, we conducted a subgroup analysis to identify the relationship between climate conditions and activity for 85 people living in the capital city of South Korea. We found that wind chill temperature significantly affected the steps per day of preoperative patients. In the spline regression models, the relationship between wind chill temperature and daily steps was linear below 25 °C wind chill temperature, with lower daily steps

corresponding to colder temperatures. This is similar to the previous findings [30,39]. The wind chill temperature used in our study is likely to be more closely related to patients' PA behavior compared to absolute temperature, as the wind chill temperature is a quantitative measure of the heat or cold that the human body feels, calculated based on air temperature and wind velocity [40]. While it is not possible to change the weather conditions, a better understanding of how weather influences PA might be helpful when developing strategies to ameliorate the impact of adverse weather conditions on future PA interventions in low-CRF patients and older adults.

There are several limitations to this study. Firstly, this was a cross-sectional study and we were not able to observe the change of PA of the same patient group in different seasons. In addition, due to enrollment periods, a limited number of patients were observed in the winter season. These factors limited our ability to observe the magnitude of seasonal variation in PA among patients. Despite these limitations, our study indicates little variability in PA by season. In fact, patients in the winter season were younger and had better lung function than patients in other season and we still see the different levels of PA depending on season. Secondly, patients who were motivated to perform PA might be more likely to participate in the study. The mean daily step figures were higher in this study compared to data reported in earlier studies [11]. This might be due to our recruitment of patients with ECOG 0 or 1 as they would be healthier and more physically active than general cancer patients. However, mean 6MWD was similar to that reported in the previous study [26], suggesting our study participants had similar CRF to other lung cancer patients. Thirdly, the physical activity device used only counts the steps and the time spent on all those steps, but it does not differentiate whether the steps are outside or inside the house. Therefore we have a limitation in assuming that patients in the winter season were less active due to decreased outdoor PA. Lastly, the results of this study might not be generalizable as it was conducted with patients at a tertiary cancer center in Seoul, Korea. Additional studies would be necessary to confirm the findings with different populations and in different settings.

5. Conclusions

In this study, we found that season and temperature levels affect PA among preoperative lung cancer patients. Patients were much less physically active in the winter season than other seasons and patients in the winter season had lower cardiorespiratory function. Health professionals need to be aware of these seasonal differences and recommend indoor physical activities that preoperative patients with lung cancer can do in winter.

Author Contributions: Conceptualization: S.K., H.Y.P., J.I.Z., and J.C.; study design: S.K., H.Y.P., J.I.Z., and J.C.; data collection: S.K., J.K.L., G.L., O.J.K., Y.M.S., and J.I.Z.; data and statistical analysis: S.K., D.K., and J.C.; drafting the manuscript: S.K., H.Y.P., J.I.Z., and J.C.; revising the manuscript: S.K., H.Y.P., D.K., J.K.L., G.L., O.J.K., Y.M.S., J.I.Z., and J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) [No. 2015R1C1A2A01055805 and 2017R1A2B2006435].

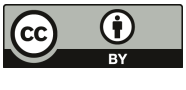
Conflicts of Interest: The authors declare no conflict of interest

References

1. Fitzmaurice, C.; Allen, C.; Barber, R.M.; Barregard, L.; Bhutta, Z.A.; Brenner, H.; Dicker, D.J.; Chimed-Orchir, O.; Dandona, R.; Dandona, L.; et al. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-years for 32 Cancer Groups, 1990 to 2015: A Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol.* **2017**, *3*, 524–548. [PubMed]
2. Sebio Garcia, R.; Yanez Brage, M.I.; Gimenez Moolhuyzen, E.; Granger, C.L.; Denehy, L. Functional and postoperative outcomes after preoperative exercise training in patients with lung cancer: A systematic review and meta-analysis. Interact. Cardiovasc. *Thorac. Surg.* **2016**, *23*, 486–497. [CrossRef] [PubMed]
3. Bade, B.C.; Thomas, D.D.; Scott, J.B.; Silvestri, G.A. Increasing physical activity and exercise in lung cancer: Reviewing safety, benefits, and application. *J. Thorac. Oncol.* **2015**, *10*, 861–871. [CrossRef]

4. Vainshelboim, B.; Muller, J.; Lima, R.M.; Nead, K.T.; Chester, C.; Chan, K.; Kokkinos, P.; Myers, J. Cardiorespiratory fitness, physical activity and cancer mortality in men. *Prev. Med.* **2017**, *100*, 89–94. [CrossRef]
5. Kehler, D.S.; Stammers, A.N.; Tangri, N.; Hiebert, B.; Fransoo, R.; Schultz, A.S.H.; Macdonald, K.; Giacomontonio, N.; Hassan, A.; Legare, J.F.; et al. Systematic review of preoperative physical activity and its impact on postcardiac surgical outcomes. *BMJ Open* **2017**, *7*, e015712. [CrossRef]
6. Steffens, D.; Beckenkamp, P.R.; Hancock, M.; Solomon, M.; Young, J. Preoperative exercise halves the postoperative complication rate in patients with lung cancer: A systematic review of the effect of exercise on complications, length of stay and quality of life in patients with cancer. *Br. J. Sports Med.* **2018**, *52*, 344. [CrossRef]
7. Chen, Y.M. Perceived barriers to physical activity among older adults residing in long-term care institutions. *J. Clin. Nurs.* **2010**, *19*, 432–439. [CrossRef] [PubMed]
8. Tucker, P.; Gilliland, J. The effect of season and weather on physical activity: A systematic review. *Public Health* **2007**, *121*, 909–922. [CrossRef] [PubMed]
9. Ridgers, N.D.; Salmon, J.; Timperio, A. Seasonal changes in physical activity during school recess and lunchtime among Australian children. *J. Sports Sci.* **2018**, *36*, 1508–1514. [CrossRef]
10. Wu, Y.T.; Luben, R.; Wareham, N.; Griffin, S.; Jones, A.P. Weather, day length and physical activity in older adults: Cross-sectional results from the European Prospective Investigation into Cancer and Nutrition (EPIC) Norfolk Cohort. *PLoS ONE* **2017**, *12*, e0177767. [CrossRef]
11. Novoa, N.M.; Varela, G.; Jimenez, M.F.; Ramos, J. Value of the average basal daily walked distance measured using a pedometer to predict maximum oxygen consumption per minute in patients undergoing lung resection. *Eur J Cardiothorac Surg.* **2011**, *39*, 756–762. [CrossRef] [PubMed]
12. Novoa, N.; Varela, G.; Jimenez, M.F.; Aranda, J.L. Influence of major pulmonary resection on postoperative daily ambulatory activity of the patients. *Interact. Cardiovasc. Thorac. Surg.* **2009**, *9*, 934–938. [CrossRef] [PubMed]
13. Granger, C.L.; McDonald, C.F.; Irving, L.; Clark, R.A.; Gough, K.; Murnane, A.; Mileshkin, L.; Krishnasamy, M.; Denehy, L. Low physical activity levels and functional decline in individuals with lung cancer. *Lung Cancer* **2014**, *83*, 292–299. [CrossRef] [PubMed]
14. Sushames, A.; Edwards, A.; Thompson, F.; McDermott, R.; Gebel, K. Validity and Reliability of Fitbit Flex for Step Count, Moderate to Vigorous Physical Activity and Activity Energy Expenditure. *PLoS ONE* **2016**, *11*, e0161224. [CrossRef]
15. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: Guidelines for the six-minute walk test. *Am. J. Respir. Crit. Care Med.* **2002**, *166*, 111–117. [CrossRef]
16. Guyatt, G.H.; Thompson, P.J.; Berman, L.B.; Sullivan, M.J.; Townsend, M.; Jones, N.L.; Pugsley, S.O. How should we measure function in patients with chronic heart and lung disease? *J. Chronic Dis.* **1985**, *38*, 517–524. [CrossRef]
17. Celli, B.; Tetzlaff, K.; Criner, G.; Polkey, M.I.; Sciurba, F.; Casaburi, R.; Tal-Singer, R.; Kawata, A.; Merrill, D.; Rennard, S. The 6-Minute-Walk Distance Test as a Chronic Obstructive Pulmonary Disease Stratification Tool. Insights from the COPD Biomarker Qualification Consortium. *Am. J. Respir. Crit. Care Med.* **2016**, *194*, 1483–1493. [CrossRef]
18. Miller, M. ATS/ERS task force: Standardisation of spirometry. *Eur. Respir. J.* **2005**, *26*, 319–338. [CrossRef]
19. American Thoracic Society. Single-breath carbon monoxide diffusing capacity (transfer factor). Recommendations for a standard technique—1995 update. *Am. J. Respir. Crit. Care Med.* **1995**, *152*, 2185–2198. [CrossRef]
20. Choi, J.K.; Paek, D.; Lee, J.O. Normal predictive values of spirometry in Korean population. *Tuberc. Respir. Dis.* **2005**, *58*, 230–242. [CrossRef]
21. Park, J.; Choi, I.; Park, K. Normal predicted standards of single breath carbon monoxide diffusing capacity of lung in healthy nonsmoking adults. *J. Korean J. Intern. Med.* **1985**, *28*, 176–183.
22. Tudor-Locke, C.; Craig, C.L.; Aoyagi, Y.; Bell, R.C.; Croteau, K.A.; De Bourdeaudhuij, I.; Ewald, B.; Gardner, A.W.; Hatano, Y.; Lutes, L.D.; et al. How many steps/day are enough? For older adults and special populations. *Int. J. Behav. Nutr. Phys. Act.* **2011**, *8*, 80. [CrossRef] [PubMed]

23. Tudor-Locke, C.; Jones, R.; Myers, A.M.; Paterson, D.H.; Ecclestone, N.A. Contribution of structured exercise class participation and informal walking for exercise to daily physical activity in community-dwelling older adults. *Res. Q. Exerc. Sport* **2002**, *73*, 350–356. [CrossRef] [PubMed]
24. American College of Sports Medicine Position Stand. The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults. *Med. Sci. Sports Exerc.* **1998**, *30*, 975–991.
25. World Health Organization (WHO). Definition of an Older or Elderly Person. *World Health Organization.* 2016. Available online: http://www.who.int/healthinfo/survey/ageingdefnolder/en/ (accessed on 3 August 2017).
26. Marjanski, T.; Wnuk, D.; Bosakowski, D.; Szmuda, T.; Sawicka, W.; Rzyman, W. Patients who do not reach a distance of 500 m during the 6-min walk test have an increased risk of postoperative complications and prolonged hospital stay after lobectomy. *Eur. J. Cardiothorac. Surg.* **2015**, *47*, e213–e219. [CrossRef]
27. Alahmari, A.D.; Mackay, A.J.; Patel, A.R.; Kowlessar, B.S.; Singh, R.; Brill, S.E.; Allinson, J.P.; Wedzicha, J.A.; Donaldson, G.C. Influence of weather and atmospheric pollution on physical activity in patients with COPD. *Respir. Res.* **2015**, *16*, 71. [CrossRef]
28. Merchant, A.T.; Dehghan, M.; Akhtar-Danesh, N. Seasonal variation in leisure-time physical activity among Canadians. *Can. J. Public Health* **2007**, *98*, 203–208. [CrossRef]
29. Klenk, J.; Buchele, G.; Rapp, K.; Franke, S.; Peter, R. Walking on sunshine: Effect of weather conditions on physical activity in older people. *J. Epidemiol. Community Health* **2012**, *66*, 474–476. [CrossRef]
30. Shoemaker, M.J.; Roper, S.E.; Calkins, T.N. Seasonal variation of daily physical activity in individuals with heart failure. *Heart Int.* **2016**, *11*, e25–e31. [CrossRef]
31. Elliot, A.J.; Cross, K.W.; Fleming, D.M. Acute respiratory infections and winter pressures on hospital admissions in England and Wales 1990–2005. *J. Public Health* **2008**, *30*, 91–98. [CrossRef]
32. Hasegawa, J.; Suzuki, H.; Yamauchi, T. Impact of season on the association between muscle strength/volume and physical activity among community-dwelling elderly people living in snowy-cold regions. *J. Physiol. Anthr.* **2018**, *37*, 25. [CrossRef] [PubMed]
33. Stefanelli, F.; Meoli, I.; Cobuccio, R.; Curcio, C.; Amore, D.; Casazza, D.; Tracey, M.; Rocco, G. High-intensity training and cardiopulmonary exercise testing in patients with chronic obstructive pulmonary disease and non-small-cell lung cancer undergoing lobectomy. *Eur. J. Cardiothorac. Surg.* **2013**, *44*, e260–e265. [CrossRef] [PubMed]
34. Hattori, K.; Matsuda, T.; Takagi, Y.; Nagaya, M.; Inoue, T.; Nishida, Y.; Hasegawa, Y.; Kawaguchi, K.; Fukui, T.; Ozeki, N.; et al. Preoperative six-minute walk distance is associated with pneumonia after lung resection. *Interact. Cardiovasc. Thorac. Surg.* **2018**, *26*, 277–283. [CrossRef] [PubMed]
35. Lai, Y.; Huang, J.; Yang, M.; Su, J.; Liu, J.; Che, G. Seven-day intensive preoperative rehabilitation for elderly patients with lung cancer: A randomized controlled trial. *J. Surg. Res.* **2017**, *209*, 30–36. [CrossRef]
36. Aspvik, N.P.; Viken, H.; Ingebrigtsen, J.E.; Zisko, N.; Mehus, I.; Wisloff, U.; Stensvold, D. Do weather changes influence physical activity level among older adults?—The Generation 100 study. *PLoS ONE* **2018**, *13*, e0199463. [CrossRef]
37. Chan, C.B.; Ryan, D.A.; Tudor-Locke, C. Relationship between objective measures of physical activity and weather: A longitudinal study. *Int. J. Behav. Nutr. Phys. Act.* **2006**, *3*, 21. [CrossRef]
38. Kimura, T.; Kobayashi, H.; Nakayama, E.; Kakihana, W. Seasonality in physical activity and walking of healthy older adults. *J. Physiol. Anthr.* **2015**, *34*, 33. [CrossRef]
39. Togo, F.; Watanabe, E.; Park, H.; Shephard, R.J.; Aoyagi, Y. Meteorology and the physical activity of the elderly: The Nakanojo Study. *Int. J. Biometeorol.* **2005**, *50*, 83–89. [CrossRef]
40. Bluestein, M. An evaluation of the wind chill factor: Its development and applicability. *J. Biomech. Eng.* **1998**, *120*, 255–258. [CrossRef]



Journal of
Clinical Medicine

MDPI

Article

Postprandial Hypotension as a Risk Factor for the Development of New Cardiovascular Disease: A Prospective Cohort Study with 36 Month Follow-Up in Community-Dwelling Elderly People

Aelee Jang

Department of Nursing, University of Ulsan, 93, Daehak-ro, Nam-gu, Ulsan 44610, Korea; aeleejang@ulsan.ac.kr; Tel.: +82-52-259-1252

Received: 2 January 2020; Accepted: 23 January 2020; Published: 27 January 2020

Abstract: Postprandial hypotension (PPH) is common among the elderly. However, it is unknown whether the presence of PPH can predict the development of new cardiovascular disease (CVD) in the elderly during the long-term period. This study aimed to prospectively evaluate the presence of PPH and the development of new CVD within a 36 month period in 94 community-dwelling elderly people without a history of CVD. PPH was diagnosed in 47 (50.0%) participants at baseline and in 7 (7.4%) during the follow-up period. Thirty participants (31.9%) developed new CVD within 36 months. We performed a time-dependent Cox regression analysis with PPH, hypertension, diabetes, and body mass index (BMI) as time-varying covariates. In the univariate analyses, the presence of PPH, higher BMI, hypertension, diabetes mellitus, and higher systolic and diastolic blood pressure were associated with the development of new CVD. The multivariate analysis indicated that the relationship between PPH and the development of new CVD remained (adjusted hazard ratio 11.18, 95% confidence interval 2.43–51.38, $p = 0.002$) even after controlling for other variables as covariates. In conclusion, the presence of PPH can predict the development of new CVD. Elderly people with PPH may require close surveillance to prevent CVD.

Keywords: cardiovascular disease; postprandial; hypotension; blood pressure; elderly

1. Introduction

Postprandial hypotension (PPH) is a common but often unrecognized disorder in the elderly [1]. PPH is defined as a ≥20 mmHg decrease in systolic blood pressure (SBP) within 2 h after a meal [2]. The prevalence of PPH is 20–91% in hospitalized geriatric patients [3–6]. The pathogenesis of PPH is unclear, but an inadequate cardiovascular response to postprandial splanchnic blood pooling is regarded as a primary mechanism [2]. In the elderly, sympathetic activation to decrease the effective circulating volume is often suppressed; hence, a continuous pooling of blood in the splanchnic bed may result in a significant decrease in blood pressure (BP) after a meal [7,8]. Changes in cardiovascular function and the neurohormonal response with aging can also contribute to diverse comorbidities associated with the dysregulation of BP homeostasis [9].

Cardiovascular diseases (CVDs), such as stroke, transient ischemic attack, angina, and myocardial infarction, are the main cause of mortality and work disability in the elderly [10,11]. Early recognition of subclinical risk factors is essential for the proper surveillance and prevention of new CVD. PPH is also recognized as an important clinical issue, particularly as a significant risk factor for subsequent CVD, but this has not yet been accurately confirmed [2,12]. Some studies have suggested an association between PPH and CVD and mortality [9,13]. However, it is difficult to determine the casual relationship between PPH and CVD development owing to differences in patient selection, study design, and diagnosis criteria among previous studies [9,14–16]. Furthermore, the diagnosis of PPH was heterogeneous in previous studies, such as with regard to the time interval between meal consumption and BP measurement [5]. Therefore, to investigate whether the presence of PPH can predict the development of new CVD in the elderly, this study aimed to prospectively evaluate the existence of PPH and the development of new CVD among community-dwelling elderly people in a period of 36 months.

2. Materials and Methods

2.1. Study Design and Participants

This prospective cohort study enrolled 94 participants from three senior community centers in South Korea between 2011 and 2015. The inclusion criteria were (1) age ≥65 years, (2) ability to eat independently and maintain a sitting position for 2 h after a meal, and (3) ability to perform activities of daily living independently. The exclusion criteria included (1) impaired cognitive function, (2) psychiatric illness, (3) recent hospitalization owing to acute illness within 1 month prior to the study, (4) medical history of CVD (congestive heart failure, myocardial infarction, angina, or cerebrovascular accidents, including transient ischemic attack), or (5) use of medications affecting gastrointestinal motility. After evaluating the presence of PPH at baseline, the new onset of PPH was evaluated during the follow-up period of 36 months. Body mass index (BMI) and the presence of hypertension and diabetes were also assessed regularly. This study was conducted following the guidelines of the Declaration of Helsinki of 1975, revised in 2013, and was approved by the institutional review board of Pusan National University Hospital (E-2014007), located in B city, South Korea. All participants provided written informed consent.

2.2. Evaluation of PPH and Acquisition of Demographic Data

BP was measured using an automated sphygmomanometer (A&D UA-851; A&D Company, Japan) following the European Society of Hypertension guidelines [17]. BP was measured by a well-trained nurse with the participant in a sitting position. Baseline BP was measured before a meal. Participants were asked to refrain from eating food, drinking coffee or alcohol, and smoking 4 h before measurement. All participants took any prescription medications after the evaluation of PPH, including antihypertensive drugs and hypoglycemic agents. BP was measured twice with a 5 min interval before a meal, and the average of the two measurements was used as the baseline BP. To measure postprandial BP, participants were provided the same standardized meal comprising 210 g rice, 100 g soup, and 70 g side dishes, which was approximately 500 kcal. After the participant consumed the meal, postprandial BP was measured every 15 min for 2 h. PPH was defined as a decrease in SBP of ≥20 mmHg from the baseline within 2 h after a meal [2]. We analyzed the reliability of measurements and found that the intraclass correlation coefficients (ICCs) for intra-rater reliability for SBP and diastolic blood pressure (DBP) were 0.995 (95% confidence interval (CI) 0.975–0.998, $p < 0.001$) and 0.989 (95% CI 0.942–0.996, $p < 0.001$), respectively.

Demographic and clinical information was also obtained by interviewing the participants using a questionnaire. Information on age, living conditions, alcohol intake, and smoking habits was collected, and information on medical history was obtained from hospital records after consent was obtained from the participants and their families. Bodyweight and height were measured using an automated measuring machine (G-Tech International Co., Ltd., Uijeongbu, Korea) to determine the BMI.

2.3. Surveillance of CVD Development

The primary endpoint was the occurrence of new CVD within 36 months. CVD in this study included coronary heart disease (such as congestive heart failure, angina, and myocardial infarction) and cerebrovascular disease (such as stroke and transient ischemic attack) according to the World Health Organization definition of CVD [18]. Participants were regularly followed every 3 months, and CVD-related information was obtained from them or their families during their visits to the centers or via telephone calls. If the participant was diagnosed with a new disease, additional information was also obtained from the medical record or prescription from the hospital at which the participant was diagnosed with new conditions, including the new onset of hypertension, diabetes, and CVDs.

2.4. Statistical Analysis

The reliability of BP measurements was calculated by determining the ICC (two-way random effects model). Variables are expressed as frequencies and percentages for categorical data, and means ± standard deviation for numerical data. Group differences were assessed using the chi-squared test or Fisher's exact test for categorical data and the independent t test or Mann–Whitney U test for numerical data as appropriate. To determine whether the distribution was normal, we used the Shapiro–Wilk test. Time to CVD was estimated using Kaplan–Meier curves. Survival curves were compared between groups using the log-rank test. Cox regression models with time-varying covariates were used to identify prognostic factors that were independently related to CVD. The main predictors were baseline PPH and PPH as a time-varying covariate (incorporating new onset PPH during the follow-up and duration of PPH as a time-varying covariate). In addition, hypertension, diabetes mellitus, and BMI were monitored during the follow-up period. These predicting variables were considered as time-varying covariates in time-dependent Cox regression models. All statistical analyses were performed using SPSS version 24.0 (IBM Corp., Armonk, NY, USA) and R version 3.5.1, and p-values <0.05 were considered statistically significant.

3. Results

3.1. Baseline Characteristics

The participants comprised 79 (84.0%) females and 15 (16.0%) males. The mean age was 73.1 ± 4.8 years, and the mean BMI was 23.7 ± 2.5 kg/m^2. Among the participants, 67 (71.3%) had lower than middle school education, and 62 (66.0%) were currently living with their family or spouse. The most common comorbidity was hypertension (50.0%), followed by diabetes mellitus (19.1%). At baseline, 47 (50%) participants were diagnosed with PPH, but there were no significant differences in baseline characteristics, with the exception of BP, between participants with and without PPH (Table 1).

Table 1. Baseline characteristics of participants with and without PPH.

Characteristics	Total (n = 94)	PPH (n = 47)	Non-PPH (n = 47)	p-Value
Age (years)	73.1 ± 4.8	73.6 ± 4.4	72.7 ± 5.1	0.355 [1]
Sex				0.778 [3]
Male	15 (16.0)	8 (17.0)	7 (14.9)	
Female	79 (84.0)	39 (83.0)	40 (85.1)	
Body Mass Index (kg/m^2)	23.7 ± 2.5	23.8 ± 2.7	23.5 ± 2.4	0.642 [2]
Education Level				0.509 [3]
Elementary School	67 (71.3)	36 (76.6)	31 (66.0)	
Middle School	13 (13.8)	5 (10.6)	8 (17.0)	
High School	14 (14.9)	6 (12.8)	8 (17.0)	
Living Status				1.000 [3]
Living Alone	32 (34.0)	16 (34.0)	16 (34.0)	
With Family or Spouse	62 (66.0)	31 (66.0)	31 (66.0)	
Alcohol Drinking				0.370 [3]
Yes	13 (13.8)	5 (10.6)	8 (17.0)	
No	81 (86.2)	42 (89.4)	39 (83.0)	
Smoking				1.000 [4]
Yes	4 (4.3)	2 (4.3)	2 (4.3)	
No	90 (95.7)	45 (95.7)	45 (95.7)	
Hypertension	47 (50.0)	26 (55.3)	21 (44.7)	0.302 [3]
Diabetes Mellitus	18 (19.1)	11 (23.4)	7 (14.9)	0.294 [3]
Baseline SBP (mmHg)	128.6 ± 20.2	140.1 ± 17.2	117.1 ± 16.1	<0.001 [2]
<120	34 (36.2)	6 (12.8)	28 (59.6)	<0.001 [3]
≥120 to <140	31 (33.0)	15 (31.9)	16 (34.0)	
≥140	29 (30.9)	26 (55.3)	3 (6.4)	
Baseline DBP (mmHg)	75.2 ± 9.9	78.7 ± 7.5	71.7 ± 10.9	<0.001 [2]
<80	64 (68.1)	23 (48.9)	41 (87.2)	<0.001 [3]
≥80	30 (31.9)	24 (51.1)	6 (12.8)	
Postprandial SBP Change (mmHg)				<0.001 [3]
<10	26 (27.7)	0 (0.0)	26 (55.3)	
≥10 to <20	21 (22.3)	0 (0.0)	21 (44.7)	
≥20	47 (50.0)	47 (100.0)	0 (0.0)	
Postprandial DBP Change (mmHg)				<0.001 [3]
<10	41 (43.6)	9 (19.1)	32 (68.1)	
≥10	53 (56.4)	38 (80.9)	15 (31.9)	

Values are either frequency with percentage in parentheses or mean ± standard deviation. PPH, postprandial hypotension; SBP, systolic blood pressure; DBP, diastolic blood pressure. [1] P values were derived from independent t tests. [2] P values were derived from Mann–Whitney's U test. [3] P values were derived using chi-square tests. [4] P values were derived using Fisher's exact test. Shapiro–Wilk's test was employed for test of normality assumption.

3.2. Incidence of CVD

During the follow-up, 30 (31.9%) patients developed new CVD, of whom 4 passed away owing to acute myocardial infarction (n = 2) and ischemic stroke (n = 2). The CVD incidence was significantly higher in the group with PPH than the group without PPH at baseline (55.3% vs. 8.5%, $p < 0.001$, Figure 1).

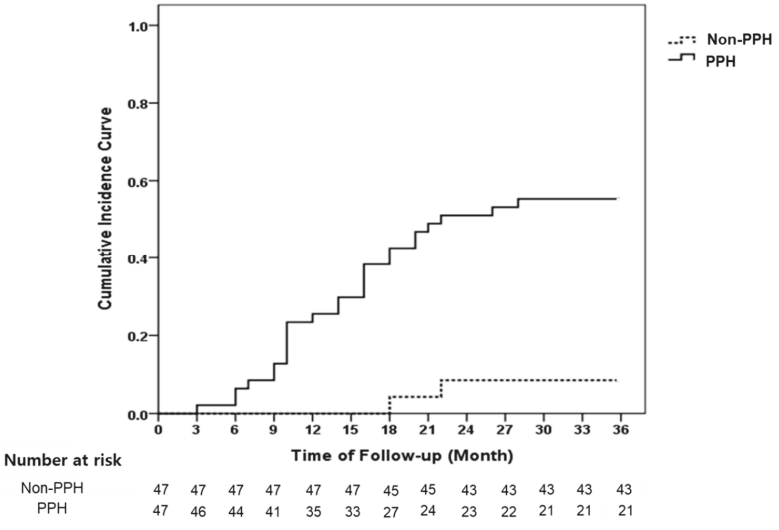


Figure 1. Kaplan–Meier curve of the incidence of new CVD between participants with and without PPH at baseline. The 3 year incidence of CVD was significantly higher in the group with PPH than in the group without PPH at baseline (55.3% vs. 8.5%, log-rank test, $p < 0.001$). PPH, postprandial hypotension; CVD, cardiovascular disease.

Table 2 indicates the differences between groups with respect to CVD development. The presence of PPH at baseline (86.7% in the CVD group vs. 32.8% in the non-CVD group, $p < 0.001$), higher BMI (24.5 ± 2.1 in the CVD group vs. 23.3 ± 2.7 in the non-CVD group), hypertension (66.7% in the CVD group vs. 42.2% in the non-CVD group), higher baseline SBP (139.3 ± 20.9 mmHg in the CVD group vs. 123.6 ± 17.9 mmHg in the non-CVD group), and higher baseline DBP (79.0 ± 8.6 mmHg in the CVD group vs. 73.4 ± 10.1 mmHg in the non-CVD group) were significantly related to the development of new CVD. The percentage of participants with an SBP change of ≥20 mmHg (86.7% vs. 32.8%, respectively, $p < 0.001$) and a DBP change of ≥10 mmHg (73.3% vs. 48.4%, respectively, $p = 0.023$) were significantly higher among participants who developed new CVD than among those who did not.

Table 2. Differences of characteristics between participants who developed and did not develop new CVD.

	Total (n = 94)	CVD (n = 30)	Non-CVD (n = 64)	p-Value
PPH				<0.001 [3]
Yes	47 (50.0)	26 (86.7)	21 (32.8)	
No	47 (50.0)	4 (13.3)	43 (67.2)	
Age (years)	73.1 ± 4.8	73.0 ± 4.1	73.2 ± 5.1	0.849 [1]
Sex				1.000 [4]
Male	15 (16.0)	5 (16.7)	10 (15.6)	
Female	79 (84.0)	25 (83.3)	54 (84.4)	
Body mass index (kg/m^2)	23.7 ± 2.5	24.5 ± 2.1	23.3 ± 2.7	0.013 [2]
Education level				0.534 [4]
Elementary school	67 (71.3)	24 (80.0)	43 (67.2)	
Middle school	13 (13.8)	3 (10.0)	10 (15.6)	
High school	14 (14.9)	3 (10.0)	11 (17.2)	
Living status				0.571 [3]
Living alone	32 (34.0)	9 (30.0)	23 (35.9)	
With family or spouse	62 (66.0)	21 (70.0)	41 (64.1)	

Table 2. *Cont.*

	Total (n = 94)	CVD (n = 30)	Non-CVD (n = 64)	p-Value
Alcohol drinking				0.336 [4]
Yes	13 (13.8)	6 (20.0)	7 (10.9)	
No	81 (86.2)	24 (80.0)	57 (89.1)	
Smoking				0.590 [4]
Yes	4 (4.3)	2 (6.7)	2 (3.1)	
No	90 (95.7)	28 (93.3)	62 (96.9)	
Hypertension	47 (50.0)	20 (66.7)	27 (42.2)	0.027 [3]
Diabetes mellitus	18 (19.1)	9(30.0)	9 (14.1)	0.067 [3]
Baseline SBP (mmHg)	128.6 ± 20.2	139.3 ± 20.9	123.6 ± 17.9	<0.001 [2]
<120	34 (36.2)	5 (16.7)	29 (45.3)	<0.001 [3]
≥120 to <140	31 (33.0)	7 (23.3)	24 (37.5)	
≥140	29 (30.9)	18 (60.0)	11 (17.2)	
Baseline DBP (mmHg)	75.2 ± 9.9	79.0 ± 8.6	73.4 ± 10.1	0.005 [2]
<80	64 (68.1)	15 (50.0)	49 (76.6)	0.010 [3]
≥80	30 (31.9)	15 (50.0)	15 (23.4)	
Postprandial SBP change (mmHg)				<0.001 [3]
<10	26 (27.7)	3 (10.0)	23 (35.9)	
≥10 to <20	21 (22.3)	1 (3.3)	20 (31.3)	
≥20	47 (50.0)	26 (86.7)	21 (32.8)	
Postprandial DBP change (mmHg)				0.023 [3]
<10	41 (43.6)	8 (26.7)	33 (51.6)	
≥10	53 (56.4)	22 (73.3)	31 (48.4)	

Values are either frequency with percentage in parentheses or mean ± standard deviation. PPH, postprandial hypotension; CVD, cardiovascular disease; SBP, systolic blood pressure; DBP, diastolic blood pressure. [1] P values were derived from independent t tests. [2] P values were derived from Mann–Whitney's U test. [3] P values were derived using chi-square tests. [4] P values were derived using Fisher's exact test. Shapiro–Wilk's test was employed for test of normality assumption.

3.3. Predictive Factors for the Development of New CVD

Ninety-four patients were available for the Cox regression models to assess the relationship between PPH and CVD using PPH as the time-varying covariate. There were 30 cases of CVD. We considered time-varying covariates, including PPH, hypertension, diabetes, and BMI, in the time-dependent Cox regression models. Among the 47 patients without PPH at baseline, 7 were diagnosed with PPH during the follow-up period. The main predictive factors were baseline PPH, with PPH as a time-varying covariate incorporating new onset PPH during follow-up, and the duration of PPH as a time-varying covariate. New onset hypertension and diabetes mellitus during the follow-up period were observed in 12 and 11 patients, respectively. In addition, BMI was measured at baseline, 1 year, and 2 years. All predictive variables were considered as time-varying covariates in time-dependent Cox regression models. In the univariate analyses, the presence of PPH, higher BMI, hypertension, diabetes mellitus, and higher SBP and DBP were found to be associated with the development of new CVD, while age, sex, education level, living status, alcohol consumption, and smoking were not. Patients with PPH were more likely to develop new CVD (crude hazard ratio (HR) 15.97, 95% CI 3.80–67.08, $p < 0.001$). The multivariate analysis that included significant factors in the univariate analyses as covariates revealed that the relationship between PPH and the development of new CVD remained (adjusted HR 11.18, 95% CI 2.43–51.38, $p = 0.002$) even after controlling for other variables as covariates in the multivariate analysis. Hypertension was also found to be a significant factor affecting CVD (adjusted HR 3.26, 95% CI 1.22–8.76, $p = 0.019$) (Table 3).

Table 3. Summary of time-dependent Cox regression analyses using time-varying covariates (a total of 94 subjects, 30 CVDs).

Characteristic	Univariate			Multivariate		
	HR	95% CI	*p*-Value	HR	95% CI	*p*-Value
PPH	15.97	3.80–67.08	<0.001	11.18	2.43–51.38	0.002
Age (years)	1.00	0.93–1.07	0.940			
Sex (female)	0.90	0.34–2.35	0.830			
BMI (kg/m^2)	1.14	1.01–1.29	0.030	1.10	0.96–1.26	0.173
Education level (vs. Elementary school)						
Middle school	0.61	0.19–2.04	0.430			
High school	0.57	0.17–1.88	0.350			
Living status (Living alone)	0.80	0.37–1.75	0.580			
Alcohol drinking	1.61	0.66–3.95	0.300			
Smoking	1.77	0.42–7.46	0.430			
Hypertension	4.61	1.77–12.05	0.002	3.26	1.22–8.73	0.019
Diabetes mellitus	3.20	1.56–6.56	0.002	1.80	0.84–3.89	0.132
Baseline SBP (mmHg)	1.03	1.02–1.05	<0.001	1.02	0.99–1.05	0.148
Baseline DBP (mmHg)	1.03	1.01–1.06	0.013	0.97	0.91–1.02	0.245

HR, hazard ratio; CVD, cardiovascular disease; BMI, body mass index; SBP, systolic blood pressure; DBP, diastolic blood pressure; PPH, postprandial hypotension.

4. Discussion

This study investigated the time-varying effect of PPH associated with new CVD in community-dwelling elderly people. The study indicated that approximately half (55.3%) of the participants with PPH developed new CVD within the 36 month follow-up period, whereas only 8.5% of participants without PPH developed CVD. After adjustment for other covariates, the time-varying effect of PPH on the development of new CVD was found to be significant (HR 11.18, 95% CI 2.43–51.38). These results suggest that elderly people with PPH were more likely to develop new CVD and, thus, required close surveillance.

The reported prevalence of PPH varies widely in previous studies depending on the patient population and diagnostic methods [5]. Studies of institutionalized elderly patients reported a PPH prevalence of 24–38% [3,4]. Meanwhile, the incidence of PPH among hospitalized geriatric patients was considerably high (up to 91%) [5,6]. However, studies investigating the incidence of PPH among healthy elderly people are relatively rare [19]. Unlike previous studies, the present study investigated PPH in relatively healthy elderly people who were not institutionalized or hospitalized. Although the study population comprised elderly patients, the PPH incidence was 50%, which seems to be slightly higher than that reported in previous studies. The development of PPH is highly dependent on meal composition. Carbohydrate-rich meals predispose patients to a more immediate decrease in BP than do meals containing primarily protein or fat [20]. Therefore, it is expected that PPH might be more prevalent in people from Asian regions owing to their carbohydrate-dominant diets. However, few studies have investigated PPH in Asian patients [21]. To date, this is the first prospective cohort study to investigate PPH and its association with CVD in relatively healthy Asian elderly people. In this study, a stringent diagnostic approach was used for PPH, such as provision of a standardized meal and more frequent BP measurements. Furthermore, in this study, variance in BP changes related to food composition was reduced by administering the same standardized meal to all participants. Therefore, this study may provide relevant insights into the epidemiology of PPH in Asian regions. Large cohort studies are recommended to further determine the prevalence of PPH in different patient populations using a standardized diagnostic method.

In this study, 30 (31.9%) participants developed new CVD during the 36 month follow-up. In one study on PPH as a predictor of new CVD [14], 40.8% of older nursing home residents developed CVD during a 29 month follow-up. The patients with new CVD exhibited a significantly greater

decline in postprandial SBP than did those without ($p < 0.001$). The other previous study found a 59.2% occurrence of new CVD during a 4.7 year follow-up period in older low-level-care residents in long-term health facilities [9]. The occurrence of new CVD cases was lower in this study (31.9%) than in previous studies (40.1–59.2%) because this study was conducted in a relatively healthy cohort from a community setting. Furthermore, the previous studies did not fully exclude populations with a history of CVD. In contrast, this prospective study excluded people with a history of CVD to clearly demonstrate the relationship between PPH and CVD.

In this study, the presence of PPH, higher BMI, hypertension, diabetes mellitus, and higher SBP and DBP were predictive factors for the development of CVD in the univariate analysis using the time-dependent Cox proportional hazard model. After adjustment for these factors, the time-varying effect of PPH and hypertension on the occurrence of new CVD still remained in the elderly people. According to Aronow et al. [14], the mean maximal decrease in postprandial SBP was a significant risk factor for developing CVD. Similarly, in previous cross-sectional studies, PPH was considered an independent predictor of asymptomatic cerebrovascular damage in healthy participants and hospitalized patients with essential hypertension [15,16]. Despite the small sample size, this study is meaningful as a prospective cohort study for the risk factors of new CVD in the presence or absence of PPH among community-dwelling elderly people without a history of CVD.

The total mortality (7/94, 7.4%) and CVD-related mortality (4/94, 4.3%) rates in this study were relatively lower than those reported in previous studies (16.9–54.2%) [9,13,14]. Again, this is likely because this study included relatively healthy older adults (i.e., from a community setting) in contrast to previous studies, which generally included hospitalized geriatric patients or patients in nursing homes. The mean age of the participants in this study (73.1 years) was also relatively lower than that in previous studies (77.8–83.2 years) [9,13,14]. In a previous study of hypertensive elderly patients who were followed up at the cardiology clinic for 4 years, the total mortality rate was 16.9%, and a greater decrease in postprandial SBP was associated with a greater increase in CVD mortality. In that prospective study, CVD mortality constituted 50% of the total mortality rate, while PPH accounted for 52.9% of deaths owing to CVD [13], which was similar to the findings of the present study. In a study by Fisher et al. [9], PPH was the only risk factor for all-cause mortality, and 50% of deaths were attributed to CVD. Participants with a postprandial SBP difference of ≥20 mmHg had the lowest survival during the 4.7 year follow-up period. They also reported that PPH accounted for 47% of deaths in long-term healthcare facilities, which was lower than the 57.1% (4 of 7 deaths) found in the present study. Populations in long-term healthcare facilities typically require nursing care for monitoring and preventing serious diseases, including CVD. These perspectives may explain the difference in attributing the total mortality rate to CVD mortality between the present study and the previous study mentioned above [9]. Therefore, prevention and close monitoring are needed to reduce deaths due to new CVD by classifying patients diagnosed with PPH as high risk for CVD.

This study had some limitations. First, the sample size was relatively small. However, compared to previous studies, this study was conducted over a long-term follow-up with the time-varying effect of PPH. The post hoc power for investigating the primary objective of this study was approximately 99%. Second, despite efforts to recruit a similar ratio of males and females, the participants were predominantly female because of the higher proportion of elderly women in the Korean elderly population. Future studies should confirm CVD risk factors in sex-balanced populations with postprandial SBP reduction during long-term follow-up. Third, the level of physical activity was not considered in this study, which could affect the development of CVDs. Nevertheless, the participants had similar lifestyle patterns because they resided in the same town and enjoyed the same activity programs provided by the community welfare center. Therefore, their physical activity levels may be similar. Fourth, the pathophysiological relationship between PPH and CVD remains vague despite PPH being identified as an independent time-varying variable for the development of new CVD. Autonomic dysfunction may play a major role in the development of PPH-related new CVD because PPH has been reported to be pathologically associated with autonomic dysfunction. This needs further

evaluation using animal models. Fifth, the positive association between PPH and CVD may only reflect the high baseline SBP on the risk of CVD because the baseline SBP level in the PPH group was much higher than that in the non-PPH group. Although hypertension was adjusted in the multivariate model, based on the small sample size, the effect of high baseline SBP may persist. As it is impossible to stratify the analysis by hypertension status at baseline, owing to the limited sample size, further research in the form of a well-designed, large-sample study in a normotensive elderly population is necessary to definitively establish the effect of PPH on the risk of CVD.

This study investigated the relationship between PPH and the development of new CVD in elderly people during a long-term follow-up period. The results also suggest that the underlying reason of CVD development could be the greater decline in postprandial SBP (such as that occurring in PPH). This study demonstrated that increases in SBP variations after a meal may be a manifestation of CVD development and may reflect the presence of subclinical cardiovascular damage.

5. Conclusions

This study revealed that PPH was an independent predictor of new CVD among community-dwelling elderly people during a 36 month follow-up. Thus, CVD development may be prevented or monitored based on the presence of PPH. However, longitudinal studies in larger samples are needed to clarify the prognostic significance of PPH in the development of new CVD.

Funding: This work was supported by the 2019 Research Fund of University of Ulsan.

Conflicts of Interest: The author declares no conflicts 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. Luciano, G.L.; Brennan, M.J.; Rothberg, M.B. Postprandial hypotension. *Am. J. Med.* **2010**, *123*, e1–e6. [CrossRef] [PubMed]
2. Jansen, R.W.; Lipsitz, L.A. Postprandial hypotension: Epidemiology, pathophysiology, and clinical management. *Ann. Intern. Med.* **1995**, *122*, 286–295. [CrossRef] [PubMed]
3. Aronow, W.S.; Ahn, C. Postprandial hypotension in 499 elderly persons in a long-term health care facility. *J. Am. Geriatr. Soc.* **1994**, *42*, 930–932. [CrossRef] [PubMed]
4. Le Couteur, D.G.; Fisher, A.A.; Davis, M.W.; McLean, A.J. Postprandial systolic blood pressure responses of older people in residential care: Association with risk of falling. *Gerontology* **2003**, *49*, 260–264. [CrossRef] [PubMed]
5. Van Orshoven, N.P.; Jansen, P.A.; Oudejans, I.; Schoon, Y.; Oey, P.L. Postprandial hypotension in clinical geriatric patients and healthy elderly: Prevalence related to patient selection and diagnostic criteria. *J. Aging Res.* **2010**, *2010*, 243752. [CrossRef] [PubMed]
6. Puisieux, F.; Bulckaen, H.; Fauchais, A.L.; Drumez, S.; Salomez-Granier, F.; Dewailly, P. Ambulatory blood pressure monitoring and postprandial hypotension in elderly persons with falls or syncopes. *J. Gerontol. A Biol. Sci. Med. Sci.* **2000**, *55*, M535–M540. [CrossRef]
7. Lipsitz, L.A.; Ryan, S.M.; Parker, J.A.; Freeman, R.; Wei, J.Y.; Goldberger, A.L. Hemodynamic and autonomic nervous system responses to mixed meal ingestion in healthy young and old subjects and dysautonomic patients with postprandial hypotension. *Circulation* **1993**, *87*, 391–400. [CrossRef]
8. Jang, A.L.; Hwang, S.K.; Kim, D.U. Effects of guar gum ingestion on postprandial blood pressure in older adults. *J. Nutr. Health. Aging.* **2015**, *19*, 299–304. [CrossRef]
9. Fisher, A.A.; Davis, M.W.; Srikusalanukul, W.; Budge, M.M. Postprandial hypotension predicts all-cause mortality in older, low-level care residents. *J. Am. Geriatr. Soc.* **2005**, *53*, 1313–1320. [CrossRef] [PubMed]
10. World Health Organization (WHO). The Top 10 Causes of Death. Available online: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death (accessed on 10 September 2019).
11. Virtanen, M.; Lallukka, T.; Ervasti, J.; Rahkonen, O.; Lahelma, E.; Pentti, J.; Pietiläinen, O.; Vahtera, J.; Kivimäki, M. The joint contribution of cardiovascular disease and socioeconomic status to disability retirement: A register linkage study. *Int. J. Cardiol.* **2017**, *230*, 222–227. [CrossRef] [PubMed]

12. Trahair, L.G.; Horowitz, M.; Jones, K.L. Postprandial hypotension: A systematic review. *J. Am. Med. Dir. Assoc.* **2014**, *15*, 394–409. [CrossRef] [PubMed]
13. Zanasi, A.; Tincani, E.; Evandri, V.; Giovanardi, P.; Bertolotti, M.; Rioli, G. Meal-induced blood pressure variation and cardiovascular mortality in ambulatory hypertensive elderly patients: Preliminary results. *J. Hypertens.* **2012**, *30*, 2125–2132. [CrossRef] [PubMed]
14. Aronow, W.S.; Ahn, C. Association of postprandial hypotension with incidence of falls, syncope, coronary events, stroke, and total mortality at 29-month follow-up in 499 older nursing home residents. *J. Am. Geriatr. Soc.* **1997**, *45*, 1051–1053. [CrossRef] [PubMed]
15. Tabara, Y.; Okada, Y.; Uetani, E.; Nagai, T.; Igase, M.; Kido, T.; Ochi, N.; Ohara, M.; Takita, R.; Kohara, K.; et al. Postprandial hypotension as a risk marker for asymptomatic lacunar infarction. *J. Hypertens.* **2014**, *32*, 1084–1090. [CrossRef] [PubMed]
16. Kohara, K.; Jiang, Y.; Igase, M.; Takata, Y.; Fukuoka, T.; Okura, T.; Kitami, Y.; Hiwada, K. Postprandial hypotension is associated with asymptomatic cerebrovascular damage in essential hypotensive patients. *Hypertension* **1999**, *33*, 565–568. [CrossRef] [PubMed]
17. Mancia, G.; Grassi, G. The new European society of hypertension/European society of cardiology (ESH/ESC) guidelines. *Ther. Adv. Cardiovasc. Dis.* **2008**, *2*, 5–12. [CrossRef] [PubMed]
18. World Health Organization (WHO). Cardiovascular Diseases (CVDs). Available online: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) (accessed on 15 September 2019).
19. Peitzman, S.J.; Berger, S.R. Postprandial blood pressure decrease in well elderly persons. *Arch. Intern. Med.* **1989**, *149*, 286–288. [CrossRef] [PubMed]
20. Gentilcore, D.; Hausken, T.; Meyer, J.H.; Chapman, I.M.; Horowitz, M.; Jones, K.L. Effects of intraduodenal glucose, fat, and protein on blood pressure, heart rate, and splanchnic blood flow in healthy older subjects. *Am. J. Clin. Nutr.* **2008**, *87*, 156–161. [CrossRef] [PubMed]
21. Son, J.T.; Lee, E. Prevalence and risk factors of postprandial hypotension in Korean elderly people. *J. Korean Acad. Nurs.* **2009**, *39*, 198–206. [CrossRef]

Journal of
Clinical Medicine

MDPI

Article

Relationship between Morbidity and Health Behavior in Chronic Diseases

Munjae Lee [1,2], Sewon Park [1,2] and Kyu-Sung Lee [2,3,*]

[1] Research Institute for Future Medicine, Samsung Medical Center, Seoul 06351, Korea; emunjae@skku.edu (M.L.); se10919@g.skku.edu (S.P.)
[2] Department of Medical Device Management and Research, SAIHST, Sungkyunkwan University, Seoul 06355, Korea
[3] Department of Urology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul 06351, Korea
* Correspondence: ksleedr@skku.edu

Received: 15 November 2019; Accepted: 26 December 2019; Published: 2 January 2020

Abstract: This study aimed to analyze the demographic characteristics and health behaviors related to chronic diseases and to identify factors that may affect chronic diseases. Data from the Seventh Korea National Health and Nutrition Examination Survey were used, and 3795 adults aged above 40 years were included. The following demographic variables were obtained: sex, age, education, income, type of health insurance, and private insurance. The following health behavior factors were also analyzed: medical checkup, drinking, smoking, exercise, obesity, and hypercholesterolemia. Participants with lower socioeconomic status had a higher risk of developing chronic diseases. Meanwhile, those with private health insurance had a lower risk of developing chronic diseases. In addition, participants who underwent medical checkups and performed exercises had a lower risk, while those with obesity and hypercholesterolemia had a higher risk of developing chronic diseases. It is necessary to manage chronic diseases through comprehensive programs, rather than managing these diseases individually, and through community primary care institutions to improve health behaviors.

Keywords: chronic disease; health behavior; socioeconomic status; primary care; Korea

1. Introduction

An individual's behavior related to health may have an effect on their physical health or ability to recover from illness. In particular, health-related behavior, such as a lack of exercise, smoking, and drinking, are some of the main factors that can contribute to morbidity and mortality [1–3]. Health behavior affects 40% of premature deaths; in order to reduce premature mortality, improving health behaviors is more cost-effective than improving the social and physical environments or health-care systems [4]. These health behaviors are important in maintaining good health, which is influenced by biological and socioeconomic factors, among others [5,6]. Rapid economic growth, high health-care costs, lifestyle changes, and population aging have been associated with an increased prevalence of chronic diseases worldwide. Chronic diseases may cause complications, and thus, require continuous care and are among the types of diseases with high health-care costs due to their long disease duration [7–9].

Chronic diseases, one of the leading causes of death worldwide, especially cardiocerebrovascular diseases, diabetes, and hypertension, have a high mortality rate. However, the mortality rate of chronic diseases can be reduced through prevention [10–12]. Chronic disease is closely related to changes in health behaviors; the main health behaviors affecting the development of chronic diseases include health risk behavioral factors, such as smoking, drinking, and physical activities, and clinicopathologic factors, including obesity, hypertension, and hypercholesterolemia [13,14]. In particular, since health-related lifestyles have increased the risk of mortality, the significance of managing health risk behavioral factors

has also been increasing. Thus, it is necessary to prevent chronic diseases and delay the aggravation of symptoms by improving individual lifestyles [15–17]. In addition, individual health behaviors may differ according to sociodemographic characteristics including age and sex [18]. In the identification of the individual physical condition, sociodemographic and socioeconomic factors are known to act as important factors, and prevalence rates vary in accordance with the individual's income level, education level, and socioeconomic factors [19].

Previous studies have analyzed the relationship between chronic diseases and health promotion behaviors, but were only conducted in predetermined age groups, such as in older patients, or examined the relationship between chronic diseases and health behaviors while only targeting certain chronic diseases [20–24]. As the number of polychronic patients has increased, a comprehensive analysis of chronic diseases is required. To date, the number of studies evaluating patients with chronic diseases is limited. Accordingly, in this study, we aimed to analyze the sociodemographic characteristics and health behaviors related to the development of chronic diseases and to identify factors that may have an effect on the morbidity of chronic diseases. Through this and by suggesting measures to contribute to the effective management and prevention of chronic diseases, we intend to promote the health of the people.

2. Experimental Section

2.1. Data Source and Research Participants

In this study, we utilized the source data from the 2nd year (2017) of the 7th period of the Korea National Health and Nutrition Examination Survey, performed by the Korea Centers for Disease Control and Prevention. The Korea National Health and Nutrition Examination Survey (KNHANES) is a nationwide national survey, conducted to determine health-related parameters including the prevalence of chronic diseases and health behaviors based on Article 16 of the National Health Promotion Act. A total of 10,430 individuals from 3580 households were surveyed, but only 8127 participated in the study. Of them, 5159 aged above 40 years were extracted. In addition, 658 individuals whose answers were not related to chronic diseases were excluded; hence, only 3795 participants were analyzed, after further excluding 706 who did not respond to the questions related to health behaviors. The KNHANES is approved by the ethical committee of the Korea Centers for Disease Control and Prevention. The requirement for informed consent was waived because data in the KNHANE database are anonymized in adherence to strict confidentiality guidelines. The flowchart is shown in Figure 1.

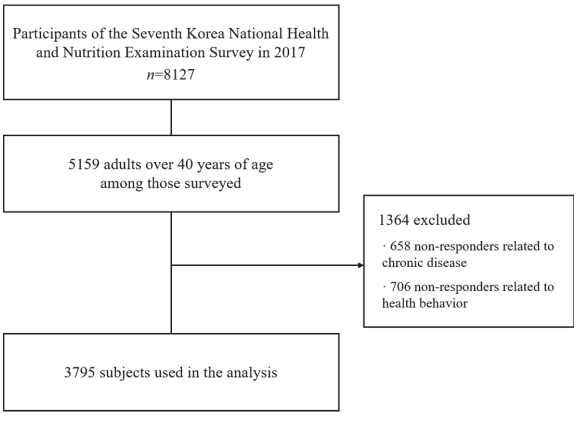


Figure 1. Flowchart of the study.

2.2. Description of Variables

In this study, questions related to sociodemographic characteristics, morbidity, and presence of chronic diseases, and health behaviors were utilized. Sex, age, education level, level of income, type of health insurance, and private insurance policy were used as sociodemographic variables. In terms of age, participants aged 19 years or older were divided into two groups: adults and older adults (aged 65 years and above). The education level was stratified into middle school graduates or less and high school graduates or more. Income status was determined by the monthly mean household gross income and was classified based on 3 million won as the cut-off point. The patients with health insurance were classified as health insurance subscribers and medical care beneficiaries [23,25–28].

Hypertension and diabetes are the main causes of cardiovascular disease, and the number of patients continues to rise due to the increase in obesity rate. In addition, the cost of medical care is proliferating more rapidly than the number of patients. Therefore, it is significant to prevent it by analyzing the factors influencing chronic diseases. Hitherto, chronic disease was defined as hypertension, dyslipidemia, stroke, myocardial infarction, and diabetes. The presence of chronic disease was determined based on a response of "Yes" to the question related to a doctor's diagnosis. Health behaviors included health checkups, drinking, smoking, exercise, obesity, and hypercholesterolemia [28–31]. Health checkup status was classified as patients who underwent health checkups and those who did not undergo health checkups. Drinking status was classified as non-drinkers and drinkers based on the monthly drinking rate; smoking status was classified as non-smokers and smokers using the current smoking rate. Exercise history was stratified as those who performed exercises and those who did not perform exercises based on the aerobic physical activity practice rate [32,33]. Furthermore, the prevalence of obesity was determined and obesity was stratified based on the following indices: a body mass index of 18.5 kg/m^2 or higher or a body mass index of 23 kg/m^2 or lower indicates normal weight, while a body mass index of 25 kg/m^2 or higher indicates obesity [34–36]. Hypercholesterolemia was stratified based on its prevalence (Table 1).

Table 1. Classification and definition of variables.

Variable	Definition
Sex	0 = Female 1 = Male
Age	0 = Adult 1 = Senior
Education	0 = ≤Middle school 1 = ≥Middle school
Income	0 = <300 1 = ≥300
Type of insurance	0 = National Health Insurance 1 = Assistance
Private insurance	0 = No 1 = Yes
Chronic disease	0 = No 1 = Yes
Medical checkup	0 = No 1 = Yes
Drinking	0 = Non-drinker 1 = Drinker
Smoking	0 = Non-smoker 1 = Smoker

Table 1. *Cont.*

Variable	Definition
Exercise	0 = No 1 = Yes
Obesity	0 = Obesity 1 = Normal
Hypercholesterolemia	0 = No 1 = Yes

2.3. Statistical Analysis

In order to analyze the relationship among sociodemographic characteristics, health behaviors, and the presence of chronic diseases, statistical analyses were conducted using the SPSS (version 25.0, https://www.ibm.com/kr-ko/analytics/spss-statistics-software).

First, cross-analysis was performed to analyze the relationship between chronic diseases and sociodemographic characteristics and between health behaviors and chronic diseases. In order to determine the relationship between sociodemographic characteristics and health behaviors and the risk for developing chronic diseases, a logistic regression analysis was performed.

3. Results

3.1. Participants' Demographic Characteristics

Of the total participants, 56% were women and the proportion of women was higher than that of men; older adults aged 65 years or higher accounted for 33% of the total study population. Most of the participants were middle school graduates or had obtained higher education (2269 persons, 59.8%) and had an income of more than 3 million won (2107 persons, 55.5%). With regard to the type of health insurance, national health insurance subscribers accounted for 95.7% of the total participants according to the characteristics of health insurance in Korea. Meanwhile, private insurance subscribers accounted for 74.2%, even though the proportion of health insurance subscribers corresponded to a majority; this finding indicates that most of the patients took a private medical insurance policy due to the lack of coverage by the national health insurance. A total of 857 patients (22.6%) underwent medical checkups, which suggests that only a few patients were able to undergo medical checkups. A total of 1762 patients (46.4%) developed chronic diseases, of whom 21.4% had two or more chronic diseases (Table 2).

Table 2. Demographic characteristics (n = 3795).

Characteristic	Type	N	%
Sex	Female	2126	56.0
	Male	1669	44.0
Age	40–65	2544	67.0
	≥65	1251	33.0
Education	<Middle school	1526	40.2
	≥Middle school	2269	59.8
Income	<300	1688	44.5
	≥300	2107	55.5
Type of insurance	National health insurance	3632	95.7
	Assistance	163	4.3
Private insurance	Y	2814	74.2
	N	981	25.8

Table 2. *Cont.*

Characteristic	Type	N	%
Medical checkup	Y	857	22.6
	N	2938	77.4
Chronic disease	0	2033	53.6
	1	950	25.0
	<2	812	21.4

3.2. Relationship between Demographic Characteristics and Chronic Diseases

In this study, we intended to analyze the relationship between sociodemographic characteristics and chronic diseases, and the results are shown in Table 3. Among chronic disease patients, 955 (44.9%) were women, this proportion being higher than that of men. Meanwhile, 807 (48.4%) of 862 male participants had chronic diseases, which indicates that men had a higher rate of chronic disease morbidity. Approximately 71.3% of the participants aged 65 years or higher had chronic diseases. In addition, most of the patients with a lower educational level and lower-income level had chronic diseases (981 patients (64.3%)and 1008 patients (59.7%), respectively). A total of 1648 (45.4%) health insurance subscribers were chronic disease patients, while 114 (69.9%) medical care recipients were chronic disease patients. Furthermore, 1128 (40.1%) chronic disease patients were private insurance subscribers.

Table 3. Relationship between demographic characteristics and chronic diseases.

Characteristic	Type	Chronic Disease				*p*-Value
		N	%	Y	%	
Sex **	Female	1171	55.1	955	44.9	0.019
	Male	862	51.6	807	48.4	
Age ***	40–65	1674	65.8	870	34.2	0.001
	≥65	359	28.7	892	71.3	
Education ***	<Middle school	545	35.7	981	64.3	0.001
	≥Middle school	1488	65.6	781	34.4	
Income ***	<300	680	40.3	1008	59.7	0.001
	≥300	1353	64.2	754	35.8	
Type of insurance ***	NHI	1984	54.6	1648	45.4	0.001
	Assistance	49	30.1	114	69.9	
Private insurance ***	Y	1686	59.9	1128	40.1	0.001
	N	347	35.4	634	64.6	

** $p < 0.05$, *** $p < 0.001$.

3.3. Relationship between Health Behavior and Chronic Diseases

We analyzed the relationship between health behaviors and chronic diseases; the results are shown in Table 4. Of the total chronic disease patients, 944 (50.6%) were alcohol drinkers, 1503 (47%) were smokers, and 605 (40.1%) performed exercises, which is less than the number of patients who did not perform exercises (1157, 50.6%). Moreover, 1263 (52.7%) and 823 (73.9%) patients with obesity and hypercholesterolemia, respectively, had chronic diseases.

Table 4. Relationship between health behavior and chronic diseases.

Characteristic	Type	Chronic Disease				*p*-Value
		N	%	Y	%	
Medical checkup	Y	453	52.9	404	47.1	0.331
	N	1580	53.8	1358	46.2	
Drinking ***	Y	923	49.4	944	50.6	0.001
	N	1110	57.6	818	42.4	
Smoking *	Y	1696	53.0	1503	47.0	0.062
	N	337	56.5	259	43.5	
Exercise ***	Y	904	59.9	605	40.1	0.001
	N	1129	49.4	1157	50.6	
Obesity ***	Y	1132	47.3	1263	52.7	0.001
	N	901	64.4	499	35.6	
Hypercholesterolemia ***	Y	291	26.1	823	73.9	0.001
	N	1742	65.0	939	35.0	

* $p < 0.1$, *** $p < 0.001$.

3.4. Factors Affecting Chronic Diseases

In order to determine the factors that may affect the development of chronic diseases, logistic regression analysis was performed, and the results are shown in Table 5. The factors with statistically significant effects in patients with chronic disease included sex, age, education, income, types of health insurance, decision to take a private insurance policy, health checkups, exercise, obesity, and hypercholesterolemia.

Table 5. Factors affecting the development of chronic diseases.

Dependent Variable	Independent Variable	Exp(B)	*p*-Value
Chronic disease	Sex ***	1.498	0.001
	Age ***	3.145	0.001
	Education ***	0.535	0.001
	Income **	0.773	0.004
	Type of insurance **	1.727	0.008
	Private insurance **	0.803	0.036
	Medical checkup **	0.782	0.009
	Drinking	1.101	0.252
	Smoking	1.061	0.606
	Exercise *	0.861	0.060
	Obesity ***	0.544	0.001
	Hypercholesterolemia ***	5.444	0.001

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.001$.

In men, the risk of developing chronic diseases was higher by 1.498 times. Further, as age increased, the risk of developing chronic diseases also increased by 3.145 times. In participants with a higher education level, the risk of developing chronic diseases increased by 0.535 times. In participants with higher income, the risk of developing chronic disease reduced by 0.773 times. With regard to the type of health insurance, the risk of developing chronic diseases increased by 1.727 times among medical care beneficiaries. In addition, for those who took a private insurance policy, the risk of developing chronic diseases increased by 0.782 times. Meanwhile, the risk of developing chronic diseases decreased by 0.782 times and 0.861 times among those who underwent medical checkups and who performed exercises, respectively. In normal-weight people, the risk of developing chronic diseases reduced by 0.544 times. In patients with hypercholesterolemia, the risk increased by 5.444 times.

4. Discussion

In this study, we analyzed the factors affecting the development of chronic diseases through logistic regression analysis using the data from the Korea National Health and Nutrition Examination Survey (2017). Of the sociodemographic characteristics, sex, age, education and income level, types of health insurance, and private insurance were found to have an effect on chronic diseases. In terms of sex, the proportion of women with chronic diseases was higher than that of men. Compared with women, men had a higher rate of chronic disease morbidity and the risk of developing chronic diseases. These results are inconsistent with those of previous studies, which reported that the prevalence of chronic diseases is higher among women than in men because men can maintain their economic level for longer than women. Women who have a lower income level than men have relatively low medical accessibility and find it difficult to manage their chronic diseases [37]. The number of chronic disease patients is increasing due to the lack of physical activity and the increasing prevalence of hypercholesterolemia and obesity, and considering that previous studies have shown that the prevalence of chronic disease was lower among men who received management, managing chronic diseases according to sex seems to be of utmost importance [38]. In addition, the number of patients aged 65 years or older who had chronic diseases was higher; therefore, the higher the age, the higher the risk of developing chronic diseases. This finding is consistent with those of a previous study, which reported that as age increases, the prevalence of chronic diseases also increases due to the decreased amount of physical activities and habit-based health risk behaviors [8,39].

It was also found that the higher the income and education levels, the lower the risk of chronic diseases. This finding is consistent with those of previous studies reporting that socioeconomic status, including income, education, and occupation levels, affects the health-related lifestyles and risk of chronic diseases [40]. Because of the low rates of physical activity and exercise practice and as the provision of medical services for managing chronic diseases has still not been ensured owing to lower educational levels or living standards, the prevalence of chronic diseases is increasing. Among medical care beneficiaries, the risk of developing chronic diseases was high, which was similar to the results of a previous study reporting that the incidence of chronic disease increased among individuals who belonged to the lower social class, like those in the low-income bracket. Social determinants, such as income, education, and social class, may cause health-related inequality but create an environment in which quality medical care can be provided for the treatment of chronic diseases. In addition, non-medical factors, such as social determinants, play a more substantial role in the management of chronic diseases than medical factors. It seems that medical care beneficiaries with low income may have more difficulty in managing chronic diseases [41–43]. There were many chronic disease patients who obtained a private medical insurance policy; the results showed that patients with private medical insurance had a lower risk of developing chronic diseases. These findings are similar to those of a previous study, which indicated that those who have private medical insurance policies tend to receive outpatient and inpatient treatments. In line with these findings, among patients with chronic diseases who require continuous health care, those with private medical insurance have a reduced burden in terms of medical expenses, leading to better health-care outcomes [44–46]. Considering these results, there are limitations in managing chronic diseases with national health insurance only. Furthermore, it is estimated that people obtain commercial medical insurance policies due to the burden of medical expenses caused by the recent increase in polychronic diseases. Therefore, since health-related inequalities in the low-income group patients, who find it difficult to pay the private medical insurance premiums, will become a serious problem if we only rely on private medical insurance for the management of chronic diseases, the coverage of the national health insurance should be reinforced for the management of chronic diseases.

Among health behaviors, the factors affecting the risk of developing chronic diseases included health checkups, exercise, obesity, and hypercholesterolemia. Those who underwent periodic health checkups had a risk of developing chronic diseases, which is similar to previous findings showing that periodic health checkups promote health and help prevent chronic diseases [8,47]. In addition,

considering the results of previous studies reporting that those who benefit from health insurance are more likely to receive health checkups depending on the nature of the health insurance system in Korea, chronic diseases could be effectively managed through modifying the nature of the insurance provided. Previous studies have shown that health behavior factors related to chronic diseases include smoking, drinking, exercise, body mass index, and regular life and eating habits [7,29,36,48,49]. However, in this study, drinking and smoking did not have a statistically significant effect on the prevalence of chronic diseases, and these results are different from those of existing research. Furthermore, exercise, obesity, and hypercholesterolemia were associated with the risk of developing chronic diseases, consistent with existing research. Among those who performed exercises, the risk of developing chronic diseases was lower, while among those with obesity and hypercholesterolemia, the risk of developing chronic diseases was higher. Weight loss via exercise programs reduces the risk of developing chronic diseases. Maintaining a standard body weight can prevent chronic diseases by alleviating hypercholesterolemia. Management of chronic diseases should be comprehensively performed with weight management through exercise; however, there seems to be a limitation in this regard according to patients' behavioral changes [50,51]. In order to overcome this limitation, wearable medical devices, which use ICT (Information & Communication Technology), have recently been developed for chronic disease management. Prevention and management of chronic diseases can be ensured through exercise [52–55]. The use of medical devices to promote physical activity leads to obesity and hypercholesterolemia management, and through the linkage between these medical devices and local clinic-centered, effective management of chronic diseases can be achieved through periodic monitoring. The results of this study also suggest that gender, age, education, and income levels have impacts on chronic disease, and it is significant to add these as risk factors and to continue monitoring in local clinic-centered facilities. Through this, a personalized chronic disease management system could be established.

This study has some limitations. First, chronic disease patients aged 40 years or below were not included. Recently, the number of younger chronic disease patients has increased owing to changes in lifestyle, therefore, further studies to analyze the factors influencing the risk of developing chronic diseases in this age group will be required, with the patients stratified as follows: youth, middle-aged, and older adults. Second, analyses according to the number of chronic diseases were not performed. In this study, only the presence or absence of chronic diseases in patients was assessed. Further studies to determine the influencing factors according to the number of chronic diseases are required. Third, there was no analysis of factors affecting chronic disease according to the residential area. Accessibility to medical services varies depending on where you live; therefore, chronic disease management may be different. Hence, it is necessary to analyze the factors affecting chronic diseases according to urban and rural areas. Despite these limitations, we comprehensively analyzed the factors influencing the prevalence of chronic diseases. Our study is significant as we were able to determine the risk factors for chronic diseases, which can be used as a basis for developing policies for the comprehensive management of chronic diseases, based on sex, age, and social factors.

5. Conclusions

In order to manage chronic diseases, the management approach should be based on patients' socioeconomic characteristics to address the differences related to sex, education, income, and medical care. The management should also include approaches to improve health behaviors, including the use of wearable medical devices and digital healthcare products. Based on our findings, we presume that chronic diseases develop due to a combination of factors. Age, socioeconomic factors, obesity, and hypercholesterolemia are factors that can be controlled to prevent and manage chronic diseases through comprehensive programs rather than through individual management. Moreover, those who belong to the lower social class, are more likely to require chronic disease management via primary healthcare institutions in the community. In order to improve health behaviors, continuous observation is required, and local clinic-centered chronic disease management can help improve health

behaviors. It is significant to establish a comprehensive management system and promote efficient medical delivery systems for chronic diseases focused on local clinic-centered facilities. However, Korea's medical delivery system urgently needs reorganization due to the concentration of university hospitals and the weakening of a local clinic-centered structure. Therefore, in order to expand the role of local clinic-centered facilities and to efficiently manage chronic diseases, the integrated local clinic-centered care chronic disease management project is being implemented. Through this, medical treatment for chronic disease management and education for improving lifestyle, are applied to lower the patient's copayment. If the burden reduction of chronic disease management is expanded, the dependency on private health insurance will be reduced, which will prevent excessive medical expenses for chronic patients. In addition, strengthening the role of local clinic-centered facilities will lead to strengthening medical access for low-income people, thereby relieving health inequalities. For older adults, when included in the community care project in line with community-based primary healthcare service, comprehensive management of chronic diseases, including health improvement and lifestyle modification, could be implemented. In particular, in Europe, where public health policies are in place, chronic diseases are effectively managed by strengthening the local clinic-centered services, such as the attending physician, to manage chronic diseases. For common goals such as chronic disease management, community care is implemented to ensure continuous health care. In view of this, chronic disease management through public health policy should be implemented prior to private medical insurance. Patients with private medical insurance have a lower risk of developing chronic diseases, but this can be seen as a problem of low insurance coverage for chronic diseases. This can be resolved through community care projects such as in Europe. Because of this, patient-centered chronic disease management will ultimately improve the health of chronic disease patients.

Author Contributions: Conceptualization, M.L.; methodology, S.P.; software, S.P.; validation, M.L. and K.-S.L.; formal analysis, M.L.; writing—original draft preparation, S.P.; writing—review and editing, M.L.; supervision, K.-S.L.; project administration, M.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2019S1A5A2A03040304).

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

References

1. Janowski, K.; Kurpas, D.; Kusz, J.; Mroczek, B.; Jedynak, T. Health-related behavior, profile of health locus of control and acceptance of illness in patients suffering from chronic somatic diseases. *PLoS ONE* **2013**, *8*, e63920. [CrossRef] [PubMed]
2. Mullen, P.D.; Hersey, J.C.; Iverson, D.C. Health behavior models compared. *Soc. Sci. Med.* **1987**, *24*, 973–981. [CrossRef]
3. Berrigan, D.; Dodd, K.; Troiano, R.P.; Krebs-Smith, S.M.; Barbash, R.B. Patterns of health behavior in U.S. adults. *Prev. Med.* **2003**, *36*, 615–623. [CrossRef]
4. McGinnis, J.M.; Williams-Russo, P.; Knickman, J.R. The case for more active policy attention to health promotion. *Health Aff.* **2002**, *21*, 78–93. [CrossRef]
5. Jarvandi, S.; Yan, Y.; Schootman, M. Income disparity and risk of death: The importance of health behaviors and other mediating factors. *PLoS ONE* **2012**, *7*, e49929. [CrossRef]
6. Nettle, D. Why are there social gradients in preventative health behavior? A perspective from behavioral ecology. *PLoS ONE* **2010**, *5*, e13371. [CrossRef]
7. Paudel, S.; Owen, A.J.; Owusu-Addo, E.; Smith, B.J. Physical activity participation and the risk of chronic diseases among South Asian adults: A systematic review and meta-analysis. *Sci. Rep.* **2019**, *9*, 9771. [CrossRef]
8. Strong, K.; Mathers, C.; Leeder, S.; Beaglehole, R. Preventing chronic diseases: How many lives can we save? *Lancet* **2005**, *366*, 1578–1582. [CrossRef]
9. Yach, D.; Hawkes, C.; Gould, C.L.; Hofman, K.J. The Global Burden of Chronic DiseasesOvercoming Impediments to Prevention and Control. *JAMA* **2004**, *291*, 2616–2622. [CrossRef]

10. Bae, E.-J.; Park, N.-J.; Sohn, H.-S.; Kim, Y.-H. Handgrip strength and all-cause mortality in middle-aged and older Koreans. *Int. J. Environ. Res. Public Health* **2019**, *16*, 740. [CrossRef]
11. Gansevoort, R.T.; Correa-Rotter, R.; Hemmelgarn, B.R.; Jafar, T.H.; Heerspink, H.J.L.; Mann, J.F.; Matsushita, K.; Wen, C.P. Chronic kidney disease and cardiovascular risk: Epidemiology, mechanisms, and prevention. *Lancet* **2013**, *382*, 339–352. [CrossRef]
12. Halpin, H.A.; Morales-Suárez-Varela, M.M.; Martin-Moreno, J.M. Chronic Disease Prevention and the New Public Health. *Public Health Rev.* **2010**, *32*, 120–154. [CrossRef]
13. Sturm, R. The effects of obesity, smoking, and drinking on medical problems and costs. *Health Aff.* **2002**, *21*, 245–253. [CrossRef] [PubMed]
14. Strine, T.W.; Chapman, D.P.; Balluz, L.S.; Moriarty, D.G.; Mokdad, A.H. The associations between life satisfaction and health-related quality of life, chronic illness, and health behaviors among US community-dwelling adults. *J. Community Health* **2008**, *33*, 40–50. [CrossRef]
15. Harris, M. The role of primary health care in preventing the onset of chronic disease, with a particular focus on the lifestyle risk factors of obesity, tobacco and alcohol. *Canberra Natl. Prev. Health Taskforce* **2008**, *1*, 21.
16. Steyn, K.; Damasceno, A. Lifestyle and related risk factors for chronic diseases. *Dis. Mortal. Sub-Sahar. Afr.* **2006**, *2*, 247–265.
17. Reeves, M.J.; Rafferty, A.P. Healthy lifestyle characteristics among adults in the United States, 2000. *Arch. Intern. Med.* **2005**, *165*, 854–857. [CrossRef]
18. Kang, Y.J.; Kang, M.-Y. Chronic diseases, health behaviors, and demographic characteristics as predictors of ill health retirement: Findings from the Korea Health Panel Survey (2008–2012). *PLoS ONE* **2016**, *11*, e0166921. [CrossRef]
19. Winkleby, M.A.; Cubbin, C. Changing patterns in health behaviors and risk factors related to chronic diseases, 1990–2000. *Am. J. Health Promot.* **2004**, *19*, 19–27. [CrossRef]
20. Speake, D.L.; Cowart, M.E.; Pellet, K. Health perceptions and lifestyles of the elderly. *Res. Nurs. Health* **1989**, *12*, 93–100. [CrossRef]
21. Cho, Y.I.; Lee, S.-Y.D.; Arozullah, A.M.; Crittenden, K.S. Effects of health literacy on health status and health service utilization amongst the elderly. *Soc. Sci. Med.* **2008**, *66*, 1809–1816. [CrossRef] [PubMed]
22. Lee, T.W.; Ko, I.S.; Lee, K.J. Health promotion behaviors and quality of life among community-dwelling elderly in Korea: A cross-sectional survey. *Int. J. Nurs. Stud.* **2006**, *43*, 293–300. [CrossRef] [PubMed]
23. Winkleby, M.A.; Jatulis, D.E.; Frank, E.; Fortmann, S.P. Socioeconomic status and health: How education, income, and occupation contribute to risk factors for cardiovascular disease. *Am. J. Public Health* **1992**, *82*, 816–820. [CrossRef] [PubMed]
24. Innes, H.; McAuley, A.; Alavi, M.; Valerio, H.; Goldberg, D.; Hutchinson, S.J. The contribution of health risk behaviors to excess mortality in American adults with chronic hepatitis C: A population cohort-study. *Hepatology* **2018**, *67*, 97–107. [CrossRef] [PubMed]
25. Lim, J.; Kim, S.; Ke, S.; Cho, B. The Association Chronic Liver Diseases with Health Related Behaviors in South Korea. *Korean J. Fam. Med.* **2010**, *31*, 302–307. [CrossRef]
26. Caspersen, C.J.; Pereira, M.A.; Curran, K.M. Changes in physical activity patterns in the United States, by sex and cross-sectional age. *Med. Sci. Sports Exerc.* **2000**, *32*, 1601–1609. [CrossRef]
27. Chang, O.; Choi, E.-K.; Kim, I.-R.; Nam, S.-J.; Lee, J.E.; Lee, S.K.; Im, Y.-H.; Park, Y.H.; Cho, J. Association between socioeconomic status and altered appearance distress, body image, and quality of life among breast cancer patients. *Asian Pac. J. Cancer Prev.* **2014**, *15*, 8607–8612. [CrossRef]
28. Fine, L.J.; Philogene, G.S.; Gramling, R.; Coups, E.J.; Sinha, S. Prevalence of multiple chronic disease risk factors: 2001 National Health Interview Survey. *Am. J. Prev. Med.* **2004**, *27*, 18–24. [CrossRef]
29. Lee, M.; Yoon, K. Catastrophic Health Expenditures and Its Inequality in Households with Cancer Patients: A Panel Study. *Processes* **2019**, *7*, 39. [CrossRef]
30. McCann, B.A.; Ewing, R. Measuring the Health Effects of Sprawl: A National Analysis of Physical Activity, Obesity and Chronic Disease. *Geography* **2003**. [CrossRef]
31. Katzmarzyk, P.; Lear, S. Physical activity for obese individuals: A systematic review of effects on chronic disease risk factors. *Obes. Rev.* **2012**, *13*, 95–105. [CrossRef] [PubMed]
32. Mukamal, K.J.; Ding, E.L.; Djoussé, L. Alcohol consumption, physical activity, and chronic disease risk factors: A population-based cross-sectional survey. *BMC Public Health* **2006**, *6*, 118. [CrossRef] [PubMed]

33. Cho, E.R.; Shin, A.; Kim, J.; Jee, S.H.; Sung, J. Leisure-time physical activity is associated with a reduced risk for metabolic syndrome. *Ann. Epidemiol.* **2009**, *19*, 784–792. [CrossRef] [PubMed]
34. Choban, P.; Flancbaum, L. Nourishing the obese patient. *Clin. Nutr.* **2000**, *19*, 305–311. [CrossRef]
35. Chan, M.; Lim, Y.; Ernest, A.; Tan, T. Nutritional assessment in an Asian nursing home and its association with mortality. *J. Nutr. Health Aging* **2010**, *14*, 23–28. [CrossRef]
36. Kim, Y.-H.; Kim, S.M.; Han, K.-D.; Jung, J.-H.; Lee, S.-S.; Oh, S.W.; Park, H.S.; Rhee, E.-J.; Lee, W.-Y.; Yoo, S.J. Waist Circumference and All-Cause Mortality Independent of Body Mass Index in Korean Population from the National Health Insurance Health Checkup 2009–2015. *J. Clin. Med.* **2019**, *8*, 72. [CrossRef]
37. Virtanen, M.; Oksanen, T.; Batty, G.D.; Ala-Mursula, L.; Salo, P.; Elovainio, M.; Pentti, J.; Lybäck, K.; Vahtera, J.; Kivimäki, M. Extending employment beyond the pensionable age: A cohort study of the influence of chronic diseases, health risk factors, and working conditions. *PLoS ONE* **2014**, *9*, e88695. [CrossRef]
38. Park, M.; Park, Y.; Kim, S.; Park, S.; Seol, H.; Woo, S.; Cho, S.; Lim, D.-S. Introduction and effectiveness of The Seoul Metabolic Syndrome Management. *Public Health Aff.* **2017**, *1*, 25–39. [CrossRef]
39. Zhao, C.; Wong, L.; Zhu, Q.; Yang, H. Prevalence and correlates of chronic diseases in an elderly population: A community-based survey in Haikou. *PLoS ONE* **2018**, *13*, e0199006. [CrossRef]
40. Cockerham, W.C.; Hamby, B.W.; Oates, G.R. The social determinants of chronic disease. *Am. J. Prev. Med.* **2017**, *52*, S5–S12. [CrossRef]
41. Martínez-García, M.; Salinas-Ortega, M.; Estrada-Arriaga, I.; Hernández-Lemus, E.; García-Herrera, R.; Vallejo, M. A systematic approach to analyze the social determinants of cardiovascular disease. *PLoS ONE* **2018**, *13*, e0190960. [CrossRef] [PubMed]
42. Taylor, L.A.; Tan, A.X.; Coyle, C.E.; Ndumele, C.; Rogan, E.; Canavan, M.; Curry, L.A.; Bradley, E.H. Leveraging the social determinants of health: What works? *PLoS ONE* **2016**, *11*, e0160217. [CrossRef] [PubMed]
43. Nicholas, S.B.; Kalantar-Zadeh, K.; Norris, K.C. Socioeconomic disparities in chronic kidney disease. *Adv. Chronic Kidney Dis.* **2015**, *22*, 6–15. [CrossRef] [PubMed]
44. Jeon, B.; Kwon, S. Effect of private health insurance on health care utilization in a universal public insurance system: A case of South Korea. *Health Policy* **2013**, *113*, 69–76. [CrossRef] [PubMed]
45. Lee, M.; Yoon, K.; Choi, M. Private health insurance and catastrophic health expenditures of households with cancer patients in South Korea. *Eur. J. Cancer Care* **2018**, *27*, e12867. [CrossRef]
46. Mulcahy, A.W.; Eibner, C.; Finegold, K. Gaining coverage through Medicaid or private insurance increased prescription use and lowered out-of-pocket spending. *Health Aff.* **2016**, *35*, 1725–1733. [CrossRef]
47. Culica, D.; Rohrer, J.; Ward, M.; Hilsenrath, P.; Pomrehn, P. Medical checkups: Who does not get them? *Am. J. Public Health* **2002**, *92*, 88–91. [CrossRef]
48. Liu, Y.; Croft, J.B.; Wheaton, A.G.; Kanny, D.; Cunningham, T.J.; Lu, H.; Onufrak, S.; Malarcher, A.M.; Greenlund, K.J.; Giles, W.H. Clustering of Five Health-Related Behaviors for Chronic Disease Prevention Among Adults, United States, 2013. *Prev. Chronic Dis.* **2016**, *13*, E70. [CrossRef]
49. Campbell, D.J.; Ronksley, P.E.; Manns, B.J.; Tonelli, M.; Sanmartin, C.; Weaver, R.G.; Hennessy, D.; King-Shier, K.; Campbell, T.; Hemmelgarn, B.R. The association of income with health behavior change and disease monitoring among patients with chronic disease. *PLoS ONE* **2014**, *9*, e94007. [CrossRef]
50. Thande, N.K.; Hurstak, E.E.; Sciacca, R.E.; Giardina, E.G.V. Management of obesity: A challenge for medical training and practice. *Obesity* **2009**, *17*, 107–113. [CrossRef]
51. Lee, M.; Yoon, K.; Lee, K.-S. Subjective health status of multimorbidity: Verifying the mediating effects of medical and assistive devices. *Int. J. Equity Health* **2018**, *17*, 164. [CrossRef] [PubMed]
52. Zhu, X.; Cahan, A. Wearable technologies and telehealth in Care Management for Chronic Illness. In *Healthcare Information Management Systems*; Springer: Cham, Switzerland, 2016; pp. 375–398.
53. Kang, S.Y. The ICT Technology for Geriatric Diseases Healthcare. In *Information Science and Applications (ICISA) 2016*; Springer: Singapore, 2016; pp. 1495–1500.
54. Bodenheimer, T.; Lorig, K.; Holman, H.; Grumbach, K. Patient self-management of chronic disease in primary care. *JAMA* **2002**, *288*, 2469–2475. [CrossRef] [PubMed]
55. Chiauzzi, E.; Rodarte, C.; DasMahapatra, P. Patient-centered activity monitoring in the self-management of chronic health conditions. *BMC Med.* **2015**, *13*, 77. [CrossRef] [PubMed]

Journal of
Clinical Medicine

MDPI

Article

Decision Support for the Optimization of Provider Staffing for Hospital Emergency Departments with a Queue-Based Approach

Fuu-Cheng Jiang [1,†], Cheng-Min Shih [2,4,†], Yun-Ming Wang [2], Chao-Tung Yang [1,*], Yi-Ju Chiang [3] and Cheng-Hung Lee [4,5]

1 Department of Computer Science, Tunghai University, No. 1727, Section 4, Taiwan Boulevard, Taichung 40704, Taiwan; admor@thu.edu.tw
2 Department of Biological Science and Technology, National Chiao Tung University, No. 1001, Daxue Rd., Hsinchu 30010, Taiwan; 10chengmin@gmail.com (C.-M.S.); ymwang@mail.nctu.edu.tw (Y.-M.W.)
3 Information and Communication Research Division, National Chung-Shan Institute of Science & Technology, Taoyuan 32546, Taiwan; yjchiang0320@gmail.com
4 Department of Orthopedics, Taichung Veterans General Hospital, No. 1650, Section 4, Taiwan Boulevard, Taichung 40705, Taiwan; leechenghung0115@gmail.com
5 Department of Food Science and Technology, Hung Kuang University, No. 1018, Section 6, Taiwan Boulevard, Taichung 43302, Taiwan
* Correspondence: ctyang@thu.edu.tw
† These authors contributed equally.

Received: 6 October 2019; Accepted: 24 November 2019; Published: 5 December 2019

Abstract: Deployment or distribution of valuable medical resources has emerged as an increasing challenge to hospital administrators and health policy makers. The hospital emergency department (HED) census and workload can be highly variable. Improvement of emergency services is an important stage in the development of the healthcare system and research on the optimal deployment of medical resources appears to be an important issue for HED long-term management. HED performance, in terms of patient flow and available resources, can be studied using the queue-based approach. The kernel point of this research is to approach the optimal cost on logistics using queuing theory. To model the proposed approach for a qualitative profile, a generic HED system is mapped into the M/M/R/N queue-based model, which assumes an R-server queuing system with Poisson arrivals, exponentially distributed service times and a system capacity of N. A comprehensive quantitative mathematical analysis on the cost pattern was done, while relevant simulations were also conducted to validate the proposed optimization model. The design illustration is presented in this paper to demonstrate the application scenario in a HED platform. Hence, the proposed approach provides a feasibly cost-oriented decision support framework to adapt a HED management requirement.

Keywords: hospital emergency department; queuing theory; decision support; cost optimization

1. Introduction

1.1. Background

Hospitals play an important role in the healthcare system of society. They have changed rapidly in parallel with improvements in medical instruments and medicine. Administrators of hospital institutions should have spared no effort in developing strategies that enable the provision of high-quality services, while operating with increased costs and pressure from competition [1]. One of the most demanding departments in terms of economic resource consumption and programming is the

hospital emergency department (HED). To this extent its operational profile should be monitored and optimized in order to provide the optimal quality of medical service subject to the budget constraint.

HEDs must be operational 24 h, and, moreover, should respond to multiple demands requiring sophisticated technical equipment and the manpower to operate them, all of which imply higher costs. Large HEDs have even higher costs because they offer a wide range of services that would be unavailable in a small rural HED [2]. Emergency Departments have traditionally been a crucial issue in the hospitals' cost containment and management. The optimization of patient flow and bottleneck elimination in key departments could be a viable way for policy makers to decrease operational costs and boost quality of care [3]. In the interest of improving patient throughput and resource utilization, appropriate key performance measures are selected, like the size of staffing providers, HED patient arrival patterns, service rate of staffing providers, waiting time, etc. To explore the tradeoff among them, the proposed queue-based optimization technique on cost may provide hospital management with a decision support for deploying staffing providers under the constraints of the kernel performance parameters.

An M/M/R/N queuing model was adopted to explore the cost profile from analyzing the relationships among relevant performance parameters in a HED, such as the number of staffing providers (i.e., servers) needed during each staffing interval. This model assumes a single queue with a limited system capacity of N that feeds into R identical medical servers (i.e., staffing providers). The fourth symbol "N" of the notation M/M/R/N indicates the restriction on the system capacity of the HED. The value of (N−R) gives the number of waiting rooms for incoming patients when all staffing providers in the HED are busy. Arrivals occur according to a time-homogeneous Poisson process with a constant rate, and the service duration (e.g., provider time associated with a patient) has an exponential distribution. In the language of queuing theory [4,5], these two assumptions are often called Markovian, hence the use of the two "M's" in the notation used for the model. One advantage of adopting the M/M/R/N model is that given an arrival rate, an average service duration, and the number of servers, formulae for performance measures such as the cost profile, the average number of patients, or the average waiting time can be easily obtained.

1.2. Contribution Profile

This novel idea in this work originated from the theory of an M/M/R/N queuing system (QS), which is used to estimate the optimal number of providers needed during each staffing interval [4,5]. At some pre-configured period (say a shift, or a day), a finite quantity of staffing providers exists to provide medical services for patients under limited waiting rooms in HEDs. On application modeling, such a finite quantity of staffing providers (i.e., R) can be regarded as the term "server" in the M/M/R/N model of queuing theory. The quantity of (N−R) can be considered to be the rather limited waiting rooms in HEDs regulated by each hospital.

The research goal for this work is to explore whether, with cost-based deployment, how many sets of staffing providers in the HED schedule would be optimal if a certain level of the server availability is kept? To explore the tradeoff between them, the proposed optimization technique may provide the HED management with decision support on the number of staffing providers. The key contributions of this paper are threefold: (1) This work provides HED administrators with an efficient deployment of staffing providers for the HED platform to optimize the cost improvement. In regards to management, the proposed system can be adopted as a decision-making methodology approaching predictive management, rather than reactive or chaotic management. (2) For quantitative analysis, the M/M/R/N queue model was applied and derived, and then the relevant system metrics were established in a brand-new manner. The mathematical expression of the cost function was established for the evaluation requirement. (3) In regards to verification, relevant experimental results were obtained in terms of configurations on cost optimization and average waiting time. The simulated results indicate that the proposed approach may provide a feasible decision support for deployment on quantities of standby servers.

The rest of the paper is organized as follows: Section 2 describes related work and the motivation behind this research. To demonstrate the framework qualitatively, an M/M/R/N model of queuing theory is adopted and the mapping profile is demonstrated in Section 3. Quantitative work is presented in Section 4, where the mathematical analysis is conducted in detail and the relevant system performance measures, such as the expected number of online servers, the expected number of spares, etc., are derived. Following this, in Section 5, the queue-based model is further addressed in terms of cost function, and the simulations of the feasibility of the proposed scheme are conducted. Finally, some concluding remarks are made in Section 6.

2. Related Work

HED crowding represents an important issue that may affect the quality and access of health care. Accordingly, the optimization of average waiting times has become a focus across many mainstream hospitals. As defined by the Canadian Association of Emergency Physicians [6], HED overcrowding is a situation in which the demand for services exceeds the ability of health care professionals to provide care within a reasonable length of time. As stated in [7], significant variation in HED patient arrival rates necessitates the adjustment of staffing patterns to optimize the timely care of patients. Green et al. [7] collected detailed HED arrival data from an urban hospital and used a queue-based analysis to gain insights on how to change provider staffing to decrease the proportion of patients who leave without being seen. However, no optimization materials in terms of mathematical theory were addressed at all in these studies [8–10].

Finamore et al. [6] described an innovative use of a satellite clinic to prevent patients returning to HED for care on a scheduled basis. Their strategy allows patients returning for follow-up diagnostics or treatment to bypass the main HED. The proposed HED satellite clinic may shorten the waiting times in multiple ways, such as increasing the capacity to remove returning patients from the pool of patients requiring care in the HED, and creating a separate registration area and a separated staffed treatment area. The visit data in the HED were used to measure crowding and completion of waiting room time, treatment time, and boarding time for all patients treated and released or admitted to a single HED during 2010. In [11], the authors conducted a relevant statistical analysis and concluded that a HED census at arrival demonstrated variation in crowding exposure over time. In the work of Wiler et al. [12], the authors developed an agent-based simulation model for the evaluation of the FTS (fast track strategies) scheme applied in the HED to reduce patient waiting time. By and large, the issues regarding cost optimization on the HED management cost are not a concern for these open studies [5,11,12].

Vass and Szabo [13] evaluated 2195 questionnaires in the HED situated in Mures County, Romania, for a period three years (2010–2013). Their research reported that long waiting times were the most important complaint in patient's satisfaction surveys. To perceive the waiting times, only a specific M/M/3 queuing model was considered in their work to demonstrate the computation details. The work of [13] has motivated us to consider whether it is possible to provide an effective and feasible approach to decision support for the optimization of provider staffing under cost constraints for the HEDs with more elaborative queue-based frameworks. This research generalizes the queuing model of [13] into the M/M/R/N queuing framework in terms of three practical aspects: (1) Numbers of medical servers (provider staffing) can be configured to one of the system parameters instead of a fixed quantity. Such a dynamic staffing level enables a hospital to quantify the cost patterns and the alleviation of HED waiting times. (2) The space available in the HED would be limited for every hospital management. The fourth factor (N) in the notation of the M/M/R/N model symbolizes the fact that only N patients can be allowed to enter the waiting rooms of the HED in order not to exacerbate the issue of overcrowding. (3) The exact mathematical expressions would be derived in an elaborative manner and the relevant cost formulation would be used to provide the generic decision support for the hospital administrators.

3. The Proposed Model of Medical Emergency Services

3.1. The Generic Platform of Medical Emergency Services

This research explores feasible decision support that is proposed to optimize running costs under the constraint of the waiting time at a HED using queue-based models. The exemplified HED is in a metropolitan hospital (Taichung Veterans General Hospital or TVGH) located in central Taiwan. It began offering medical services on 16 September 1982. Since 1991, it has been accredited as a "Medical Center and First-Class Teaching Hospital" by the Department of Health, Taiwan. Taichung Veterans General Hospital is a 1500-bed hospital with up to 3900 employees. According to the latest statistical average data of registration accessed in TVGH, it offers a capacity of about 7000 outpatients and 190 patients in the emergency room daily [14]. As a public medical center, it provides safe, high-quality medical services with advanced facilities and training programs, as well as outstanding research and development programs.

The HED building, with eight floors at the TVGH (TVGH-ED), provides comprehensive emergency services 24 h a day. The functional deployment on the ground floor of the TVGH-ED building, as shown in Figure 1, is composed of different zones, including the Registration and Triage Area, Resuscitation Areas, Internal Medicine Areas, the Pediatric Treatment Area, Waiting Areas, Clinic Areas, Monitor Rooms and the Fever Screen Center, and relevant auxiliary service units such as the X-ray service and nursing stations. As this HED is a rather complicated service system due to random arrivals, various disease chains, uncertain service times of care, and randomness in human decision-making, it is difficult to model the whole HED with a single operational model. From the perspective of the model attribute [15], a generic operational model is defined as a formal description of operations performed to deliver a health service that is applicable over a wide range of health service delivery settings. For the sake of simplicity, this research concentrates the optimization issue on a specific platform of medical service, which is used hereafter to model staffing providers for a single disease chain.

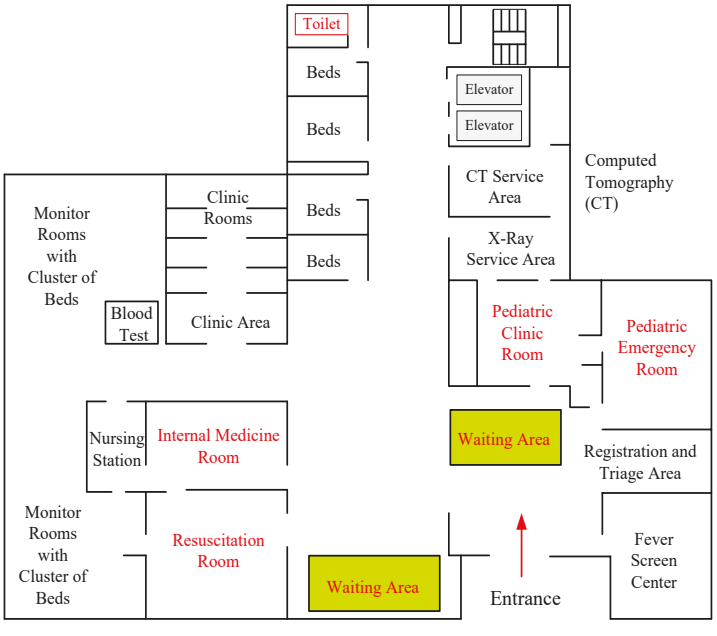


Figure 1. The functional deployment on the ground floor of the TVGH-ED building. TVGH-ED, Taichung Veterans General Hospital - Emergency Department.

3.2. Mapping Profile between the HED Service Platform and the M/M/R/N Queuing System

The proposed generic framework on the HED service platform is considered to be modeled as an M/M/R/N queuing system (QS), which is used to estimate the optimal number of providers needed during each staffing interval. An input-throughput-output framework of HED operations is used as the prototype shown in Figure 2 for a generic profile [16]. The ambulance icon symbolizes the arrivals of HED patients. Practically, patient arrivals are hard to schedule, or even control significantly. Arrivals may surge in some unpredictable time windows due to a short-term disaster, car accidents, and seasonal influences [17]. In modeling language, the busy and regular time windows can be associated with high and normal arrival rates, respectively. Patient arrivals in the proposed model are assumed to be Poisson processes [18], with average hourly rates that are forecasted for each future hour in question (say a shift, or a day) [19].

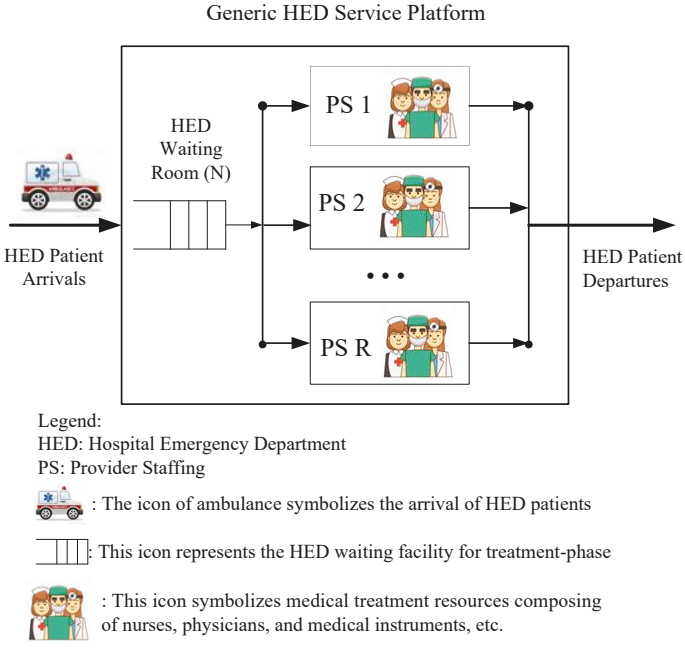


Figure 2. The generic service platform of an emergency department.

The itinerary for HED patients from arrival to exit can conceptually be divided into three phases [12]. The first phase, named "Waiting for treatment phase (*waiting phase*)", is symbolized by the icon HED Waiting Rooms in Figure 2. In the waiting phase, the patient goes through some standard processes that assist the HED to grasp the record of patients and their current medical condition. These are termed the Registration and Triage process, respectively. Registration guarantees administratively that patient demographics are captured accurately for billing and maintaining the record. Triage is the first assessment conducted by a healthcare professional after the patient arrives in the HED. The second phase (*treatment phase*) begins when the patient is placed in bed. For simplicity, the treatment phase is represented by the icons of provider staffing (medical servers) in Figure 3 for a generic profile. The whole medical service largely depends on patient acuity and physician activities. In modeling language, the duration of treatment can be regarded mathematically as the service rate of (medical) servers. The treatment phase is followed by the post-treatment phase, which is represented by the expression HED patient departures in Figure 3. Exiting from the treatment area of the HED

is reasonably assumed to mean that the patients are discharged, either as an outpatient or into the hospital [16].

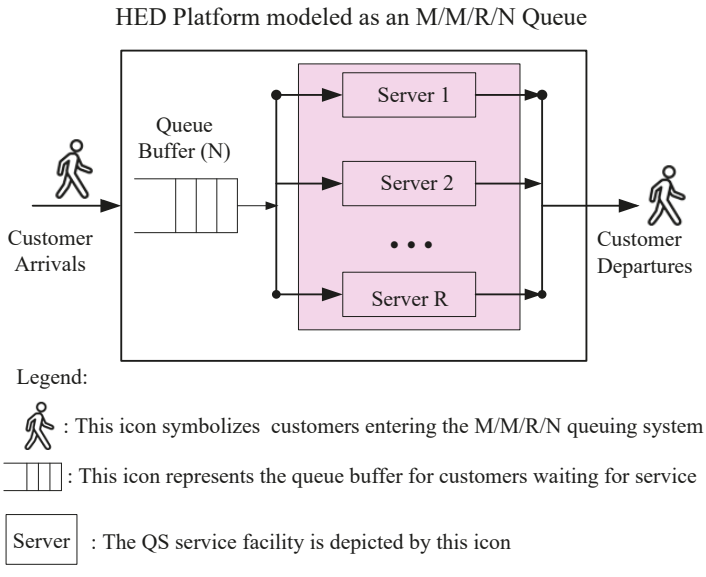


Figure 3. An M/M/R/N queue system mapped by the HED service platform.

The mapping scenarios for the theoretical approach are illustrated in Figures 2 and 3. An M/M/R/N queuing model was used to estimate the number of providers needed during each staffing time window. In Figure 3, the proposed model assumes a single queue with regulated and finite waiting rooms that feed into R identical servers with blue highlights, which is mapped to providers in the HED. The walking-man (customers) icons symbolize HED patient arrivals. Based on the proposed queuing model, relevant system metrics, such as average waiting times, expected number of customers in the queue buffer, and the probability that all servers are busy, can be analyzed and derived mathematically [7]. For instance, a patient's total length-of-stay from arrival to departure from the HED platform is termed as the patient throughput time, which is equivalent to the waiting time in the QS. Patient throughput time has a significant impact on operational and economic efficiency as well as overall patient satisfaction, which is a measure of medical service quality [20].

Generally, the performance metric on average waiting times may provide the HED administrator with decision support on how to alleviate patients' complaints. To avoid the deterioration of average patient throughput time (i.e., the average wait times in the QS), the optimization approach on the average waiting times, under some constraints such as a limited number of servers in the QS (i.e., mapped counterpart: level of staffing in the HED platform), is explored further in this article. The metric is the probability that all servers can be used to reveal the possibility and scenario in which notorious HED crowding may occur. The question is how to reduce this HED crowding phenomenon in some specific time windows. Such a metric can provide decision support for the administrator in order to properly configure or deploy hospital resources.

4. Quantitative Modeling and System Measures for the HED Platform

4.1. Theoretical Analysis

The service-oriented model on the HED platform in Figure 3 is considered to have R servers with an adequate level of staffing and a finite size (N) of waiting rooms for HED arrivals. The birth-and-death

process is adopted to derive analytic steady-state solutions to the M/M/R/N queuing system (QS). Let the states n ($n = 0, 1, 2, \ldots, N$) represent the number of customers in the QS. McManus et al. [18] studied all admissions to the medical–surgical intensive care unit (ICU) of a large, urban children's hospital during a 2-year period. Their statistical analysis confirmed that the arrival rate of patients to ICUs follows a Poisson distribution, and the durations of stay (service times) were found to follow an exponential distribution. Hence, it is reasonably assumed that the customers arrive according to a Poisson process with mean arrival rate $\lambda_n = \lambda$ if $0 \le n \le N$ and $\lambda_n = 0$ if $n > N$ due to a finite system capacity. The QS has R servers, each having an exponential distribution of service times with an identical service rate $\mu_n = \mu$. The service volume can be classified into two parts as follows:

Mean Service Rate:

$$\mu_n = \begin{cases} n\,\mu, \text{ if } 1 \le n \le R \\ R\mu, \text{ if } (R+1) \le n \le N \end{cases} \quad (1)$$

To approach analytic steady-state results for the proposed model, we first construct the state-transition-rate diagram depicted in Figure 4. The number inside the circle represents the number of customers (patients) in the system. Each circle in Figure 4 shows the steady-state probability scenario that may occur during the service period in the system. For each circle except the first one (n = 0) and the last one (n = N), there are four arrows marked with the corresponding values of the state-transition rate. The quantity marked along each arrow gives either the flow-in probability into that state or the flow-out probability off that state.

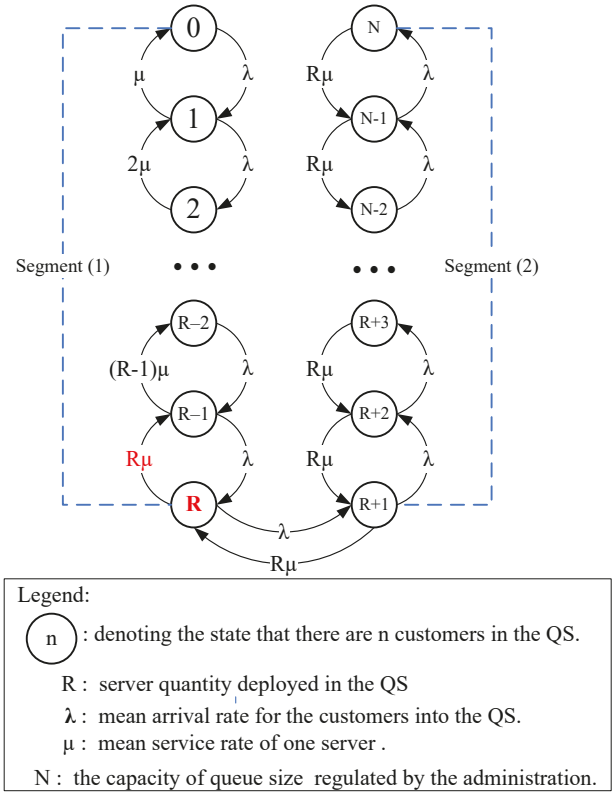


Figure 4. State-transition-rate diagram for the proposed model.

Let the notation P(n) = the probability that there are n customers in the system, where n = 0, 1, 2, . . . , N. Hence, for a steady-state case, the state probability functions P(n) can be obtained from the birth-and-death formula [5] in association with the state-transition-rate diagram shown in Figure 4. We define notation $\rho = \lambda/\mu$ for the server utilization and $\rho_S = \rho/R = \lambda/(R\mu)$ for the system utilization. According to the value n (number of customers in the QS that may be present), two segments are defined by the vector: (Segment 1, Segment 2) = ($1 \le n \le R$, $(R+1) \le n \le N$). The state probability functions P(n) can then be derived in terms of two segments as follows:

(A) Segment (1): $1 \le n \le R$

$$P(n) = \frac{\lambda_0 \cdot \lambda_1 \cdot \lambda_2 \cdots \lambda_{n-1}}{\mu_1 \cdot \mu_2 \cdot \mu_3 \cdots \mu_n} P(0) = \frac{\lambda^n}{\mu\,(2\mu)(3\mu)\cdots(n\mu)} P(0) = \frac{\lambda^n}{\mu^n\, n!} P(0) = \frac{\rho^n}{n!} P(0) \quad (2)$$

(B) Segment (2): $(R+1) \le n \le N$,

$$P(n) = \frac{\lambda_0 \cdot \lambda_1 \cdot \lambda_2 \cdots \lambda_{n-1}}{(\mu_1 \cdot \mu_2 \cdots \mu_R)(\mu_{R+1} \cdots \mu_n)} P(0) = \frac{\lambda^n}{[\mu \cdot (2\mu) \cdots (R\mu)][(R\mu \cdots R\mu)} P(0) = \\ = \frac{\lambda^n}{R!\ \mu^R\ (R\mu)^{n-R}} P(0) = \frac{\rho^n}{R!\ (R)^{n-R}} P(0) \quad (3)$$

There are (n–R) terms of Rμ in the parenthesis (Rμ. Rμ) of the above denominator. Equations (2) and (3) are the closed-forms for the state probability functions P(n) in terms of two segments in which the number of customers may be present. To obtain P(0), we substitute expressions (2) and (3) in the normalizing equation $\sum_{n=0}^{N} P(n) = 1$, which yields:

$$\sum_{n=0}^{R} \frac{\rho^n}{n!} P(0) + \sum_{n=R+1}^{N} \left(\frac{\rho^n}{R!\ R^{n-R}} \right) P(0) = 1$$

$$P(0) = \left[\sum_{n=0}^{R} \frac{\rho^n}{n!} + \sum_{n=R+1}^{N} \left(\frac{\rho^n}{R!\ R^{n-R}} \right) \right]^{-1} = \left[\sum_{n=0}^{R} \frac{\rho^n}{n!} + \frac{\rho^R \left(1 - \rho_s^{N-R+1}\right)}{R!\ (1 - \rho_s)} \right]^{-1} \quad (4)$$

4.2. System Performance Measures

Mathematical expectations are crucially important for the long-run theoretical average values of relevant parameters in the system. To formulate the expressions of the system performance metrics, it is necessary to construct average-based functions, such as the expected number of customers in the queue, expected number of busy servers in the system, etc. The following mathematical analyses are all necessary for the system performance measures of an M/M/R/N QS.

Let

Ls = expected number of customers in the system,
Lq = expected number of customers in the queue buffer,
E[I] = expected number of idle servers,
E[B] = expected number of busy servers,
P_B = Probability that all servers are busy,
Ws = average waiting times in the system,
Wq = average waiting times in the queue buffer.

With steady-state probability functions (2) and (3), it yields

$$L_s = \sum_{n=0}^{N} n\, P(n) \quad (5)$$

$$L_q = \sum_{n=R}^{N} (n - R)\, P(n) \quad (6)$$

$$E[I] = \sum_{n=0}^{R-1} (R - n)\, P(n) \tag{7}$$

$$E[B] = R - E[I] \tag{8}$$

$$P_B = \sum_{n=R}^{N} P(n) \tag{9}$$

To express the above parameters in terms of (R, N, ρ, ρ_S, P_0), the system performance measures can be derived as follows:

$$\begin{gathered} L_s = \sum_{n=0}^{N} n\, P(n) = \sum_{n=0}^{R-1} n\, P(n) + \sum_{n=R}^{N} n\, P(n) = \sum_{n=0}^{R-1} n \cdot \frac{\rho^n}{n!} P(0) + \sum_{n=R}^{N} (n - R + R) P(n) = \\ = \sum_{n=0}^{R-1} n \cdot \frac{\rho^n}{n!} P(0) + \sum_{n=R}^{N} (n - R) P(n) + R \sum_{n=R}^{N} P(n) = \sum_{n=0}^{R-1} n \cdot \frac{\rho^n}{n!} P(0) + L_q + R\, P_B \end{gathered} \tag{10}$$

$$P_B = \sum_{n=R}^{N} P(n) = \sum_{n=R}^{N} \frac{\rho^n}{R!\, R^{n-R}} P(0) = \frac{\rho^R}{R!} \frac{[1 - (\rho_s)^{N-R+1}]}{(1 - \rho_s)} P(0) \tag{11}$$

By changing the indices of **j** = **n**—R so that n = R is changed to j = 0, and n = N is changed to j = N—R,

$$L_q = \sum_{n=R}^{N} (n - R)\, P(n) = \sum_{n=R}^{N} (n - R) \frac{\rho^n}{R!\, R^{n-R}} P(0) = \frac{\rho^R P(0)}{R!} \sum_{j=0}^{N-R} [j \cdot (\rho_s)^j] \tag{12}$$

The average waiting times in the system and in the queue buffer (Ws, Wq) can be derived by applying Little's formula, which gives $W_s = \frac{L_s}{\lambda}$ and $W_q = \frac{L_q}{\lambda}$, respectively.

4.3. An Illustrative Example with Computation Details

To gain prompt perception on the theoretical implication of the quantitative modeling, an example is given by a detailed calculation. Let (R, N) = (4, 5) and (λ, μ) = (2, 1), then the server utilization ρ = λ/μ= 2 and the system utilization ρ_S = ρ/R= 0.5, which is less than unity for the stable system.

(1) $0 \le n \le (R-1) = 3$, $P(n) = \frac{\rho^n}{n!} P(0) = \frac{2^n}{n!} P(0)$

$$P(1) = \frac{2^1}{1!} P(0) = 2P(0);\ P(2) = \frac{2^2}{2!} P(0) = 2P(0);\ P(3) = \frac{2^3}{3!} P(0) = (1.33) P(0) \tag{13}$$

(2) $R \le n \le N$, i.e., For $4 \le n \le 5$, $P(n) = \frac{\rho^n}{R!\, R^{n-R}}\, P(0)$

$$P(4) = \frac{2^4}{4!\, 4^{4-4}}\, P(0) = 0.667\, P(0);\ P(5) = \frac{2^5}{4!\, 4^{5-4}}\, P(0) = 0.333\, P(0) \tag{14}$$

$$\Rightarrow [P(1), P(2), P(3), P(4), P(5)] = = [2, 2, 1.333, 0.667, 0.333]\, P(0). \tag{15}$$

The complete five state probabilities are assembled and expressed in terms of P(0) as follows: Using the normalization condition, $\sum_{n=0}^{N} P(n) = 1 \Rightarrow \sum_{n=0}^{5} P(n) = (7.33) P(0)$

$$\Rightarrow P(0) = 0.136\ \# \tag{16}$$

And from Equations (5)–(9), the system metrics can be computed sequentially as follows:

$$L_s = \sum_{n=0}^{N} n\, P(n) = P(1) + 2P(2) + 3P(3) + 4P(4) + 5P(5) = (14.333) \cdot (0.136) = 1.949 \tag{17}$$

$$L_q = \sum_{n=R}^{N} (n - R)\, P(n) = \sum_{n=4}^{5} (n - 4) P(n) = P(5) = 0.0453 \tag{18}$$

$$E[I] = \sum_{n=0}^{R-1}(R-n)P(n) = \sum_{n=0}^{3}(4-n)P(n) = 4P(0) + 3P(1) + 2P(2) + P(3) = 2.085 \quad (19)$$

$$E[B] = R - E[I] = 4 - 2.085 = 1.915 \quad (20)$$

$$P_B = \sum_{n=R}^{N} P(n) = \sum_{n=4}^{5} P(n) = P(4) + P(5) = 0.136 \quad (21)$$

$$Ws = Ls/\lambda = 1.949/2 = 0.9745 \text{ and } Wq = Lq/\lambda = 0.0453/2 = 0.0223 \quad (22)$$

The distribution of steady-state probabilities is depicted in Figure 5. The relevant system performance measures, such as Ls, Lq, E[B] and Ws, are shown in the left-lower part of Figure 5. The average number of customers in the QS and the queue buffer are 1.949 and 0.0457, respectively. The average waiting times in the system and the queue buffer are 0.9745 and 0.0223, respectively.

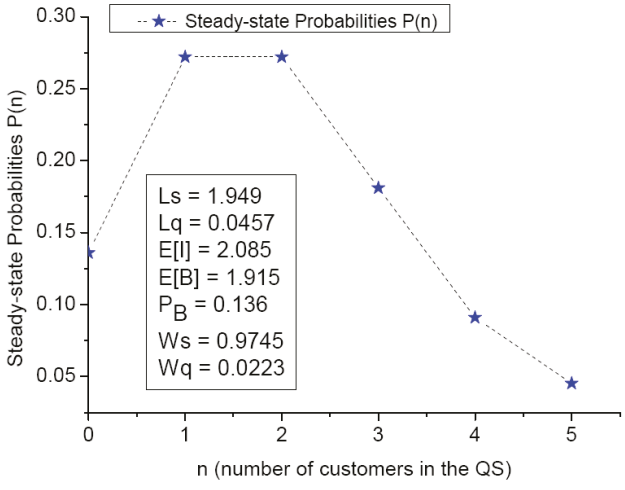


Figure 5. Steady-state probabilities with parameters (R, N, λ, μ) = (4, 5, 2, 1).

5. Issue on Decision Support for HED Management

5.1. Evaluation Formulation on Cost

The strategy to minimize the total cost of the operating horizon is referred to as the optimal policy. Like all medical institutions, the cost pattern is important for gaining long-term and stable hospital management. To optimize the cost, we developed a steady-state expected cost function per unit time for an M/M/R/N queuing system, in which the parameter vector of (R, N, λ, μ) and the average waiting times (Lq) are considered as decision variables. The cost element C_W is defined as the waiting cost per unit time (or cost rate) per customer (HED patient) present in the system. Our goal is to provide decision support for determining the optimal number of servers R, say R*, to optimize the cost function. To formulate the cost function, some cost parameters are defined in the following vector form as follows:

Cq = cost per unit time when one customer is waiting for service,
Cs = cost per unit time when one customer joins the system and is served,
(C_B, C_I) = cost per unit time when one server is (busy, idle).

Using the definitions of each cost element with its corresponding feature, the cost function F(R, N) can be developed in association with the system metrics Ls, P_B, Lq, E[I], and E[B], which are given in Equations (10)–(12), (7) and (8), respectively. It is noted that the steady-state probabilities for two segments are given in Equations (2) and (3). The probability that there is no customer in the system, P(0), is given by Equation (4).

$$\begin{aligned} F(R,\ N) &= Cq\ Lq + Cs\ (Ls - Lq) + C_B\ E[B] + C_I\ E[I] = (Cq - Cs)\ Lq + Cs\ Ls + C_B\ E[B] + C_I\ E[I] \\ &= (Cq - Cs)\ \frac{\rho^R P(0)}{R!} \sum_{j=0}^{N-R} [j\cdot(\rho_s)^j] + Cs\ \{\sum_{n=0}^{R-1} n\cdot\frac{\rho^n}{n!}P(0) + \frac{\rho^R P(0)}{R!} \sum_{j=0}^{N-R} [j\cdot(\rho_s)^j] + R\cdot \\ &\quad \frac{\rho^R}{R!}\frac{[1-(\rho_s)^{N-R+1}]}{(1-\rho_s)}\} + C_B\ (R - \sum_{n=0}^{R-1} (R-n)\ P(n)) + C_I\ [\sum_{n=0}^{R-1} (R-n)\ P(n)] \end{aligned} \quad (23)$$

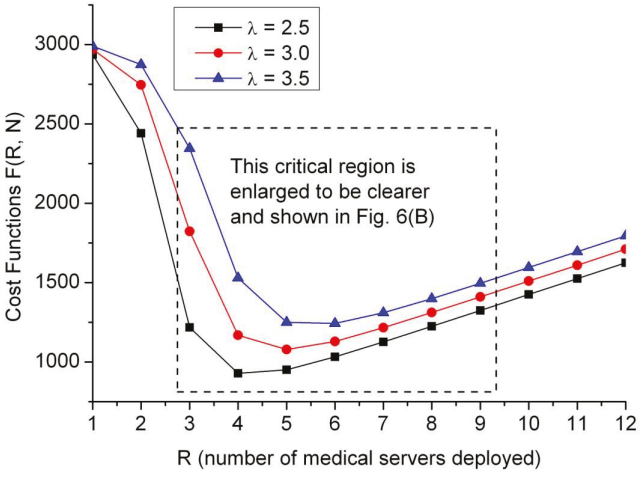
The cost function F(R, N) in Equation (23) is expressed in terms of basic parameters, such as (R, N, λ, μ), and cost elements. It is noted that the utilization parameters of the unit server and system is given by (ρ, ρ_S) =(λ/μ, λ/(Rμ)), respectively. The state probability functions P(n) for two segments are given in Equations (2) and (3), which are quite complex for the control parameter R. To find the optimal profile on the cost function, it is necessary to show the existence of convexity or unimodality of F(R, N). However, this mathematical task is difficult to implement. The cost function F(R, N) is unimodal; that is, it has a single relative minimum.

5.2. Evaluation of Cost Optimization

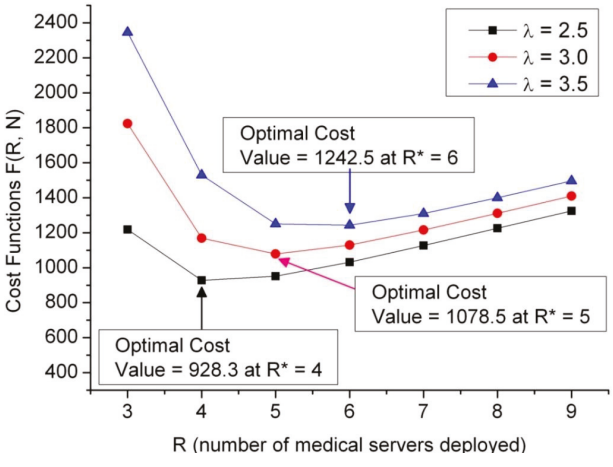
Equation (23) shows that the parameter R occurs not only at the location of in-line items, but also at the upper limit of the summation symbol Σ, which makes F(R, N) a highly nonlinear and complex function. Instead, practical numerical examples are presented and intensively studied by applying the proposed models. The optimization evaluation is firstly probed in terms of cost patterns in this subsection. For illustrative purposes, we first study the effect of varying R while keeping N constant, and then varying N while keeping R constant. All simulations are performed with the MATLAB platform with custom MATLAB scripts. The exemplified system parameters are listed as vector forms as follows:

(a) Average arrival rate of patients (λ) = 2.5, 3.0, and 3.5,
(b) Average service rate of a server (μ) = 1,
(c) Cost rate: (C_q, C_s, C_B, C_I,) = (200, 150, 120, 100),
(d) N = 15 for emergency departments of small and medium size.

Contour plots may provide the best graphical representation of the optimization problem, and also possess a powerful visualization that permits the solutions of the optimization problem by inspection. To validate the analytical solution, the graphical results were obtained and are shown in Figure 6A, where three cost contours with the black box, red circle, and blue triangle icons are depicted along the Y-axis in terms of λ = 2.5, 3.0, and 3.5, respectively. Generally, a higher patient arrival rate implies that the medical service cost is higher, so the blue line marked with the triangle icon (λ = 3.5) is situated over the red line marked with a circle (λ = 3.0). To clearly show the crucial region surrounded by the dashed-line rectangle in Figure 6A, enlarged detail is depicted in Figure 6B. In Figure 6B, the critical region is between R = 3 and R = 9 on the X-axis. The optimal cost value with the corresponding optimal R* is shown for each contour.



(**A**)



(**B**)

Figure 6. (**A**). Optimal cost patterns shown in terms of three average arrival rates. (**B**) An enlarged diagram showing the optimal cost data from Figure 6A.

5.3. Issues on Cost Profile under the Constraint of Average Waiting Time

In view of performance evaluation, the average waiting time (AWT) can also be regarded as a measure of performance committed to the HED patients, and of a yardstick for comparing the effectiveness of the deployment of the staffing providers in a quantitative manner. Practically, it is reasonable for management experts to guarantee an AWT target level when they want to alleviate the sense of worry for potential HED patients. Logically, the higher the number of staffing providers deployed, the higher the cost. Hence, the proposed approach explores the issue of decision support for optimal cost under the constraint of the AWT at some target level.

In Figure 7 with the double-Y axis, the left Y-axis and the right Y-axis are set to be the cost values and the average waiting time (AWT), respectively. Observing the solid-line contour marked with black rectangles (i.e., the left Y-axis), the optimal cost value F(R, N) = 1242.5, which occurs at R* = 6 based on the similar parameters in Figure 6A with the average arrival rate λ = 3.5. However, the corresponding

AWT approaches 6.84 units, which is a reference metric for decision-making. The proposed generic model could be used for general insights into the issues faced in deploying multiple staffing providers for a disease chain or a single department like the Department of Pediatrics shown in the middle right-handed location of the ground floor in Figure 1. On the right Y-axis, the dash-curve marked with a red star shows the variation profile of the performance metric for AWT.

During the busy time-window for a specific disease chain in HED, patients may spend hours in crowded waiting rooms before seeing a doctor. Those who choose to tolerate longer waiting times expose themselves to others who may have a contagious illness. To alleviate such an occasional impact, one straight approach for reducing the waiting time is to deploy more staffing providers for that specific disease chain. The simulation results in Figure 7 provide an exemplified decision reference on re-deploying the amount of staffing providers to alleviate the waiting time.

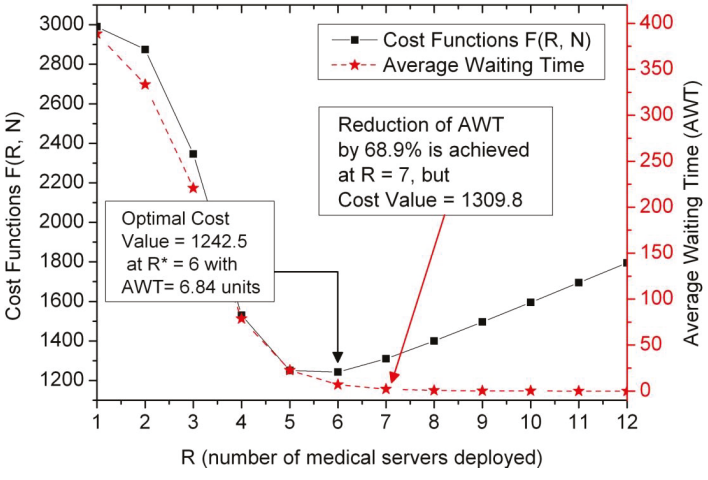


Figure 7. Decision support on optimal cost at R* = 7 under the constraint of reduction of AWT (average waiting time) by 68.9%, which is calculated from ((6.84–2.13)/6.84) × 100%.

Then an issue emerges from the judgment: how many extra staffing providers are needed to gain a reduction in the AWT by some level (for example, 50%) without over-provisioning? Observing the red-star contour with the right Y-axis in Figure 7, it is found that the AWT can be reduced by 68.9% at R* = 7 (shown by the red dash-line) at the expense of only adding one staffing provider and cost values F(R* = 7, N) = 1309.8, compared to the minimum cost F(R = 6, N) = 1242.5. The detailed numerical data are listed in Table 1 with a range of R from unity to 12. The value of AWT for R = 12 is less than 0.01 and then marked to be 0 for clarity. In other words, the proposed approach can provide a quantitative decision support on the trade-off study between the cost profile and the amount of staffing providers in HED deployment.

Table 1. Numerical data on AWT and the corresponding cost values for the range of R from unity to 12.

R	Cost Values	AWT	R	Cost Values	AWT
1	2990.0	388.57	7	1309.9	2.13
2	2873.9	333.42	8	1399.5	0.65
3	2345.8	220.77	9	1496.3	0.19
4	1530.2	78.57	10	1595.4	0.05
5	1250.7	22.51	11	1695.1	0.01
6	1242.5	6.84	12	1795.0	0

AWT, average waiting time.

5.4. Application Profile in a Window-by-Window Way

This work addressed the issue of the mathematical modelling to evaluate scenarios for deployment of medical resources to the HED, and also aimed to provide feasible applications iteratively to approaching an effective decision support in terms of deploying appropriate staffing providers to alleviate the impact on HED crowding. Patients who want to receive medical services always arrive randomly, and they require immediate services available at that time. If the service facility is operating at peak capacity when they arrive, they are obliged to wait in line (queue) with patience in the case of a shortage in staffing level. The surges and changes in HED activity may occur from time to time in terms of various time-widows with associated system parameters, as shown in Figure 8.

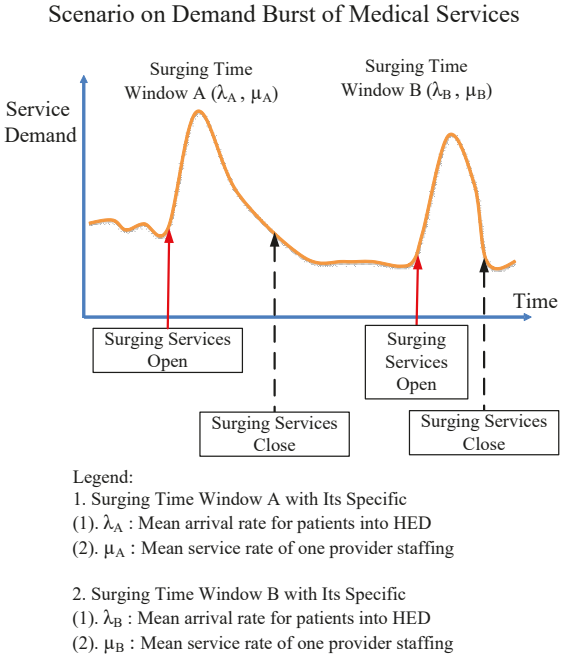


Figure 8. Iterative applications exemplified by various time-windows.

In Figure 8, each unique time-window may represent a specific surge time-period, wherein larger numbers of patients nearby the hospital are delivered to HED after some disaster or traffic event has occurred. In modeling language, the proposed modeling approaches can be applied in a window-by-window way that each specific time window can be approximated by its λ (average arrival rate) and μ (average service rate) in association with various practical historical data. For example, supposing that the surge in time window B of Figure 8 represents some middle-level traffic event, then the system parameters λ_B (average arrival rate) and μ_B (average service rate) may be approximated by some existing past and experienced parameters for the baseline, and then the cost function F(R, N) (23) may be applied iteratively to approach the cost optimization in a window-by-window way.

6. Conclusions

In terms of patient flow and available resources, an efficient generic methodology to optimize the performance of the HED platform has been addressed in this research. The proposed queue-based approach provides HED administrators with an efficient deployment of staffing providers to optimize the cost profile. Conceptually, the HED service platform was mapped into an M/M/R/N queuing

system, and illustrated using appropriate figures and materials in the work. To gain insight into the queuing model, the mathematical derivation was detailed for the application need as well.

Based on the quantitative analysis, the M/M/R/N queue model was applied and derived, and then the relevant system metrics were established in a brand-new manner. The mathematical expression for cost function was established for evaluation requirements. In regards to verification, the relevant experimental results were obtained in terms of integration configurations on cost optimization and average waiting time. Instead of chaotic management, the proposed generic methodology may provide feasible applications for approaching an effective decision support in terms of deploying appropriate staffing providers to alleviate the impact on HED crowding.

Author Contributions: Conceptualization, F.-C.J.; Data curation, Y.-J.C. and C.-H.L.; Formal analysis, C.-M.S.; Investigation, C.-H.L.; Methodology, C.-M.S. and Y.-M.W.; Project administration, Y.-J.C.; Resources, C.-T.Y.; Software, C.-T.Y.; Supervision, F.-C.J.; Validation, C.-M.S.; Writing—original draft, F.-C.J. and Y.-M.W.; Writing—review & editing, F.-C.J. and C.-T.Y.

Funding: This research was supported in part by the Ministry of Science and Technology research grants MOST 108-2119-M-029-001-A and 108-2221-E-029-010. This research was also supported in part from Taichung Veterans General Hospital (TCVGH-1075103B to Cheng-Min Shih).

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

References

1. Zengin, S.; Guzel, R.; Al, B.; Kartal, S.; Sarcan, E.; Yildirim, C. Cost analysis of a university hospital's adult emergency service. *J. Acad. Emerg. Med.* **2013**, *12*, 71–75. [CrossRef]
2. Simonet, D. Cost reduction strategies for emergency services: Insurance role, practice changes and patients accountability. *Health Care Anal.* **2009**, *17*, 1–19. [CrossRef] [PubMed]
3. Cremonesi, P.; Di Bella, E.; Montefiori, M. Cost analysis of emergency department. *J. Prev. Med. Hyg.* **2010**, *51*, 157–163. [PubMed]
4. Cooper, R.B. *Introduction to Queuing Theory*, 2nd ed.; Elsevier Science Publishing Co., Inc.: New York, NY, USA, 1990.
5. Gross, D.; Shortle, J.F.; Thompson, J.M.; Harris, C. *Fundamentals of Queuing Theory*, 4th ed.; John Wiley &Sons, Inc.: New York, NY, USA, 2008.
6. Finamore, S.R.; Sheila, S.A. Shorting the wait: A strategy to reduce waiting times in the emergency department. *J. Emerg. Nurs.* **2009**, *35*, 509–514. [CrossRef] [PubMed]
7. Green, L.V.; Soares, J.; Giglio, J.F.; Green, R.A. Using queuing theory to increase the effectiveness of emergency department provider staffing. *Acad. Emerg. Med.* **2006**, *13*, 61–68. [CrossRef] [PubMed]
8. Derlet, R.W.; Richards, J.R.; Kravitz, R.L. Frequent overcrowding in U.S. emergency departments. *Acad. Emerg. Med.* **2001**, *8*, 151–155. [CrossRef] [PubMed]
9. Kennedy, J.; Rhodes, K.; Walls, C.A.; Asplin, B.R. Access to emergency care: Restricted by long waiting times and cost and coverage concerns. *Ann. Emerg. Med.* **2004**, *43*, 567–573. [CrossRef] [PubMed]
10. Hoot, N.R.; Aronsky, D. Systematic review of emergency department crowding: Causes, effects, and solutions. *Ann. Emerg. Med.* **2008**, *52*, 126–136. [CrossRef] [PubMed]
11. Ding, R.; Zeger, S.L.; McCarthy, M.L.; Pines, J.M. Comparison of methods for measuring crowding and its effects on length of stay in the emergency department. *Acad. Emerg. Med.* **2011**, *18*, 1269–1277.
12. Kaushal, A.; Zhao, Y.; Peng, Q.; Strome, T.; Weldon, E.; Zhang, M.; Chochinov, A. Evaluation of fast track strategies using agent-based simulation modeling to reduce waiting time in a hospital emergency department. *Socio Econ. Plan. Sci.* **2015**, *50*, 18–31. [CrossRef]
13. Vass, H.; Szabo, Z.K. Application of queuing model to patient flow in emergency department. Case study. *Procedia Econ. Financ.* **2015**, *32*, 479–487. [CrossRef]
14. Taichung Veterans General Hospital (TVGH): Departments of Medical Services. Available online: http://www.vghtc.org.tw/ (accessed on 5 December 2019).
15. Fletcher, A.; Worthington, D. What is a 'generic' hospital model?—A comparison of 'generic' and 'specific' hospital models of emergency patient flows. *Health Care Manag. Sci.* **2009**, *12*, 374–391. [CrossRef]
16. Hoot, N.R.; Leblanc, L.J.; Jones, I.; Levin, S.R.; Zhou, C.; Gadd, C.S.; Aronsky, D. Forecasting emergency department crowding: A discrete event simulation. *Ann. Emerg. Med.* **2008**, *52*, 116–125. [CrossRef]

17. Cochran, J.K.; Roche, K.T. A multi-class queuing network analysis methodology for improving hospital emergency department performance. *Comput. Oper. Res.* **2009**, *36*, 1497–1512. [CrossRef]
18. McManus, M.L.; Long, M.C.; Cooper, A.; Litvak, E. Queuing theory accurately models the need for critical care resources. *Anesthesiology* **2004**, *100*, 1271–1276. [CrossRef]
19. Marmor, Y.N.; Wasserkrug, S.; Zeltyn, S.; Mesika, Y.; Greenshpan, O.; Carmeli, B.; Shtub, A.; Mandelbaum, A. Toward simulation-based real-time decision-support systems for emergency departments. In Proceedings of the 2009 Winter Simulation Conference (WSC), Austin, TX, USA, 13–16 December 2009.
20. Oh, C.; Novotny, A.M.; Carter, P.L.; Ready, R.K.; Campbell, D.D.; Leckie, M.C. Use of a simulation-based decision support tool to improve emergency department throughput. *Oper. Res. Health Care* **2016**, *9*, 29–39. [CrossRef]

Journal of
Clinical Medicine

MDPI

Article

Predicting Long-Term Health-Related Quality of Life after Bariatric Surgery Using a Conventional Neural Network: A Study Based on the Scandinavian Obesity Surgery Registry

Yang Cao [1,*], Mustafa Raoof [2], Scott Montgomery [1,3,4], Johan Ottosson [2] and Ingmar Näslund [2]

1 Clinical Epidemiology and Biostatistics, School of Medical Sciences, Örebro University, 70182 Örebro, Sweden; scott.montgomery@oru.se
2 Department of Surgery, Faculty of Medicine and Health, Örebro University, 70182 Örebro, Sweden; mustafa.raoof@regionorebrolan.se (M.R.); johan.ottosson@regionorebrolan.se (J.O.); ingmar.naslund@regionorebrolan.se (I.N.)
3 Clinical Epidemiology Division, Department of Medicine, Karolinska Institutet, 17177 Stockholm, Sweden
4 Department of Epidemiology and Public Health, University College London, London WC1E 6BT, UK
* Correspondence: yang.cao@oru.se; Tel.: +46-19-602-6236

Received: 4 November 2019; Accepted: 3 December 2019; Published: 5 December 2019

Abstract: Severe obesity has been associated with numerous comorbidities and reduced health-related quality of life (HRQoL). Although many studies have reported changes in HRQoL after bariatric surgery, few were long-term prospective studies. We examined the performance of the convolution neural network (CNN) for predicting 5-year HRQoL after bariatric surgery based on the available preoperative information from the Scandinavian Obesity Surgery Registry (SOReg). CNN was used to predict the 5-year HRQoL after bariatric surgery in a training dataset and evaluated in a test dataset. In general, performance of the CNN model (measured as mean squared error, MSE) increased with more convolution layer filters, computation units, and epochs, and decreased with a larger batch size. The CNN model showed an overwhelming advantage in predicting all the HRQoL measures. The MSEs of the CNN model for training data were 8% to 80% smaller than those of the linear regression model. When the models were evaluated using the test data, the CNN model performed better than the linear regression model. However, the issue of overfitting was apparent in the CNN model. We concluded that the performance of the CNN is better than the traditional multivariate linear regression model in predicting long-term HRQoL after bariatric surgery; however, the overfitting issue needs to be mitigated using more features or more patients to train the model.

Keywords: prediction; deep learning; conventional neural network; health-related quality of life; bariatric surgery

1. Introduction

Severe obesity, defined as having a body mass index (BMI) greater than 40 kg/m^2 or greater than 35 kg/m^2 plus at least one obesity-related comorbidity [1,2], has been associated with numerous health outcomes and reduced health-related quality of life (HRQoL) [3–8]. HRQoL measures population health multi-dimensionally from physical, mental, emotional, and social functioning domains, which have already been identified as an important indication for bariatric surgery and recognized by the United States National Institutes of Health Conference as early as 1991 [9,10]. Although many studies have reported changes in HROoL after bariatric surgery, few are long-term prospective studies. A systematic review of seven prospective cohort studies with a follow-up time of ≥5 years revealed that bariatric surgery patients reported considerably improved HRQoL and the improvement was maintained over

the long term [11]. However, many patients still experience reduced HRQoL after surgery. In our study, 39% of patients had significant improvements in physical functioning (PF) (increased by >25 in the original score or >0.25 in the scaled score), and the rest had no significant improvement and some patients (2%) even had significant deterioration (reduced by >25 in the original score or >0.25 in the scaled score). No relationship between the PF scores before and 5 years after surgery was identified (Figure S1).

Although some preoperative psychological factors, including personality change, severe psychiatric disorder, or depressive symptoms, are associated with postoperative HRQoL after bariatric surgery [12,13], whether long-term HRQoL after bariatric surgery can be predicted based on patients' baseline features has not been investigated. The present study examined the performance of the convolution neural network (CNN) for predicting 5-year HRQoL after bariatric surgery based on the available preoperative information from a national quality registry, and compared CNN with a conventional linear regression estimator.

2. Material and Methods

2.1. Patients and Features

Data for the patients registered in the Scandinavian Obesity Surgery Registry (SOReg) were used for the current study. The SOReg was launched in 2007 and covers 98% of bariatric surgery in Sweden since 2009. SOReg is validated regularly and has been shown to have high data quality [14]. In total, 27 - of 42 operating centers in Sweden participate in the HRQoL registration in SOReg. HRQoL was measured using the RAND-SF-36 and the obesity-related problems (OP) scale preoperatively and 1, 2, and 5 years after surgery. In the present study, preoperative and 5-year HRQoL data, including PF, role physical (RP), bodily pain (BP), general health (GH), vitality (VT), social functioning (SF), role emotional (RE), mental health (MH) scale, summary physical scale (PCS), summary mental scale (MCS), and OP, were used. All scale scores ranged from 0 to 100, with higher scores indicating better health status except for OP, where low values represent good health. Eight baseline features, including sex, age, BMI, sleep apnea syndrome (SAS), hypertension, diabetes, dyslipidemia, and depression, were also used as predictors.

In total, 6687 patients with complete information on 19 baseline features and 11 5-year HRQoL measures were used in the machine learning study.

The data that support the study are not publicly available because they contain information that could compromise research participant privacy and confidentiality. The authors will make the data available upon reasonable request and with permission of the Committee of Scandinavian Obesity Surgery Registry in Örebro, Sweden.

2.2. Feature Scaling

Before machine learning, the features in the dataset were scaled. The binary features were converted into dummy variables, and the continuous features were scaled to between 0 and 1 using a min-max scaler. In the sensitivity analysis, the normalizer and standardizer scalers were also used to evaluate the influence of scalers on the model's performance.

2.3. Conventional Neural Network

A CNN is a regularized version of a multi-layer perceptron neural network, which was inspired by a biological process where the connectivity pattern between neurons resembles the organization of the visual cortex [15]. Although not specifically developed for non-image data, CNN may achieve state-of-the-art results on regression prediction problems, especially for data with time series or spatial patterns. The CNN input is traditionally two-dimensional (2D) but can also be changed to be one-dimensional (1D), allowing it to develop an internal representation of a 1D sequence. In our study, we used a CNN with seven hidden layers, including two 1D convolution layers (with 10 filters for

each), two 1D max pooling layers, one flattened layer, and two dense layers (with 1000 computation units). The rectified linear unit (relu) activation function was used for the convolution layers and dense layers, and the normal distribution was used to initialize weights in the layers. The mean squared error was used as the loss function and the an Adadelta algorithm was used as optimizer when compiling the model [16]. The structure of the CNN model is shown in Figure S2.

2.4. Model Validation and Evaluation

In total, 20% of the patients were randomly selected as a test dataset for the final evaluation of the data, and the rest of the patients were used as the training dataset. To find optimal high-level parameters (like the number, size, and type of layers in the networks) and lower-level parameters (like the number of epochs, choice of loss function and activation function, and optimization procedure) in the CNN model, the K-fold cross validation method was used during the training phase [17]. We split the training data into 5 partitions, instantiated 5 identical models, and trained each one on 4 partitions while validating on the remaining partition. The performance of each model was evaluated using the mean squared error (MSE) because of the existence of zero values in the outcome variables. We then computed the average performance over the 5 folds. In the end, the choice of the parameters was a compromise between the model's performance and computing time, i.e., the model with both the smallest validation error and a shorter computing time was deemed an optimal model. The training, validation, and final evaluation process is shown in Figure S3.

To avoid overuse of the deep learning method for prediction, we also applied a simple multivariate linear regression model as an estimator to predict the 5-year HRQoL scores, and compared the performance between the linear regression model and the CNN model.

2.5. Software and Hardware

The descriptive and inferential statistical analyses were performed using Stata 15.1 (StataCorp LLC, College Station, TX, USA). The CNN and multiple linear regression models were achieved using packages scikit-learn 0.21.2 and Keras 2.2.4 in Python 3.6 (Python Software Foundation, https://www.python.org/).

All of the computation was conducted in a computer with a 64-bit Windows 7 Enterprise operation system (Service Pack 1), Intel ®Core TM i5-4210U CPU @ 2.40 GHz, and 16.0 GB random access memory.

3. Results

3.1. Descriptive Analysis of the Data

In total, 6687 patients registered in SOReg between 2008 and 2012 with complete demographic and preoperative comorbidity information, and preoperative and 5-year HROoL scores were included in the study. The characteristics of the patients are shown in Table 1. Briefly, the average age and BMI of the patients were 42.7 years and 42.3 kg/m2, respectively. More than three quarters (77%) were female and 45% had at least one of the five comorbidities (SAS, hypertension, diabetes, depression, and dyslipidemia) before bariatric surgery.

Table 1. Characteristics of the patients (n = 6687) included in the study, mean (SD) or n (%).

	Preoperative		Five Years after Bariatric Surgery	
	Original	**Scaled**	**Original**	**Scaled**
Age (year)	42.7 (11.0)	0.494 (0.197)	47.7 (11.0)	0.494 (0.197)
BMI (kg/m^2)	42.3 (5.2)	0.241 (0.103)	30.3 (5.2)	0.358 (0.127)
Female	5259 (77%)	NA	5259 (77%)	NA
SAS	680 (10%)	NA	NA	NA
Hypertension	1851 (27%)	NA	NA	NA
Diabetes	990 (15%)	NA	NA	NA
Depression	884 (13%)	NA	NA	NA
Dyslipidemia	747 (11%)	NA	NA	NA
PF	61.6 (21.9)	0.616 (0.219)	84.2 (20.7)	0.842 (0.207)
RP	60.2 (38.9)	0.602 (0.389)	77.8 (36.6)	0.778 (0.366)
BP	56.0 (26.8)	0.560 (0.268)	65.1 (30.8)	0.651 (0.308)
GH	58.2 (21.4)	0.582 (0.214)	68.0 (24.7)	0.680 (0.247)
VT	47.3 (23.0)	0.473 (0.230)	54.5 (26.9)	0.545 (0.269)
SF	74.8 (26.1)	0.748 (0.261)	79.5(26.5)	0.795 (0.265)
RE	75.9 (36.2)	0.759 (0.362)	76.7 (37.9)	0.767 (0.379)
MH	71.5 (19.4)	0.715 (0.194)	72.0 (23.0)	0.720 (0.230)
PCS	38.3 (10.7)	0.567 (0.177)	47.6 (11.1)	0.653 (0.163)
MCS	46.8 (11.7)	0.621 (0.172)	44.6 (13.8)	0.621 (0.192)
OP	61.0 (26.3)	0.610 (0.263)	25.6 (27.4)	0.256 (0.274)

SD, standard deviation; NA, not applicable; BMI, body mass index; SAS, sleep apnea syndrome; PF, physical functioning; RP, role-physical; BP, bodily pain; GH, general health; VT, vitality; SF, social functioning; RE, role-emotional; MH, mental health; PCS, summary physical scale; MCS, summary mental scale; OP, obesity-related problems.

3.2. *Performance of the CNN Model in the K-Fold Cross-Validation*

We analyzed 11 HRQoL scores in the study. To make our description concise, we used the PF score as an example of our data analysis as follows.

In general, the performance of the CNN model (measured as the MSE) increased with more convolution layer filters, computation units, and epochs, and decreased with a larger batch size. Although the performance increased with the model's complexity, the computing time increased exponentially. When we set the number of computation units and filters to be large enough (1000 and 10, respectively) and the batch size was small enough (10), the performance of the CNN model in K-fold cross-validation is shown in Figure 1. The performance was not stable when the number of epochs was small and changed dramatically depending on the random seed used in training (Figure 1). When the number of epochs was >40, the model presented smaller MSE than the linear regression model (0.032 vs. 0.035, Figures 1 and 2). Although more epochs reduced the MSE in the CNN model, the computing time increased exponentially, indicating a higher cost in machine learning (Figure 1). The MSE of the linear regression model appeared constant when the number of epochs >40 (Figure 2), which means the prediction cannot be improved with more epochs. The cross-validation indicates that the CNN model may provide better prediction but at the expense of the computing time.

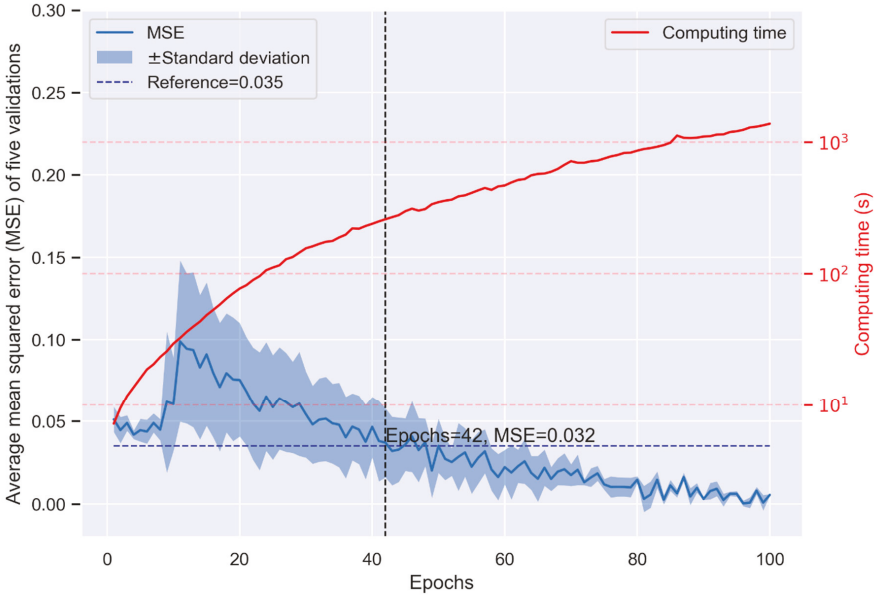


Figure 1. Performance of the convolution neural network (CNN) model in K-fold cross-validation.

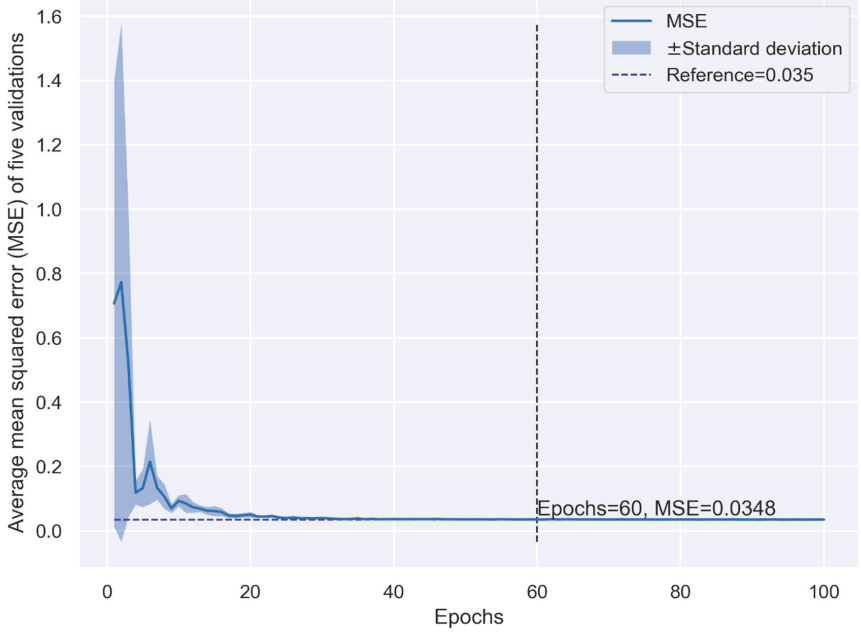


Figure 2. Performance of the simple multivariate linear regression model in K-fold cross-validation.

3.3. Performance of the CNN Model in the Final Evaluation

When the models were evaluated using the test data that were not seen previously by the models, in general, the CNN model presented a better performance (solid line in Figure 3b) than

the linear regression model (solid line in Figure 3a) with epochs >40. Although overfitting was presented sporadically in the CNN model (comparing the solid line with the dotted line in Figure 3b), the performance improved gradually with an increased number of epochs while remaining constant in the linear regression model.

Finally, we used 40 epochs for the CNN model, and predicted PF scores for both the training data and the test data. Clear correlations can be seen between the predicted values and observed values in the training data, with an MSE of 0.032 for the CNN model (Figure 3d and Table 2) compared to the MSE of 0.033 seen in the linear regression model (Figure 3c and Table 2). For the test data, the CNN model had an MSE of 0.035 (Figure 3f and Table 2) compared with 0.034 (Figure 3e and Table 2) from the linear regression model. Although the CNN model provided better prediction than the linear regression model for the test data, the overfitting became apparent in some situations when the model learned patterns more specific to the training data.

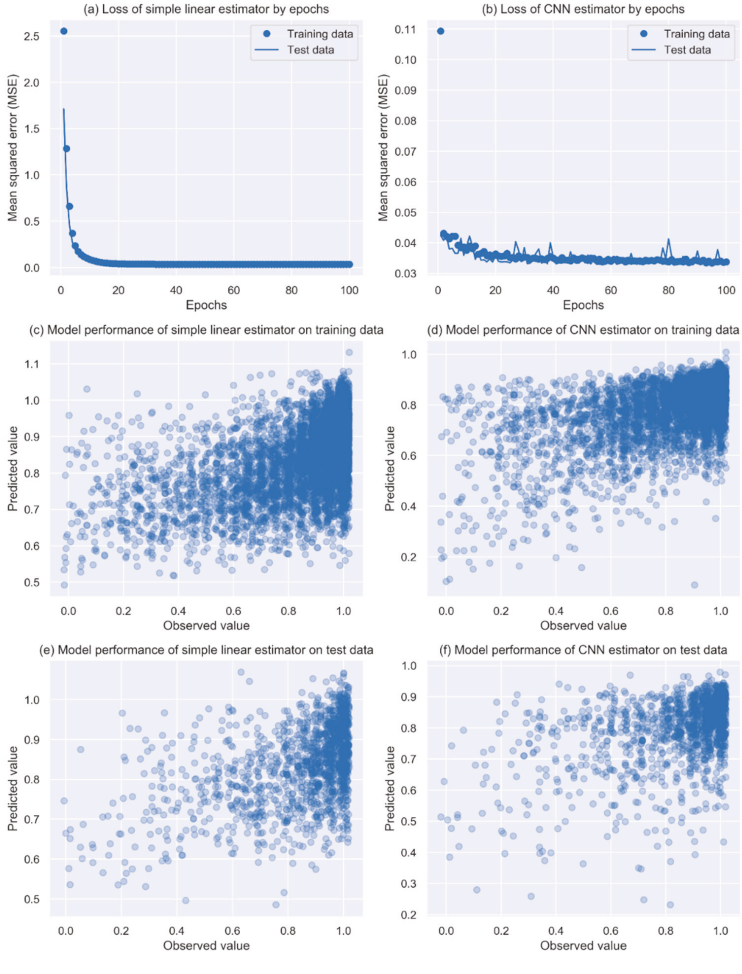


Figure 3. Model performance of the simple linear estimator and the CNN estimator. The dots in the plots (c)–(f) were jittered to avoid a heavy overlap of patients with the same coordinates. CNN, convolution neural network.

3.4. Performance of CNN in Predicting Other HRQoL Measures

The relationships between the baseline and the 5-year scores of other HRQoL measures in the test data are shown in Figure 4. Except for GH and VT, no clear relationship between the baseline and the observed 5-year scores is seen for the HRQoL measures (Figure 4, plots a1–j1). However, the predicted 5-year scores based on the baseline scores and the CNN model show clear correlations with the observed 5-year scores for BP, GH, VT, MH, MCS, and OP (Figure 4, plots a2–j2).

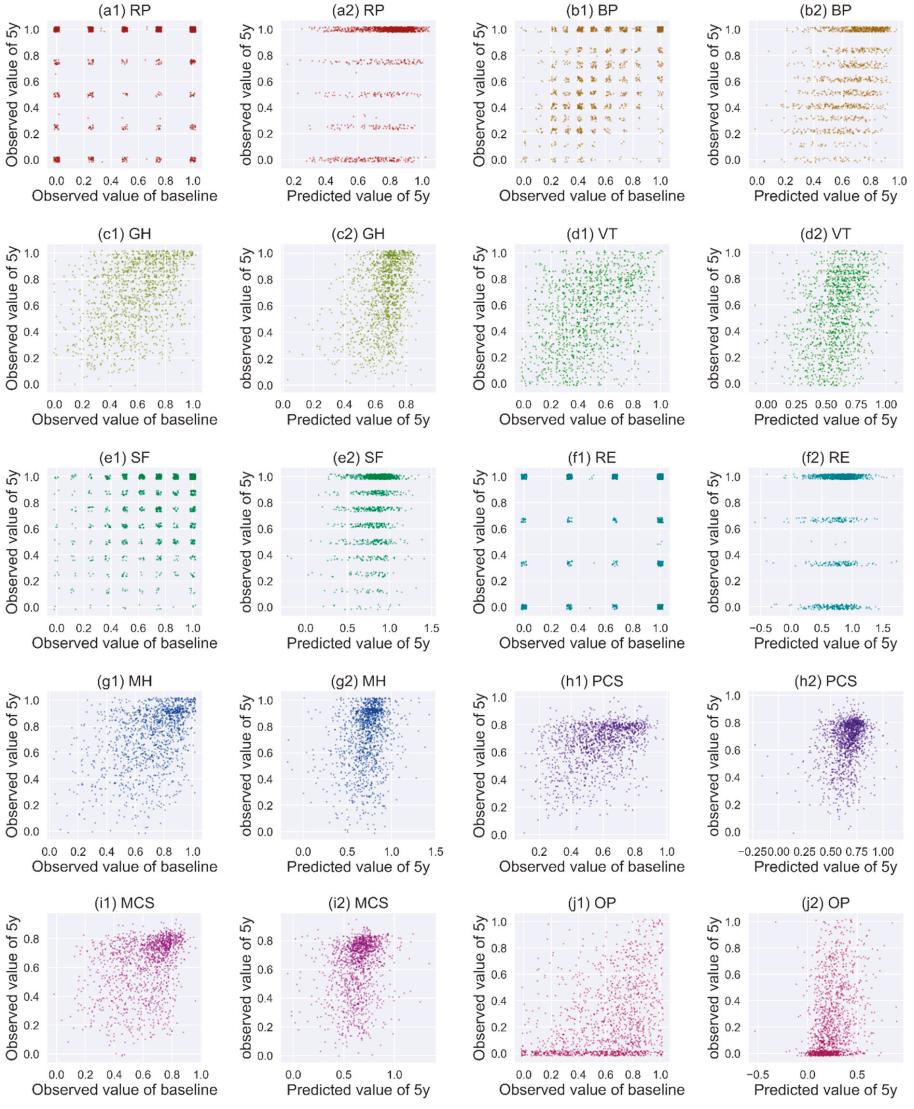


Figure 4. Correlation of the observed 5-year scores with the observed baseline scores and predicted scores for test data. The dots in the plots were jittered to avoid a heavy overlap of patients with the same coordinates. RP, role physical; BP, bodily pain; GH, general health; VT, vitality; SF, social functioning; RE, role emotional; MH, mental health; PCS, summary physical scale; MCS, summary mental scale; OP, obesity-related problems.

We compared the performance of the CNN model and the linear regression model for all the HRQoL measures in both the training data and the test data. The CNN model showed an overwhelming advantage in predicting all the HRQoL measures. The MSEs of the CNN model for the training data were 8% to 80% smaller than those of the linear regression model (Table 2). The overfitting was also apparent in the CNN model, i.e., the MSEs of the CNN model for the test data were all greater than those of the linear regression model (Table 2).

Table 2. Mean squared errors (MSEs) of the CNN model and the multivariate linear regression model.

HRQoL Measure	Training Data		Test Data	
	CNN Model	Linear Regression Model	CNN Model	Linear Regression Model
PF	0.0316	0.0329	0.0350	0.0343
RP	0.1078	0.1178	0.1324	0.1211
BP	0.0604	0.0763	0.0898	0.0772
GH	0.0280	0.0497	0.0618	0.0508
VT	0.0303	0.0572	0.0914	0.0625
SF	0.0213	0.0600	0.0995	0.0588
RE	0.0393	0.1275	0.2118	0.1269
MH	0.0119	0.0427	0.0807	0.0416
PCS	0.0087	0.0210	0.0333	0.0219
MCS	0.0075	0.0301	0.0584	0.0305
OP	0.0450	0.0625	0.0750	0.0608

PF, physical functioning; RP, role physical; BP, bodily pain; GH, general health; VT, vitality; SF, social functioning; RE, role emotional; MH, mental health; PCS, summary physical scale; MCS, summary mental scale; OP, obesity-related problems.

3.5. Sensitivity Analysis and Computing Time

We also conducted sensitivity analysis using different scalers and optimizers in data preparation and model compiling, and tuned the hyperparameters using the exhaustive grid search method [18]. Although they showed more or less influence on the models' performance, the influence was negligible when the number of epochs was large and the batch size was small. The computing time for the CNN model largely depends on the hyperparameter settings of the layers, number of epochs and the batch size for training, and the software and hardware used. In our study, with the model structures and hyperparameters shown in Figure S2, the running time ranged from 70 (epoch = 40, batch size = 10, without cross-validation) to 595 s (epoch = 400, batch size = 10, with five cross-validations) on our computer.

4. Discussion

Machine learning methods to predict HRQoL have been used in elderly with chronic diseases [19], cervical cancer patients [20], and osteoarthritis patients [21]. However, to our knowledge, they have not been used to predict the postoperative HRQoL of patients undergoing bariatric surgery. We explored the feasibility and capacity of a deep learning method, i.e., convolution neural network, to predict long-term HRQoL after bariatric surgery using a national register. The study can only be achieved based on a well-maintained and high-quality longitudinal database with long-term follow-up like SOReg [22].

Our results indicate that 5-year HRQoL after bariatric surgery may be well predicted preoperatively for some scale domains like PF, BP, GH, VT, MH, MCS, and OP. In our study, we aimed to evaluate and predict the quality of life of patients after bariatric surgery. Some patients were not "satisfied" even when they lost weight. Other factors, such as complications during follow-up and preoperative pharmacologic drug treatment, are associated with a change of the quality of life after bariatric surgery, whereas age, sex, and preoperative metabolic comorbidity may also play a role [11,23–25]. Our findings may provide important information for postoperative care and rehabilitation for this group of patients.

Our research question was about predicting continuous outcomes using supervised deep learning methods, which could be converted to a question of supervised two- or multi-class classification, i.e., to predict whether the quality of life of the patient has improved, remained unchanged, or deteriorated. Although the precision of prediction might be reduced in classification, the accuracy might be enhanced, and the method might be more applicable for clinical use. We would like to investigate the question in future studies.

There has been a warning that healthcare researchers should not be overly enthralled by the promises of deep learning methods [26]. Therefore, to avoid abusing the deep learning method in our study, we also compared the performance of the CNN model with a conventional statistical learning method for continuous variables, using a multivariate linear regression model. Although the conventional statistical methods require sometimes complex processing (feature engineering) to extract the requisite discriminative features, they may provide more interpretable results compared to the deep learning methods. In contrast, the biggest advantage of deep learning methods is that they try to learn high-level features from data in an incremental manner, which eliminates the need for domain expertise and hard-core feature extraction. However, the generalizability of deep learning models relies largely on the data they learned, and overfitting on unseen data is more apparent, as observed in our study. Although there are some ways in which we may reduce overfitting in deep learning models, the rule of thumb is to use more training data.

There are potential limitations to our study. In total, 28,293 patients underwent surgery for a primary gastric bypass between 2008 and 2012 and had a follow-up longer than 5 years when the study was initiated. However, only less than one quarter of the patients who had complete HRQoL information could be used for the machine learning. Compared to the patients who had no or incomplete HRQoL information, the patients with complete relevant data were older (42.7 ± 11.0 vs. 40.4 ± 10.8 years), had fewer males (21.2% vs. 25.1%), and lower BMI (42.3 ± 5.2 vs. 42.8 ± 5.5 kg/m^2). These factors have already been shown to influence HRQoL [27–29]. Because of these systematic differences in HRQoL between the patients with and without HRQoL measures, the generalizability of our CNN model may be questionable. The missing information needs to be imputed in the future for deep machine learning. We would also point out that the CNN built in our study was only based on features from gastric bypass patients, which cannot be generalized to other surgical procedures or health conditions. The application of CNN in predicting prognosis after surgeries still needs to be investigated using large data from the real world.

5. Conclusions

CNN can be used to predict long-term HRQoL after bariatric surgery based on the baseline features of patients. The performance of the CNN was found to be better than the traditional multivariate linear regression model; however, its overfitting on unseen data needs to be mitigated by using more features of patients or greater use of training data in the future.

Supplementary Materials: The following are available online at http://www.mdpi.com/2077-0383/8/12/2149/s1, Figure S1: Physical functioning (PF) scores before and after bariatric surgery of 6687 patients used in the study. Figure S2: Structure of the conventional neural network (CNN) model used in the study. Figure S3: Process of training, validation and evaluation

Author Contributions: Data curation, M.R.; Formal analysis, Y.C.; Funding acquisition, Y.C.; Investigation, Y.C., M.R., S.M. and I.N.; Methodology, Y.C. and M.R.; Project administration, J.O. and I.N.; Resources, J.O. and I.N.; Software, Y.C.; Validation, Y.C. and S.M.; Visualization, Y.C.; Writing—original draft, Y.C., M.R. and I.N.; Writing—review and editing, Y.C., M.R., S.M., J.O. and I.N.

Funding: Yang Cao's work was supported by Örebro Region County Council (OLL-864441). 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. Kopelman, P.G.; Caterson, I.D.; Dietz, W.H. *Clinical Obesity in Adults and Children*; John Wiley & Sons: New York, NY, USA, 2009.
2. Organization, W.H. *Obesity: Preventing and Managing the Global Epidemic*; World Health Organization: Geneva, Switzerland, 2000.
3. Prospective Studies Collaboration; Whitlock, G.; Lewington, S.; Sherliker, P.; Clarke, R.; Emberson, J.; Halsey, J.; Qizilbash, N.; Collins, R.; Peto, R. Body-mass index and cause-specific mortality in 900 000 adults: Collaborative analyses of 57 prospective studies. *Lancet* **2009**, *373*, 1083–1096. [CrossRef] [PubMed]
4. Guh, D.P.; Zhang, W.; Bansback, N.; Amarsi, Z.; Birmingham, C.L.; Anis, A.H. The incidence of co-morbidities related to obesity and overweight: A systematic review and meta-analysis. *BMC Public Health* **2009**, *9*, 88. [CrossRef] [PubMed]
5. Wawok, P. Evaluation of Quality of Life of patient with obesity after surgical procedure. *Eur. J. Med. Res.* **2009**, *14*, 114.
6. Kawatkar, A.A.; Nichol, M.B. The Impact of Obesity on Medical Expenditure and Health Related Quality of Life. *Value Health* **2009**, *12*, A134. [CrossRef]
7. Kolotkin, R.L.; Williams, V.S.L.; Ervin, C.M.; Williams, N.; Meincke, H.H.; Qin, S.S.; Smith, L.V.; Fehnel, S.E. Validation of a new measure of quality of life in obesity trials: Impact of Weight on Quality of Life-Lite Clinical Trials Version. *Clin. Obes.* **2019**, *9*, e12310. [CrossRef]
8. Dreber, H.; Thorell, A.; Reynisdottir, S.; Hemmingsson, E. Health-Related Quality of Life 5Years After Roux-en-Y Gastric Bypass in Young (18-25Years) Versus Older (26Years) Adults: A Scandinavian Obesity Surgery Registry Study. *Obes. Surg.* **2019**, *29*, 434–443. [CrossRef]
9. Consensus Development Conference Panel. NIH conference. Gastrointestinal surgery for severe obesity. Consensus Development Conference Panel. *Ann. Intern. Med.* **1991**, *115*, 956–961. [CrossRef]
10. Mazer, L.M.; Azagury, D.E.; Morton, J.M. Quality of Life After Bariatric Surgery. *Curr. Obes. Rep.* **2017**, *6*, 204–210. [CrossRef] [PubMed]
11. Andersen, J.R.; Aasprang, A.; Karlsen, T.I.; Natvig, G.K.; Vage, V.; Kolotkin, R.L. Health-related quality of life after bariatric surgery: A systematic review of prospective long-term studies. *Surg. Obes. Relat. Dis.* **2015**, *11*, 466–473. [CrossRef]
12. Wimmelmann, C.L.; Dela, F.; Mortensen, E.L. Psychological predictors of mental health and health-related quality of life after bariatric surgery: A review of the recent research. *Obes. Res. Clin. Pract.* **2014**, *8*, e314–e324. [CrossRef]
13. Masheb, R.M.; White, M.A.; Toth, C.M.; Burke-Martindale, C.H.; Rothschild, B.; Grilo, C.M. The prognostic significance of depressive symptoms for predicting quality of life 12 months after gastric bypass. *Compr. Psychiatry* **2007**, *48*, 231–236. [CrossRef] [PubMed]
14. Hedenbro, J.L.; Naslund, E.; Boman, L.; Lundegardh, G.; Bylund, A.; Ekelund, M.; Laurenius, A.; Moller, P.; Olbers, T.; Sundbom, M.; et al. Formation of the Scandinavian Obesity Surgery Registry, SOReg. *Obes. Surg.* **2015**, *25*, 1893–1900. [CrossRef] [PubMed]
15. Matsugu, M.; Mori, K.; Mitari, Y.; Kaneda, Y. Subject independent facial expression recognition with robust face detection using a convolutional neural network. *Neural Netw.* **2003**, *16*, 555–559. [CrossRef]
16. Zeiler, M.D. ADADELTA: An adaptive learning rate method. *arXiv* **2012**, arXiv:1212.5701.
17. Raschka, S.; Mirjalili, V. *Python Machine Learning*; Packt Publishing Ltd.: Birmingham, UK, 2017.
18. Wilson, A.C.; Roelofs, R.; Stern, M.; Srebro, N.; Recht, B. The Marginal Value of Adaptive Gradient Methods in Machine Learning. *Adv. Neural Inf.* **2017**, *30*, 4148–4158.
19. Lee, S.K.; Son, Y.J.; Kim, J.; Kim, H.G.; Lee, J.I.; Kang, B.Y.; Cho, H.S.; Lee, S. Prediction Model for Health-Related Quality of Life of Elderly with Chronic Diseases using Machine Learning Techniques. *Healthc. Inf. Res.* **2014**, *20*, 125–134. [CrossRef]
20. Kumar, S.; Rana, M.L.; Verma, K.; Singh, N.; Sharma, A.K.; Maria, A.K.; Dhaliwal, G.S.; Khaira, H.K.; Saini, S. PrediQt-Cx: Post treatment health related quality of life prediction model for cervical cancer patients. *PLoS ONE* **2014**, *9*, e89851. [CrossRef]
21. Persson, P.-V.; Rietz, H. *Predicting and Analyzing Osteoarthritis Patient Outcomes with Machine Learning*; Lund University: Lund, Sweden, 2017.

22. Tao, W.; Holmberg, D.; Naslund, E.; Naslund, I.; Mattsson, F.; Lagergren, J.; Ljung, R. Validation of Obesity Surgery Data in the Swedish National Patient Registry and Scandinavian Obesity Registry (SOReg). *Obes. Surg.* **2016**, *26*, 1750–1756. [CrossRef]
23. Hell, E.; Miller, K.A.; Moorehead, M.K.; Norman, S. Evaluation of health status and quality of life after bariatric surgery: Comparison of standard Roux-en-Y gastric bypass, vertical banded gastroplasty and laparoscopic adjustable silicone gastric banding. *Obes. Surg.* **2000**, *10*, 214–219. [CrossRef]
24. Dymek, M.P.; Le Grange, D.; Neven, K.; Alverdy, J. Quality of life after gastric bypass surgery: A cross-sectional study. *Obes. Res.* **2002**, *10*, 1135–1142. [CrossRef]
25. Major, P.; Matlok, M.; Pedziwiatr, M.; Migaczewski, M.; Budzynski, P.; Stanek, M.; Kisielewski, M.; Natkaniec, M.; Budzynski, A. Quality of Life After Bariatric Surgery. *Obes. Surg.* **2015**, *25*, 1703–1710. [CrossRef] [PubMed]
26. Chen, D.; Liu, S.J.; Kingsbury, P.; Sohn, S.; Storlie, C.B.; Habermann, E.B.; Naessens, J.M.; Larson, D.W.; Liu, H.F. Deep learning and alternative learning strategies for retrospective real-world clinical data. *NPJ Digit. Med.* **2019**, *2*, 43. [CrossRef] [PubMed]
27. Chang, C.Y.; Hung, C.K.; Chang, Y.Y.; Tai, C.M.; Lin, J.T.; Wang, J.D. Health-related quality of life in adult patients with morbid obesity coming for bariatric surgery. *Obes. Surg.* **2010**, *20*, 1121–1127. [CrossRef] [PubMed]
28. Kolotkin, R.L.; Crosby, R.D.; Williams, G.R. Health-related quality of life varies among obese subgroups. *Obes. Res.* **2002**, *10*, 748–756. [CrossRef] [PubMed]
29. White, M.A.; O'Neil, P.M.; Kolotkin, R.L.; Byrne, T.K. Gender, race, and obesity-related quality of life at extreme levels of obesity. *Obes. Res.* **2004**, *12*, 949–955. [CrossRef]

Journal of
Clinical Medicine

MDPI

Article

Clinical Validation of Innovative Optical-Sensor-Based, Low-Cost, Rapid Diagnostic Test to Reduce Antimicrobial Resistance

Suman Kapur [1,*,†], Manish Gehani [1,†], Nagamani Kammili [2], Pankaj Bhardwaj [3], Vijayalakshmi Nag [3], Sudha M. Devara [2] and Shashwat Sharad [4,*]

1 Department of Biological Sciences, Birla Institute of Technology and Science, Pilani, Hyderabad Campus, Hyderabad, Telangana 500078, India; dr.manishgehani@gmail.com
2 Department of Microbiology, Gandhi Medical College and Hospital, Hyderabad, Telangana 500003, India; nagamaniy2k03@rediffmail.com (N.K.); sudha_devara@yahoo.com (S.M.D.)
3 All India Institute of Medical Sciences, Jodhpur, Rajasthan 342005, India; pankajbhardwajdr@gmail.com (P.B.); vijayalakshmi005@gmail.com (V.N.)
4 Center for Prostate Disease Research, Department of Surgery, Uniformed Services University of the Health Sciences and the Walter Reed National Military Medical Center, Bethesda, MD 20817, USA
* Correspondence: skapur@hyderabad.bits-pilani.ac.in (S.K.); ssharad@cpdr.org (S.S.); Tel.: +91-40-6630-3563 (S.K.); +1-240-694-4931 (S.S.)
† Suman Kapur and Manish Gehani contributed equally to this work.

Received: 18 October 2019; Accepted: 27 November 2019; Published: 1 December 2019

Abstract: The antibiotic susceptibility test determines the most effective antibiotic treatment for bacterial infection. Antimicrobial stewardship is advocated for the rational use of antibiotics to preserve their efficacy in the long term and provide empirical therapy for disease management. Therefore, rapid diagnostic tests can play a pivotal role in efficient and timely treatment. Here, we developed a novel, rapid, affordable, and portable platform for detecting uropathogens and reporting antibiogram to clinicians in just 4 h. This technology replicates the basic tenets of clinical microbiology including bacterial growth in indigenously formulated medium, and measurement of inhibition of bacterial growth in presence of antibiotic/s. Detection is based on chromogenic endpoints using optical sensors and is analyzed by a lab-developed algorithm, which reports sensitivity to the antibiotic's panel tested. To assess its diagnostic accuracy, a prospective clinical validation study was conducted in two tertiary-care Indian hospitals. Urine samples from 1986 participants were processed by both novel/index test and conventional Kirby Bauer Disc Diffusion method. The sensitivity and specificity of this assay was 92.5% and 82%, respectively ($p < 0.0005$). This novel technology will promote evidence-based prescription of antibiotics and reduce the burden of increasing resistance by providing rapid and precise diagnosis in shortest possible time.

Keywords: urinary tract infection; rapid culture; antibiotic susceptibility testing (AST), evidence-based prescription; antibiotics; antimicrobial resistance (AMR), rapid diagnostics

1. Introduction

Healthcare challenges faced by developing countries are vastly different from those in developed nations. With very limited budget for healthcare, developing countries have not been able to put up any significant infrastructure to address their huge disease burden. In vitro diagnostic (IVD) tests provide the basis for most medical decision-making and play crucial role in limiting healthcare costs, since appropriate diagnostic tests performed in a timely manner i) improve patient care, ii) contribute to protecting consumers' health, iii) help to limit healthcare spending, iv) reduce the risk of trial-and-error treatment and over-prescription, v) shorten the time before treatment begins, and vi)

decrease the length of hospital stays. Appropriate diagnosis can improve the effectiveness of treatments and avoid long-term complications for the infected patient. India harbors the world's largest burden of drug-resistant pathogens. Easy access, availability, and higher consumption of medicines have led to a disproportionately higher incidence of inappropriate use of antibiotics and greater levels of antimicrobial resistance (AMR) compared to developed countries [1]. It has been shown that the health sector in India suffers from gross inadequacy of funds, which will further result in conditions favorable for the development of drug resistance [2]. The high resistance of pathogens in the country, even to newer antibiotics, has led to the emergence of superbugs like New Delhi Metallo-beta-lactamase (NDM-1) [3]. By 2050, 2 million Indians are projected to die as a result of AMR [3]. Indians are the largest consumers of antibiotics worldwide, despite a decline in communicable diseases [3], due to a liberal policy for over-the-counter sale of antibiotics and irrational prescription of antibiotics. A study by Ganguly et al. highlighted the importance of rationalizing antibiotic use to limit AMR in India [4]. Irrational prescription happens due to a lack of fast point-of-use tests for evidence-based prescription, lack of infrastructure for bacterial culture and antibiotic susceptibility test (AST), and lack of awareness worldwide. Selective pressure from inappropriate use of antibiotics can lead to resistance via the emergence of mutant strains [5]. Unavailability of rapid point-of-use diagnostics to distinguish bacterial infections and suggest appropriate therapy is a major reason for irrational prescriptions of antibiotic/s.

Urinary tract infections (UTIs) lead to 23% of all antibiotic prescriptions in primary healthcare. Even in India, UTIs account for about 8.1 million prescriptions each year. Diagnosis of UTI is a multistep process including determination of pathogen load, identification, and AST requiring culture of sample, which takes around 48–72 h. Even if high-throughput automated systems like Vitek 2, Microscan Walkaway, or Phoenix are used, the results are not available faster than 28 h [6,7]. Conventional urine culture and AST method is not accessible to most clinicians practicing in low-resource settings. Even with access to lab testing facilities, but in the absence of any rapid test, clinicians are forced to prescribe antibiotics empirically. The empirical antibiotics used in the first 48–72 h prove to be ineffective against infection in approximately 33% cases [8–10]. Unresolved, relapsed UTIs tend to be resistant to previously used antibiotics [11]. Nearly 23% to 33% of the prescriptions for UTIs have been found to have no clinical justification [8–10]. Moreover, UTIs are also caused by non-bacterial organisms such as Candida (3% cases) [12], Trichomonas (17% cases) [13], Chlamydia (~16% cases) [14], and rarely Mycobacterium, Schistosoma haematobium, Adenovirus, BK polyomavirus, and mycoplasma [15], and cannot be treated by antibiotics that are empirically prescribed. Such unnecessary drug use is often harmful, and results in multidrug-resistant infections and reduced options for antimicrobial therapy [16]. An urgent need is perceived for developing suitable field operable test for prescribing targeted antibiotics [17]. Addressing the menace of antimicrobial resistance needs a scalable rapid diagnostic test, which gives detection, identification, quantification, and phenotypic antimicrobial susceptibility of bacteria within a minimum turn-around time and has an integrated technology platform for clinical adoption. This test should show high sensitivity and specificity, should be low cost for adoption in low-resource settings, and should be easy to use with minimal training [18,19].

Assays used for upstream screening to improve diagnostic yield of positive samples, like gram staining, dipstick with leukocyte esterase and nitrite, pus cell count [20], urine analysis and microscopy [21], chlorhexidine [22], interleukin-8 [23], Griess test [24], microstix [25], serum procalcitonin level [26], and urine catalase-based uriscreen test [27] have shown poor sensitivity and specificity. Novel antibody-based lateral flow immunoassay (RapidBac) [28], chromogenic limulus amoebocyte lysate assay [29], and flow cytometry-based systems (Accuri-6, UF 100, UF-1000i) [18,30,31] have shown high sensitivity and specificity but they do not provide identification of bacteria and its antimicrobial susceptibility. Forward light scattering systems like Uro-Quick (Alifax) and BacterioScan model 216 (BacterioScan Inc., St. Louis, MO, USA) provide detection of bacteria with antimicrobial susceptibility but do not identify the causative bacteria [32,33].

Molecular and proteomic technologies require overnight incubation on culture plates, and do not provide antimicrobial susceptibility [18]. Matrix Assisted Laser Desorption Ionization-Time Of Flight

(MALDI-TOF) is expensive to install [34], Fluorescence In Situ Hybridization (FISH) requires multiple probes for all uropathogens [18], while multiplex Polymerase Chain Reaction (PCR) platforms like GeneXpert Omni and Cepheid [18] do not provide quantification of significant bacteriuria and need multiple probes for all uropathogens. Application of these test to direct urine testing needs extensive sample preparation.

Genetic signature identification Confirming Active Pathogens Through Unamplified RNA Expression (CAPTURE) assay [35] identifies bacteria but does not give antimicrobial susceptibility. Time-lapse microscopy-based systems like oCelloScope (Phillips BioCell) [36] and Accelerate ID/AST (Accelerate Diagnostics) [37] provide both identification and antimicrobial susceptibility, but phenotypic measures for identification in direct urine are not precise and they are not easy to use in a clinical lab setting. Integrated microfluidic-Biosensor assays based on ion mobility spectrometry or colorimetric sensor arrays are cost effective and sensitive but their results get confounded by urine variability and in presence of low bacterial count [18]. Most of the newer technologies mentioned above are neither easy to use nor affordable in a resource-poor setting like public hospitals of India. In this study, we evaluated a rapid, portable, easy-to-use, less resource-intensive, and affordable technology, which provides bacterial identification and AST results within 4 h. This new technology integrates the basic tenets of clinical microbiology including bacterial growth in a medium optimized for uropathogens and measurement of inhibition of bacterial growth in presence of specific antibiotic, with detection of bacteria based on chromogenic endpoint by enzymatic hydrolysis of specific media cocktails by UTI causing bacteria. The optical sensor-based measurement of endpoint output is analyzed using indigenous software, based on a lab-developed statistical algorithm, which reports both the sensitivity of the pathogen to a customizable panel of antibiotics and bacterial load in the sample. This integrated technology platform can be used for diagnosing UTIs caused by bacteria and for suggesting effective antibiotics in all types of clinical settings as a preliminary triage test [38] to promote evidence-based prescription and minimize irrational use of antibiotics. The low cost of the test obliviates the need for upstream screening with poor sensitivity screening tests and promotes scalability for use in mass population. The objective of the present study was to evaluate the diagnostic accuracy of the novel test in UTI cases as compared to the gold standard urine culture method.

2. Materials and Methods

2.1. Study Design, Setting, and Population

The study was conducted over a 2-year period from January 2017 to December 2018, simultaneously in Gandhi Medical College and Hospital, Secunderabad located in Southern India, and All India Institute of Medical Sciences (AIIMS), Jodhpur located in Northern India. To ensure sufficient case load for achieving required sample size and to ensure that good lab practices are followed, Laboratory of Gandhi Hospital, which is the referral laboratory of State of Telangana, and AIIMS, which is a premiere tertiary care hospital, were chosen for this study.

2.2. Ethical Approval

The study was reviewed and approved by Institutional Ethics Review Committee of both institutions. Objectives of the study were explained to all participants in their native language and they were enrolled after obtaining a written informed consent. The study was conducted according to the principles expressed in the Declaration of Helsinki.

2.3. Study Oversight

This prospective clinical validation study was designed to evaluate diagnostic accuracy of the novel/index test with the reference gold standard urine culture and AST method. Eligible participants were referred by clinicians for urine culture and sensitivity test, based on a provisional diagnosis of UTI. Patients who received antibiotics in the preceding two weeks or had indwelling or suprapubic

catheter were excluded. Consenting participants were evaluated in microbiology laboratory by history taking and review of medical records.

2.4. Test Methods for Bacterial Culture and Identification

Clean-catch mid-stream urine samples were collected from each enrolled participant in a sterile container and divided into two parts under sterile conditions. One part was used for routine urine culture and AST and the second for conducting the index test in the hospital premises itself. All samples were processed within 2 h of collection to avoid contamination/bacterial growth.

The index test was the novel test designed for direct quantitative detection and antibiotic sensitivity of bacteria found in human urine [39,40]. The test identifies common UTI-causing bacteria, namely *Escherichia coli*, *Klebsiella*, *Pseudomonas*, *Enterococcus*, *Proteus*, and *Staphylococcus* sp. This rapid method replicates the basic tenets of clinical microbiology, namely (1) growth of bacteria in a specialized medium, and (2) measuring the inhibition of growth of bacteria in the presence of an antibiotic. Detection is based on chromogenic endpoints. The output was analyzed using lab-developed algorithm-based software, which reports the sensitivity of the pathogen to the panel of antibiotics tested. The urine sample was collected in a sterile container. To harvest the bacteria, 10 mL urine was filtered through a sterile syringe with the help of a micro-filter attached to it and filtrate was discarded. After that, BITGEN, specially designed media for accelerated growth of uropathogens, was pushed through the filter in the vial to recover bacteria from the filter, shaken well, and then closed with the dropper cap. The bacteria were harvested in 3 mL of proprietary BITGEN medium. This was then set side at room temperature for about 5 min. Subsequently, four drops (~110–120 µL) of proprietary BITGEN medium containing harvested bacterial suspension was added into all the three strips—one pre-functionalized strip for identification of bacteria and two different 8-well strips, pre-loaded with antibiotics. All the strips were resealed and incubated at 37 °C for 4 h. A 4-h incubation period was found to be sufficient for all commonly found uropathogens accounting for 98% cases of UTIs. The media was optimized for nutrients and supports growth up to 8 h with a start bacterial number of 10^5 cells/mL [39]. BITGEN is a proprietary media that has chromogens sensitive to bacterial growth even at low numbers of bacteria and for rapid culture. The enzymatic hydrolysis of specific media cocktails used in this proprietary media metabolizes the chromogens. For identification of bacteria, the 8 wells in the identification strip had a cocktail of specific substrates, which were metabolized by specific bacterial types. Growth of bacteria in the well led to end product formation during the 4-h incubation. The use of optical sensor enables measuring of all color combinations and the lab-developed analytical software interprets the identification of the bacteria based on specific chromogenic endpoints produced as a consequence of specific metabolic activity of each bacterial type. For both identification and AST, the sample was loaded at the same time and incubated for the same length of time.

To identify susceptibility of pathogen, the above-mentioned two pre-functionalized antibiotic strips were used. Each of the antibiotic strips had 8 compartments and, except the first compartment (or reference well) of each of the two antibiotic strips, all the remaining 14 compartments were subjected to preloading by the chosen antibiotics. The preloaded antibiotics used were Amoxicillin, Gentamicin, Amikacin, Cefepime, Ofloxacin, Ciprofloxacin, Ceftriaxone, Piperacillin-Tazobactum, Cefotaxime, Cefuroxime, Tobramycin, Levofloxacin, Cefazolin, and Imipenem. The concentration and composition of the antibiotics were chosen as per Clinical and Laboratory Standards Institute (CLSI) guidelines [41].

In the case that the urine sample had pathogens, it was reflected in the first well of the antibiotic strips, referred to as the reference well of both the antibiotic strips as there is no inhibition of bacterial growth in this well. As per phenotypic AST of bacteria present, the remaining 14 compartments showed varied levels of bacterial growth depending on the bacterial susceptibility to the chosen antibiotics. The bacterial growth within the preloaded antibiotic compartment was represented by a change in color of the BITGEN, measured by chromogenic and nephelometric endpoints using an array of 64 photodiodes in an electronic optical sensor. The intensity of the color is a measure of the number of growing cells in the presence and absence of a particular antibiotic. The sensor output was

analyzed using a proprietary lab-developed statistical algorithm, pre-installed on the reader, which provides ready-to-use results for sensitivity of the pathogen to the antibiotics tested, both as a display on liquid crystal display (LCD) screen and a printout for permanent records. The reader was also enabled to transfer results to other storage devices using a wireless module and/or a universal serial bus (USB) interface. In case of insufficient growth, the analytical software prompts for incubating for one additional hour and then if no growth is detected, the software reports the sample to be negative for presence of bacteria.

Further, for reference, standard universally accepted, conventional gold standard urine culture and Kirby Bauer method for AST was chosen. First, 10 µL of each urine sample was streaked on a chromogenic culture medium, chromID® CPS Elite Translucent using a calibrated ni-chrome wire loop of 4 mm by semi-quantitative method using surface streaking. The inoculated plates were incubated for 18–24 h at 37 °C. After incubation, in case growth of colonies was up to the tertiary streaking, it was considered as significant bacteriuria with 10^5 Colony Forming Units (CFU)/mL. Positive cultures were further processed for determining the AST by Kirby Bauer Disc Diffusion Method as per Clinical and Laboratory Standards Institute guidelines [41]. A suspension of each isolate was prepared to a McFarland standard and spread over Muller Hilton Agar using lawn culture method. Himedia discs with defined concentrations of antibiotics were placed over the culture. After incubation for 18 to 24 h at 37 °C, zones of growth inhibition around each antibiotic disc were measured to the nearest millimeter and a reference table was used to determine susceptibility. The American Type Culture Collection (ATCC) bacterial strains, namely *Enterococcus faecalis*, *Escherichia coli*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, and *Staphylococcus aureus*, were used for quality control in the entire process.

The cut-off for labelling both index as well as reference test as positive was pre-specified as 10^5 CFU/mL based on Infectious Diseases Society of America guidelines [42]. Neither the team performing the index test nor the one conducting urine culture and sensitivity was provided any clinical information about the participant/s. Both teams were also not informed about the results of the other test and, hence, the index test was conducted in a completely blinded manner.

2.5. Data Analysis

Collected data and results from both tests for each participant were compiled and analyzed by Statistical Package for the Social Sciences (SPSS) software (version 24). A contingency table was used for determining diagnostic accuracy and kappa statistics was used for agreement analysis. Further, 95% confidence interval (CI) was used to describe diagnostic accuracy, with p values of <0.05 considered as significant. Sample size was calculated to be 600 for estimating the sensitivity of the index test, based on a precision of 4% and confidence level of 95%, when the sensitivity of the new test was expected to be at least 50%. "Best-case scenario method" was used for indeterminate results and mixed growth. Samples with rare species, budding yeast cells, and contaminated samples were removed from final analysis for a "complete case analysis". No analysis of variability in diagnostic accuracy was performed with respect to age group or department, as it was not pre-specified in the study. The raw data generated from the study which was used to analyze these results has been made publicly available as a safe harbor file in online repository "Harvard Dataverse" [43].

2.6. Reagents

Analytical-grade chemicals required for preparation of BITGEN, identification strips, and antibiotic strips were procured from Sigma Chemicals, St Louis, MO, USA. Chromogenic culture media, Muller Hilton media, and antibiotic discs were procured from Himedia, India; chromID® CPS Elite Translucent from BioMérieux, France; 8-well strips and syringe filters from NUNC, Denmark; sterile syringes from Dispovan, India. The scanner/reader machine for novel test was obtained from Micro Lab Instruments, Ahmedabad, India. The bacterial strains *Enterococcus faecalis* (ATCC29212), *Escherichia coli* (ATCC25922), *Klebsiella pneumoniae* (ATCC13883), *Pseudomonas aeruginosa* (ATCC27853), and *Staphylococcus aureus* (ATCC25923) were purchased from Himedia, India.

3. Results

3.1. Study Characteristics

Overall, 2001 eligible participants (1030 in AIIMS and 971 in Gandhi Hospital) were identified and 1986 participants (1022 in AIIMS and 964 in Gandhi Hospital) were enrolled in the study. Data of 1835 participants (982 in AIIMS and 853 in Gandhi Hospital) were included in the final analysis. A total of 55 samples (20 from AIIMS and 35 from Gandhi Hospital) with low sample volume could not be processed by the index test.

There were no indeterminate results reported by the index test in both the hospitals. One hundred and eleven participants (97 in AIIMS and 14 in Gandhi Hospital) with indeterminate reference standard urine culture results were reported as having no bacterial growth and were reclassified as true negatives using best-case scenario. Samples with mixed growth in both index and reference standard tests were considered positive for UTI. Fifty-five samples (5 from AIIMS and 50 from Gandhi Hospital) were reported as contaminated and not considered for final analysis. Since the index test is designed for identifying the most common bacteria only, 19 samples (6 from AIIMS and 13 from Gandhi Hospital) with budding yeast cells and 22 samples (9 from AIIMS and 13 from Gandhi Hospital) with rare bacteria (*Citrobacter*, *Acinetobacter*, *Morganella*, and *Providencia*) were also excluded from the final analysis. Thus, a total of 96 cases were excluded from final analysis after performing both tests (Figure 1). Table 1 summarizes the mean age, gender distribution, and referring departments. The majority of participants had cystitis, and more male patients were referred at AIIMS than Gandhi Hospital. Ninety-seven cases (10%) cases in AIIMS cohort and nine cases (1%) in Gandhi Hospital cohort had progressed to frank pyelonephritis.

Table 1. Characteristics of participants.

Characteristics	AIIMS [1] (*n* = 982)	Gandhi (*n* = 853)
Demographic characteristics		
Age (in years)		
Mean age	43.4	35.7
Minimum age	<1	1
Maximum age	95	90
Gender		
Male	631 (64.3%)	399 (46.8%)
Female	351 (35.7%)	454 (53.2%)
Referring Department		
Medical Specialties	193	481
Surgical Specialties	632	94
Pediatrics	67	138
Obstetrics and Gynecology	87	140
Radio-diagnosis	3	0
Clinical Syndrome		
Pyelonephritis	10%	1%
Cystitis	90%	99%

[1] AIIMS = All India Institute of Medical Sciences.

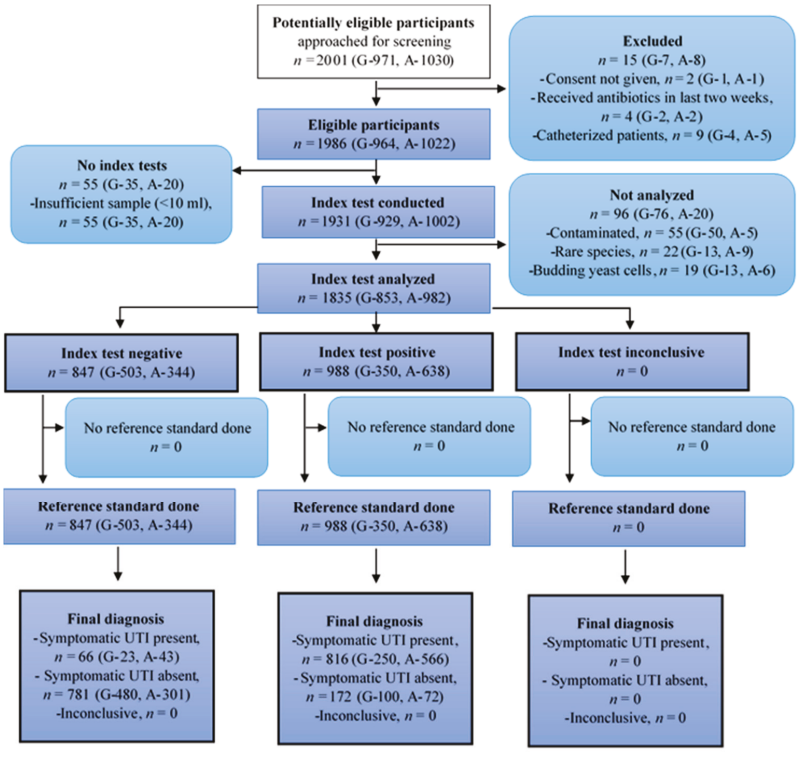


Figure 1. Flow of participants through the study—Standards for Reporting Diagnostic Accuracy (STARD) diagram (G = Gandhi Hospital, A = AIIMS).

3.2. Test Performance

There was no time gap between processing of samples by both tests. No adverse event occurred while performing index test or reference standard test since only urine sample collection was involved. In AIIMS, 609 cases, while in Gandhi Hospital, 273 cases, were diagnosed with symptomatic UTI based on positive culture results by conventional method. Furthermore, 953 participants (373 in AIIMS and 580 in Gandhi Hospital) with symptoms of UTI, showed low colony count on culture plates. Out of these, the index test reported 172 cases (72 in AIIMS and 100 in Gandhi Hospital) as positive, which were otherwise reported as negative by conventional method, 48 h post incubation.

3.2.1. Diagnostic Accuracy

AIIMS cohort showed a higher sensitivity (92.9%) while Gandhi Hospital cohort showed a marginally higher specificity (82.8%). The sensitivity and specificity in both the validation sites were within 95% confidence interval of the other hospital (Table 2). The sensitivity and specificity obtained by use of the novel test was well within the stipulated limits laid down in the recommendations issued by the European Urinalysis Guidelines for rapid tests.

Table 2. Comparison of test results obtained by novel test and urine culture method.

Contingency Tables			
AIIMS (*n* = 982)	**Urine Culture Positive**	**Urine Culture Negative**	**Total**
Index test Positive	566 (92.9%)	72 (19.3%)	638
Index test Negative	43 (7.1%)	301 (80.7%)	344
Total	609	373	982
Gandhi (*n* = 853)	**Urine Culture Positive**	**Urine Culture Negative**	**Total**
Index test Positive	250 (91.6%)	100 (17.2%)	350
Index test Negative	23 (8.4%)	480 (82.8%)	503
Total	273	580	853
Combined (*n* = 1835)	**Urine Culture Positive**	**Urine Culture Negative**	**Total**
Index test Positive	816 (92.5%)	172 (18.0%)	988
Index test Negative	66 (7.5%)	781 (82.0%)	847
Total	882	953	1835
Diagnostic Accuracy			
Parameters	**AIIMS (*n* = 982)**	**Gandhi (*n* = 853)**	**Combined (*n* = 1835)**
Sensitivity	92.9% (95% CI: 90.6–94.8%)	91.6% (95% CI: 87.6–94.6%)	92.5% (95% CI: 90.6–94.2%)
Specificity	80.7% (95% CI: 76.3–84.6%)	82.8% (95% CI: 79.4–85.8%)	82.0% (95% CI: 79.4–84.3%)
Agreement Analysis			
Parameters	**AIIMS (*n* = 982)**	**Gandhi (*n* = 853)**	**Combined (*n* = 1835)**
Kappa value [1]	0.748	0.692	0.741
Standard error [1]	0.022	0.025	0.016
p value	<0.0005	<0.0005	<0.0005

[1] Kappa value and its standard error measures agreement between results of two dichotomous variables (here two diagnostic tests providing positive or negative results).

3.2.2. Agreement Analysis

Good agreement was observed at both validation sites, as seen by a Kappa = 0.741. The observed agreement is statistically significant as reflected by a *p* value of <0.0005.

3.3. Identification of Bacteria

The index test correctly reported the causative bacteria as reported positive by urine culture in 82% cases in AIIMS cohort and 80% cases in Gandhi Hospital cohort (Table 3). In Gandhi Hospital, four cases of *Streptococcus* were not reported by the index test as it is not designed to identify the same. Out of 17 mixed growth in Gandhi cohort, the index test identified seven as individual bacteria, while among 49 in AIIMS cohort, it identified 41 as individual bacteria.

Table 3. Identification of bacteria in the two cohorts.

Identification of Bacteria (Single Species Identification)							
AIIMS	***E. coli***	***Enterococcus***	***Klebsiella***	***Proteus***	***Staphylococcus***	***Pseudomonas***	**Total**
"*n*" (based on urine culture)	324	104	79	4	1	48	560
% correct identification by index test	93%	74%	68%	75%	100%	48%	82%
Gandhi	***E. coli***	***Enterococcus***	***Klebsiella***	***Proteus***	***Staphylococcus***	***Pseudomonas***	**Total**
"*n*" (based on urine culture)	92	22	100	8	22	8	252
% correct identification by index test	85%	82%	83%	63%	50%	88%	80%

3.4. Antibiotic Susceptibility

The index test used the same set of 14 antibiotics for every sample, while AIIMS and Gandhi Hospital laboratories used specific antibiotics based on identified bacteria. Hence, only a subset of antibiotics overlapped for both tests. Further, as the conventional method relied on the choice of antibiotics by the microbiologist in-charge, sets of antibiotics tested in both tests were also not used

for all the samples tested. Antibiotics tested for at least 30 samples in both the tests were included in analysis presented in Table 4. The rapid index test correctly reported sensitivity and resistance to antibiotics in 91% and 96% cases, respectively, in AIIMS cohort, and these numbers were 87% in the case of sensitivity to tested antibiotics and 92% in the case of resistance to antibiotics reported for the Gandhi Hospital cohort.

Table 4. Comparison of antibiotic susceptibility report in the two cohorts using both tests.

AIIMS				
	AIIMS Result	**Total Tests Done Together**	**Agreement in Results**	**Disagreement in Results**
Gentamycin	R	149	141 (95%) [a]	8 (5%) [b]
	S	259	237 (92%) [a]	22 (8%) [c]
	I	7	0	7 [d]
Amikacin	R	65	65 (100%) [a]	0 [b]
	S	29	26 (90%) [a]	3 (10%) [c]
	I	3	0	3 [d]
Ciprofloxacin	R	49	43 (88%) [a]	6 (12%) [b]
	S	1	1 (100%) [a]	0 [c]
	I	0	0	0 [d]
Ceftriaxone	R	269	263 (98%) [a]	6 (2%) [b]
	S	99	85 (86%) [a]	14 (14%) [c]
	I	4	0	4 [d]
Piperacillin-Tazobactum	R	121	118 (98%) [a]	3 (2%) [b]
	S	286	266 (93%) [a]	20 (7%) [c]
	I	22	0	22 [d]
Cefazolin	R	34	24 (71%) [a]	10 (29%) [b]
	S	12	10 (83%) [a]	2 (17%) [c]
	I	0	0	0 [d]
Imipenem	R	56	56 (100%) [a]	0 [b]
	S	21	19 (90%) [a]	2 (10%) [c]
	I	3	0	3 [d]
Overall	R	743	710 (96%) [a]	33 (4%) [b]
	S	707	644 (91%) [a]	63 (9%) [c]
Gandhi				
	Gandhi Result	**Total Tests Done Together**	**Agreement in Results**	**Disagreement in Results**
Gentamycin	R	56	54 (96%) [a]	2 (4%) [b]
	S	120	103 (86%) [a]	17 (14%) [c]
	I	0	0	0 [d]
Amikacin	R	12	12 (100%) [a]	0 [b]
	S	27	23 (85%) [a]	4 (15%) [c]
	I	0	0	0 [d]
Cefepime	R	35	34 (97%) [a]	1 (3%) [b]
	S	10	8 (80%) [a]	2 (20%) [c]
	I	0	0	0 [d]
Piperacillin-Tazobactum	R	13	12 (92%) [a]	1 (8%) [b]
	S	21	20 (95%) [a]	1 (5%) [c]
	I	0	0	0 [d]
Cefotaxime	R	28	28 (100%) [a]	0 [b]
	S	14	14 (100%) [a]	0 [c]
	I	0	0	0 [d]
Levofloxacin	R	23	15 (65%) [a]	8 (35%) [b]
	S	8	6 (75%) [a]	2 (25%) [c]
	I	0	0	0 [d]
Cefazolin	R	56	50 (89%) [a]	6 (11%) [b]
	S	20	18 (90%) [a]	2 (10%) [c]
	I	0	0	0 [d]
Overall	R	223	205 (92%) [a]	18 (8%) [b]
	S	220	192 (87%) [a]	28 (13%) [c]

R = Resistant; S = Sensitive; and I= Intermediate; a—complete agreement, b—very major error, c—major error, and d—minor error. (Same test results either susceptible or resistant by both tests, were classified as "complete agreement" and result reported as resistant by culture and susceptible by novel test was labelled as "very major error"; susceptible by culture but resistant by novel test was labelled as "major error"; intermediate by culture and susceptible or resistant by novel test was labelled as "minor error"). Please note that no intermediate results were reported by novel test and by Gandhi Hospital culture reports.

4. Discussion

Treating patients, including UTIs caused by bacteria, is a challenging task, and development of rapid AST is very important to provide better healthcare services. Use of microbiological culture method and Kirby–Bauer disc diffusion tests are well established for diagnosis of UTIs in healthcare facilities worldwide. However, this entire method needs trained microbiologists and its major limitations are long turn-around time, resource intensiveness in the form of lab infrastructure, and requirement of cold chain for supply and storage of reagents [44]. In resource-constrained settings with poor or limited access to laboratory-based testing, performing urine culture and AST is not feasible. Therefore, initial antibiotic therapy in infectious diseases such as UTIs which accounts for ~40% cases of all infections as per World Health Organization (WHO), is mostly empirical. Hence, an alternative method, like the index test described herein, for reporting antibiotic sensitivity in a short period of 4 h, with no ancillary resource requirement, will not only be beneficial for patient care, but also curtail unnecessary antibiotic prescriptions. Additionally, availability of results in 4 h saves the repeat visit of patients to collect lab reports made available only after three days under best conditions and often even longer in remote and hard-to-reach geographical locations.

The high-cost, resource-intensive, non-portable, most commonly used automated systems Vitek 2 and MicroScan Walkaway provide AST and identification results in more than 28 h [6,7]. Reports evaluating the susceptibility of only Gram-negative bacilli to 11 antibacterial using these two systems showed the results in 92.7% of isolates and overall concurrence with the standard test being 94% with a 3.4% major error rate [45]. With reference to preventing emergence of resistance to antibiotics, they still pose a major limitation in terms of time taken to complete the identification and antibiogram profile of UTI causing pathogen. Most of the newer technologies tried for UTI [18] are facing limitations like the need for an overnight incubation, extensive initial sample preparation, need for an upstream screening test, and lack of integrated technology platform for clinical adoption. These technologies are expensive and not easy to use. Most of them do not give antimicrobial susceptibility. Previously tried strip-based tests also showed less sensitivity [46]. Even automated urine analyzers have resulted in low sensitivity [47]. In comparison, this index test is portable, can be used in all healthcare settings, costs less than 0.4 million INR (~5000 USD), needs no ancillary equipment or dedicated space, and provides ready-to-use antibiogram results and microbial identification within 4 h. The sensitivity and limitation of other tests are summarized in Table 5. The higher sensitivity, >90%, and specificity, >80%, of the index test, with kappa values indicating very good agreement with gold reference standard test, show that it has good diagnostic accuracy as a rapid test [48] for its role as a preliminary triage test and its intended use of diagnosing bacteriuria and preventing irrational prescription of antibiotics. Although the gold standard for diagnosing UTI remains as urine culture, the high cost, laboratory requirements, and long turnaround times (24–72 h) are its disadvantages. Further, in this type of testing, recognition and classification of bacteria is associated with the experience of laboratory technicians. The novel/index test developed is a simpler system and shows better agreement (Kappa = 0.741, significant substantial agreement) with the gold standard and is therefore best suited for routine use in clinical laboratories.

This novel test correctly identified sensitivity to multiple antibiotics in more than 75% instances (and in several cases with 100% accuracy), which then becomes the basis for evidence-based, rational use of antibiotics for specific therapy. The availability of results within 4 h will discourage unnecessary prescription of antibiotics in case of absence of bacterial disease and help the physician to prescribe antibiotics that are identified to be effective against the causative pathogen.

Table 5. Accuracy of different methods for estimating presence of bacteria.

Sl. No.	Method	Sensitivity	Specificity	Detection of Bacteriuria	Identification of Bacteria	Antibiotic Susceptibility	Limitations	Reference
1	MALDI-TOF	67% to 86%	Nearly 60%	Yes	Yes	No	Overnight incubation needed, expensive, extensive sample preparation	[18]
2	FISH	>96%	>96%	Yes	Yes	No	Requires multiple probes for all pathogens	[18]
3	PCR	82%	60%	Yes	Yes	No	Does not provide quantification, needs extensive initial processing and multiple probes	[18]
4	Integrated microfluidics-biosensor systems	91% to 95%	95% to 99%	Yes	Yes	Yes	Confounded results by urine variability and low bacterial count	[18]
5	Gram staining	85.1%	98.9%	Yes	No	No		[20]
6	Dipstick with nitrite & leucocyte esterase	53.1%	100%	Yes	No	No		[20]
7	Pus cell count	42.5%	95.5%	Yes	No	No		[20]
8	Urine analysis and microscopy	46.4%	89%	Yes	No	No		[21]
9	Chlorhexidine	100%	54%	Yes	No	No		[22]
10	Interleukin-8	70%	67%	Yes	No	No		[23]
11	Griess test	63.3%	99.5%	Yes	No	No		[24]
12	Serum procalcitonin level	30%	100%	Yes	No	No		[26]
13	Uriscreen Test	100%	68.6%	Yes	No	No		[27]
14	Antibody-based Lateral flow immunoassay	86%	94%	Yes	No	No		[28]
15	Chromogenic amoebocyte lysate assay	88.7%	98.7%	Yes	No	No		[29]
16	Flow cytometry-based systems	99%	58%	Yes	No	No		[30,31]
17	Genetic signature identification CAPTURE assay	100%	90%	Yes	Yes	No	Expensive, resource intensive, infrastructure, highly skilled manpower	[35]
18	Time-lapse microscopy	96% (agreement)	96% (agreement)	Yes	Yes	Yes	Imprecise phenotypic identification measures in direct urine, not easy-to-use	[36,37]
19	Kirby Bauer Method	51%	99%	No	No	Yes	Time-consuming, resource intense, and not user-friendly	[41,42]

Table 5. *Cont.*

Sl. No.	Method	Sensitivity	Specificity	Detection of Bacteriuria	Identification of Bacteria	Antibiotic Susceptibility	Limitations	Reference
20	Semiquantitative culture method	72.7%	95.7%	Yes	Yes	Yes	Time-consuming, resource intense, and not user-friendly	[44]
21	Quantitative culture method	59.3%	94.4%	Yes	Yes	Yes	Time-consuming, resource intense, and not user-friendly	[44]
22	Strip-based Urinalysis DongJiu	31.1%	91.8%	Yes	No	No		[46]
23	Automated Urinalysis system (URISED)	47%	91.1%	Yes	No	No		[47]
24	Detection of bacteriuria by a non-culture method	90% to 95%	90% to 95%	Not Applicable	Not Applicable	Not Applicable	European Guidelines for Urinalysis	[48]
25	Detection of bacteriuria by a rapid non-culture method	80% to 90%	80% to 90%	Not Applicable	Not Applicable	Not Applicable	European Guidelines for Urinalysis	[48]
26	Index Test	91.6 to 93.8%	80.7 to 96.7%	Yes	Yes	Yes	User-friendly, portable, affordable, rapid-fastest available, no special training required	Data from current study described above

The incidence of UTI from this hospital-based study may not be generalized over the entire population as the present study was conducted in tertiary hospitals to enroll enough participants in the shortest possible time and simultaneous comparison with reference gold standard test without any loss of time in processing or transport of collected samples. Due to the conventional practice in microbiology labs on the choice of antibiotics, which may also be governed by the availability of antibiotics discs, antibiotics could be compared for sensitivity in a subgroup of samples tested. In spite of this, the novel test reported resistance to antibiotics with an accuracy of 80% to 100%, except for two antibiotics, cefazolin (71%) and levofloxacin (65%). This was primarily seen in samples with more than one bacterial entity with a high probability of quorum sensing, impacting the sensitivity to the given antibiotic/s.

Rapid assay described herein determines the efficacy of an antibiotic not only in the shortest possible time, but also with literally no dependence on trained manpower and lab infrastructure. Although this novel test is suitable for use in all healthcare settings, it can prove to be of immense and unsurpassable value for healthcare facilities in low-resource settings due to features like portability, point-of-use testing, and no additional requirements. A consensus using Delphi technique, obtained from experts regarding criteria required for an acceptable point-of-care test for UTI detection, was reported by Weir et al. [49]. This novel test fulfils 25 out of 26 accepted criteria, except for just one, i.e., use of small sample volume. The novel test also fulfils six out of seven of the WHO's ASSURED criteria for ideal characteristics for a point-of-care test in resource-limited settings, the sole unfulfilled one being equipment free, and also matches all the revised criteria suggested by Paul et al. [50]. Effective diagnosis is a prerequisite for successful therapy, and early and accurate diagnosis results in timely and appropriate treatment.

5. Conclusions

In conclusion, it can be said that this novel test, with high sensitivity and specificity for detecting bacterial UTI and reporting antibiogram, can be used as a triage test for diagnosing UTIs and suggesting appropriate treatment for an evidence-based prescription for antibiotics in any kind of healthcare settings. In the wake of growing AMR, the prevalent "lack of priority for diagnostics over treatment" needs to be addressed urgently and this novel test enables physicians and labs to achieve this by adopting this affordable and portable IVD test before prescribing antibiotics for treatment of infectious diseases.

Author Contributions: Conceptualization, S.K., and S.S.; methodology, M.G.; software, S.K.; validation, M.G., S.K., N.K., P.B.,V.N., and S.M.D.; formal analysis, M.G.; investigation, M.G. and S.M.D.; resources, S.K., N.K., and V.N.; data curation, P.B. and M.G.; writing—original draft preparation, M.G., S.K., and S.S.; writing—review and editing, M.G., S.K., and S.S.; visualization, S.K. and S.S.; supervision, S.K., S.S., N.K., P.B., V.N., and S.M.D.; project administration, M.G., V.N., S.M.D., and S.K.; funding acquisition, S.K.

Funding: This technology platform was developed with funding received as a research grant from Defence Research and Development Organization, India under their NPMASS program and the clinical validation was undertaken from grants received from DBT under their SPARSH and GYTI-SRISTI schemes. The study and researchers were independent of the funding agency and the funding agencies had no role in design of the study, collection of data, collection of samples, processing of the novel/index test, analysis, and interpretation of data, writing of paper or submission for publication. The authors had full control of all primary data.

Acknowledgments: We gratefully acknowledge the support provided by the Government of Telangana supported Microbiology Department of Gandhi Medical College and All India Institute of Medical Sciences, Jodhpur for conducting the study and providing space and facilities for hosting the index test in their premises. We also acknowledge the support of Mr. M. Rajesh Bhatt, the lab technician for coordinating sample collection at Gandhi Hospital. Technical help from the technicians in Genomics Laboratory of BITS Pilani, Hyderabad Campus is also gratefully acknowledged.

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. World Health Organization. *The World Health Report 1996: Fighting DISEASE, Fostering Development*; Report of the Director-General; World Health Organization: Geneva, Switzerland, 1996.
2. Lal, P. *Report of the National Commission on Macroeconomics and Health*; Report of the National Commission on Health; Ministry of Health and Family Welfare: New Delhi, India, 2005.
3. Dixit, A.; Kumar, N.; Kumar, S.; Trigun, V. Antimicrobial Resistance: Progress in the Decade since Emergence of New Delhi Metallo-β-Lactamase in India. *Indian J. Community Med.* **2019**, *44*, 4–8. [PubMed]
4. Ganguly, N.K.; Arora, N.K.; Chandy, S.J.; Fairoze, M.N.; Gill, J.P.S.; Gupta, U.; Hossain, S.; Joglekar, S.; Joshi, P.C.; Kakkar, M.; et al. Rationalizing antibiotic use to limit antibiotic resistance in India. *Indian J. Med. Res.* **2011**, *134*, 281–294. [PubMed]
5. Baquero, F. Low-level antibacterial resistance: A gateway to clinical resistance. *Drug Resist. Updat.* **2001**, *4*, 93–105. [CrossRef] [PubMed]
6. Eigner, U.; Schmid, A.; Wild, U.; Bertsch, D.; Fahr, A.-M. Analysis of the comparative workflow and performance characteristics of the VITEK 2 and Phoenix systems. *J. Clin. Microbiol.* **2005**, *43*, 3829–3834. [CrossRef] [PubMed]
7. Thomson, K.S.; Cornish, N.E.; Hong, S.G.; Hemrick, K.; Herdt, C.; Moland, E.S. Comparison of Phoenix and VITEK 2 extended-spectrum-beta-lactamase detection tests for analysis of Escherichia coli and Klebsiella isolates with well-characterized beta-lactamases. *J. Clin. Microbiol.* **2007**, *45*, 2380–2384. [CrossRef]
8. Dolk, F.C.K.; Pouwels, K.B.; Smith, D.R.M.; Robotham, J.V.; Smieszek, T. Antibiotics in primary care in England: Which antibiotics are prescribed and for which conditions? *J. Antimicrob. Chemother.* **2018**, *73*, ii2–ii10. [CrossRef]
9. Chua, K.-P.; Fischer, M.A.; Linder, J.A. Appropriateness of outpatient antibiotic prescribing among privately insured US patients: ICD-10-CM based cross sectional study. *BMJ* **2019**, *364*, k5092. [CrossRef]
10. Schmiemann, G.; Kniehl, E.; Gebhardt, K.; Matejczyk, M.M.; Hummers-Pradier, E. The diagnosis of urinary tract infection: A systematic review. *Deutsches Ärzteblatt Int.* **2010**, *107*, 361–367.
11. Kim, H.Y.; Lee, S.-J.; Lee, D.S.; Yoo, J.M.; Choe, H.-S. Microbiological Characteristics of Unresolved Acute Uncomplicated Cystitis. *Microb. Drug Resist.* **2016**, *22*, 387–391. [CrossRef]
12. Seifu, W.D.; Gebissa, A.D. Prevalence and antibiotic susceptibility of Uropathogens from cases of urinary tract infections (UTI) in Shashemene referral hospital, Ethiopia. *BMC Infect. Dis.* **2018**, *18*, 30. [CrossRef]
13. Chang, P.-C.; Hsu, Y.-C.; Hsieh, M.-L.; Huang, S.-T.; Huang, H.-C.; Chen, Y. A pilot study on Trichomonas vaginalis in women with recurrent urinary tract infections. *Biomed. J.* **2016**, *39*, 289–294. [CrossRef] [PubMed]
14. Frej-Mądrzak, M.; Teryks-Wołyniec, D.; Jama-Kmiecik, A.; Sarowska, J.; Choroszy-Król, I. Diagnosing Chlamydia Trachomatis Urinary Tract Infections—Preliminary Report. *Adv. Clin. Exp. Med.* **2015**, *24*, 441–445. [CrossRef] [PubMed]
15. Lee, J.B.L.; Neild, G.H. Urinary tract infection. *Medicine* **2007**, *35*, 423–428. [CrossRef]
16. Mathai, E.; Thomas, R.J.; Chandy, S.; Mathai, M.; Bergstrom, S. Antimicrobials for the treatment of urinary tract infection in pregnancy: Practices in southern India. *Pharmacoepidemiol. Drug Saf.* **2004**, *13*, 645–652. [CrossRef] [PubMed]
17. U.S. Preventive Services Task Force. Screening for asymptomatic bacteriuria in adults: Reaffirmation recommendation statement. *Am. Fam. Phys.* **2010**, *81*, 505.
18. Davenport, M.; Mach, K.E.; Shortliffe, L.M.D.; Banaei, N.; Wang, T.-H.; Liao, J.C. New and developing diagnostic technologies for urinary tract infections. *Nat. Rev. Urol.* **2017**, *14*, 296–310. [CrossRef]
19. Gehani, M.; Kapur, S.; Bhardwaj, P.; Nag, V.; Balasubramaniam, S.; Kammili, N.; Madhuri, S. Unmet need of antenatal screening for asymptomatic bacteriuria: A risk factor for adverse outcomes of pregnancy. *Indian J. Community Med.* **2019**, *44*, 193. [CrossRef]
20. Jayalakshmi, J.; Jayaram, V.S. Evaluation of various screening tests to detect asymptomatic bacteriuria in pregnant women. *Indian J. Pathol. Microbiol.* **2008**, *51*, 379–381. [CrossRef]
21. Van Nostrand, J.D.; Junkins, A.D.; Bartholdi, R.K. Poor Predictive Ability of Urinalysis and Microscopic Examination to Detect Urinary Tract Infection. *Am. J. Clin. Pathol.* **2000**, *113*, 709–713. [CrossRef]
22. Okusanya, B.O.; Aigere, E.O.S.; Eigbefoh, J.O.; Okome, G.B.O.; Gigi, C.E. Is a chlorhexidine reaction test better than dipsticks to detect asymptomatic bacteriuria in pregnancy? *J. Obstet. Gynaecol.* **2014**, *34*, 21–24. [CrossRef]

23. Shelton, S.D.; Boggess, K.A.; Kirvan, K.; Sedor, F.; Herbert, W.N. Urinary interleukin-8 with asymptomatic bacteriuria in pregnancy. *Obstet. Gynecol.* **2001**, *97*, 583–586. [PubMed]
24. Mathews, J.E.; George, S.; Mathews, P.; Mathai, E.; Brahmadathan, K.N.; Seshadri, L. The Griess test: An inexpensive screening test for asymptomatic bacteriuria in pregnancy. *Aust. N. Z. J. Obstet. Gynaecol.* **1998**, *38*, 407–410. [CrossRef] [PubMed]
25. Archbald, F.J.; Verma, U.; Tejani, N.A. Screening for asymptomatic bacteriuria with Microstix. *J. Reprod. Med.* **1984**, *29*, 272–274. [PubMed]
26. Bilir, F.; Akdemir, N.; Ozden, S.; Cevrioglu, A.S.; Bilir, C. Increased serum procalcitonin levels in pregnant patients with asymptomatic bacteriuria. *Ann. Clin. Microbiol. Antimicrob.* **2013**, *12*, 25. [CrossRef] [PubMed]
27. Waisman, Y.; Zerem, E.; Amir, L.; Mimouni, M. The Validity of the Uriscreen Test for Early Detection of Urinary Tract Infection in Children. *Pediatrics* **1999**, *104*, e41. [CrossRef]
28. Stapleton, A.E.; Cox, M.E.; DiNello, R.K.; Geisberg, M.; Abbott, A.; Roberts, P.L.; Hooton, T.M. Performance of a New Rapid Immunoassay Test Kit for Point-of-Care Diagnosis of Significant Bacteriuria. *J. Clin. Microbiol.* **2015**, *53*, 2805–2809. [CrossRef]
29. Nachum, R.; Arce, J.J.; Berzofsky, R.N. Gram-negative bacteriuria of pregnancy: Rapid detection by a chromogenic Limulus amoebocyte lysate assay. *Obstet. Gynecol.* **1986**, *68*, 215–219.
30. Moshaver, B.; de Boer, F.; van Egmond-Kreileman, H.; Kramer, E.; Stegeman, C.; Groeneveld, P. Fast and accurate prediction of positive and negative urine cultures by flow cytometry. *BMC Infect. Dis.* **2016**, *16*, 211. [CrossRef]
31. Duong, H.P.; Wissing, K.M.; Tram, N.; Mascart, G.; Lepage, P.; Ismaili, K. Accuracy of Automated Flow Cytometry-Based Leukocyte Counts to Rule Out Urinary Tract Infection in Febrile Children: A Prospective Cross-Sectional Study. *J. Clin. Microbiol.* **2016**, *54*, 2975–2981. [CrossRef]
32. Hale, D.C.; Wright, D.N.; McKie, J.E.; Isenberg, H.D.; Jenkins, R.D.; Matsen, J.M. Rapid screening for bacteriuria by light scatter photometry (Autobac): A collaborative study. *J. Clin. Microbiol.* **1981**, *13*, 147–150.
33. Jenkins, R.D.; Hale, D.C.; Matsen, J.M. Rapid semiautomated screening and processing of urine specimens. *J. Clin. Microbiol.* **1980**, *11*, 220–225. [PubMed]
34. Tran, A.; Alby, K.; Kerr, A.; Jones, M.; Gilligan, P.H. Cost Savings Realized by Implementation of Routine Microbiological Identification by Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry. *J. Clin. Microbiol.* **2015**, *53*, 2473–2479. [CrossRef] [PubMed]
35. O'Keefe, L.C.; Koelle, P.; McGee, Z.; Dewberry, L.S.; Wright, C.; Stallings, J.E.; Gates, E.; Chittur, K. Innovations in Worksite Diagnosis of Urinary Tract Infections and the Occupational Health Nurse. *Workplace Health Saf.* **2019**, *67*, 268–274. [CrossRef] [PubMed]
36. Fredborg, M.; Rosenvinge, F.S.; Spillum, E.; Kroghsbo, S.; Wang, M.; Sondergaard, T.E. Rapid antimicrobial susceptibility testing of clinical isolates by digital time-lapse microscopy. *Eur. J. Clin. Microbiol. Infect. Dis.* **2015**, *34*, 2385–2394. [CrossRef] [PubMed]
37. Price, C.S.; Kon, S.E.; Metzger, S. Rapid antibiotic susceptibility phenotypic characterization of Staphylococcus aureus using automated microscopy of small numbers of cells. *J. Microbiol. Methods* **2014**, *98*, 50–58. [CrossRef] [PubMed]
38. Bossuyt, P.M.; Irwig, L.; Craig, J.; Glasziou, P. Comparative accuracy: Assessing new tests against existing diagnostic pathways. *BMJ* **2006**, *332*, 1089–1092. [CrossRef]
39. Kapur, S.; Gupta, S. Indigenous rapid diagnostic technology for antibiotic susceptibility testing in urinary tract infection: From bench side to bedside. *BMJ Innov.* **2017**, *3*, 7–11. [CrossRef]
40. Kapur, S.; Gupta, S.; Dv, P.; Pal, A.; Pant, J.; Jain, R. Rapid Sensor Based Technology: A Novel Tool for Direct Antimicrobial Susceptibility Testing in Urinary Tract Infection. *Transl. Med. Biotecnol.* **2014**, *2*, 22–28.
41. Clinical and Laboratory Standards Institute (CLSI). *Performance Standards for Antimicrobial Susceptibility Testing*, 26th ed.; CLSI Supplement M100S; Clinical and Laboratory Standards Institute: Wayne, PA, USA, 2016; ISBN1 56238-923-8, ISBN2 56238-924-6.
42. Infectious Diseases Society of America; American Society of Nephrology; American Geriatric Society; Nicolle, L.E.; Bradley, S.; Colgan, R.; Rice, J.C.; Schaeffer, A.; Hooton, T.M. Infectious Diseases Society of America Guidelines for the Diagnosis and Treatment of Asymptomatic Bacteriuria in Adults. *Clin. Infect. Dis.* **2005**, *40*, 643–654. [CrossRef]
43. Gehani, M.; Kapur, S. Rapid Diagnostic Test for Antimicrobial Susceptibility. *Harv. Dataverse* **2019**, *V1*. [CrossRef]

44. Riboli, D.F.M.; Lyra, J.C.; Silva, E.P.; Valadão, L.L.; Bentlin, M.R.; Corrente, J.E.; de Souza Rugolo, L.M.S.; de Souza, M.D.L.R. Diagnostic accuracy of semi-quantitative and quantitative culture techniques for the diagnosis of catheter-related infections in newborns and molecular typing of isolated microorganisms. *BMC Infect. Dis.* **2014**, *14*, 283. [CrossRef] [PubMed]
45. Salmanzadeh-Ahrabi, S.; Kocagöz, T. Evaluation of a liquid-phase colorimetric method for rapid antibacterial susceptibility testing. *Iran. J. Microbiol.* **2013**, *5*, 259–262. [PubMed]
46. Kocer, D.; Sariguzel, F.M.; Ciraci, M.Z.; Karakukcu, C.; Oz, L. Diagnostic Accuracy of a New Urinalysis System, DongJiu, for Diagnosis of Urinary Tract Infection. *Ann. Clin. Lab. Sci.* **2015**, *45*, 686–691. [PubMed]
47. Huysal, K.; Budak, Y.U.; Ulusoy Karaca, A.; Aydos, M.; Kahvecioğlu, S.; Bulut, M.; Polat, M. Diagnostic accuracy of urised automated urine microscopic sediment analyzer and dipstick parameters in predicting urine culture test results. *Biochem. Medica* **2012**, *23*, 211–217. [CrossRef]
48. Kouri, T.; Fogazzi, G.; Gant, V.; Hallander, H.; Hofmann, W.; Guder, W.G. European Urinalysis Guidelines. *Scand. J. Clin. Lab. Investig. Suppl.* **2000**, *60*, 1–96. [CrossRef]
49. Weir, N.-J.M.; Pattison, S.H.; Kearney, P.; Stafford, B.; Gormley, G.J.; Crockard, M.A.; Gilpin, D.F.; Tunney, M.M.; Hughes, C.M. Criteria required for an acceptable point-of-care test for UTI detection: Obtaining consensus using the Delphi technique. *PLoS ONE* **2018**, *13*, e0198595. [CrossRef]
50. Drain, P.K.; Hyle, E.P.; Noubary, F.; Freedberg, K.A.; Wilson, D.; Bishai, W.R.; Rodriguez, W.; Bassett, I.V. Diagnostic point-of-care tests in resource-limited settings. *Lancet Infect. Dis.* **2014**, *14*, 239–249. [CrossRef]

Journal of
Clinical Medicine

MDPI

Article

Characteristics of Mild Cognitive Impairment in Northern Japanese Community-Dwellers from the ORANGE Registry

Yu Kume [1], Tomoko Takahashi [2], Yuki Itakura [3], Sangyoon Lee [4], Hyuma Makizako [5], Tsuyosi Ono [6], Hiroyuki Shimada [4] and Hidetaka Ota [3,*]

1 Department of Occupational Therapy, Graduate School of Medicine, Akita University, Akita 010-8543, Japan; kume.yuu@hs.akita-u.ac.jp
2 Integrated Community Support Center, Public Health and Welfare Department, City Hall of Yokote, Akita 013-0525, Japan; takahashi-tomoko-d@city.yokote.lg.jp
3 Advanced Research Center for Geriatric and Gerontology, Akita University, Akita 010-8543, Japan; itacie15@gmail.com
4 Center for Gerontology and Social Science, National Center for Geriatrics and Gerontology, Obu, Aichi 474-8511, Japan; sylee@ncgg.go.jp (S.L.); shimada@ncgg.go.jp (H.S.)
5 Department of Physical Therapy, School of Health Sciences, Faculty of Medicine, Kagoshima University, Kagoshima 890-8544, Japan; makizako@health.nop.kagoshima-u.ac.jp
6 Omori Municipal Hospital, Akita 013-0525, Japan; onot@oomorihp.jp
* Correspondence: hidetaka-ota@med.akita-u.ac.jp; Tel.: +81-18-801-7061; Fax: +81-18-801-7062

Received: 12 September 2019; Accepted: 8 November 2019; Published: 10 November 2019

Abstract: A gradually increasing prevalence of mild cognitive impairment (MCI) is recognized in the super-aging society that Japan faces, and early detection and intervention in community-dwellers with MCI are critical issues to prevent dementia. Although many previous studies have revealed MCI/non-MCI differences in older individuals, information on the prevalence and characteristics of MCI in rural older adults is limited. The aim of this study was to investigate differential characteristics between older adults with and without MCI. The investigation was conducted over one year from 2018 to 2019. Participants were recruited from Akita in northern Japan. Neuropsychological assessments were applied to classify MCI, including the National Center for Geriatrics and Gerontology Functional Assessment Tool (NCGG-FAT) and the Touch panel-type Dementia Assessment Scale (TDAS) based on the Alzheimer's disease assessment scale. Our samples consisted of 103 older adults divided into 54 non-MCI and 49 MCI. The MCI group had lower scores of all cognitive items. Our results showed that individuals with MCI had significantly slower walking speed (WS) and worse geriatric depression scale (GDS) compared to non-MCI. In addition, WS was significantly associated with some cognitive items in non-MCI, but not in MCI. Finally, we showed that predictive variables of MCI were WS and GDS. Our study provides important information about MCI in rural community-dwellers. We suggest that older adults living in a super-aging society should receive lower limb training, and avoiding depression in older adults through interaction of community-dwellers may contribute to preventing the onset of MCI.

Keywords: older adults living in super-aging society; mild cognitive impairment; walking speed; depression

1. Introduction

Mild cognitive impairment (MCI) is a transitional state of cognition between normal ageing and dementia that may progress to dementia. MCI is defined by subjective or objective evidence of cognitive decline greater than expected for the individual's age and education level but that does not interfere

notably with activities of daily life, and the early detection and prevention of MCI are a challenge to prevent dementia in older adults [1]. Within established processes for making a diagnosis of MCI [2], some factors in its early detection remain unclear, as well as predictors of reversion from MCI to normal cognition. Differences between individuals without MCI and those with MCI have been studied and reported, which shows that many factors such as a lack of exercise [3], cerebrovascular factors [4], and anxiety [5] affect cognitive function. Especially, it is appreciated that cognitive and physical impairments in older adults are related through shared pathophysiological mechanisms [6]. Some studies show that older adults with MCI compared to individuals without MCI perform more poorly not just on neurocognitive performance, but also on complex motor and psychomotor domains [7–9], and exhibit greater gait impairment [10–14]. Recently, it has become clear that MCI and physical frailty are related. The physical phenotype of frailty is represented by low levels of lean body mass, muscle strength, gait performance, physical activity, and exhaustion [15]. Gait performance of the frailty is associated with cognitive decline and MCI conversion to AD as reported by [9,16]. Therefore, investigating close associations between MCI and physical function has important implications for improving diagnostic acuity of MCI and targeting interventions to prevent dementia and disability among older adults.

To clarify the dementia risk associated with MCI or early stage dementia, a nationwide clinical registry called the Organized Registration for the Assessment of dementia on Nationwide General consortium toward Effective treatment (ORANGE) is ongoing in Japan [17]. The recruitment of many registrants has been in progress in several regions of Japan from 2017, and we performed an extending preclinical trial in a cohort in northern Japan up to 2019. As is well known, gradual growth of the older population has been experienced in Japan. Especially, northern rural areas in Japan (Akita prefecture) are the most super-aging society in the world (e.g., the number of individuals over aged 75 in Akita is estimated to reach 205,000 people by 2025 [18]). Although there are few epidemiological data regarding MCI in rural areas of Japan, several studies have reported MCI profiles in older adults [19,20]. Most of them are focused on the prevalence of MCI or the conversion rate to dementia, and the detailed cognitive profile (e.g., attention, executive function, information coding skill, etc.) of MCI is not covered, as well as scarce epidemiological data regarding health-related variables such as physical performance and mental status. Therefore, we analyzed the data of a prospective cohort in northern Japan. In this study, we investigated which factors were related to MCI status according to the National Center for Geriatrics and Gerontology Functional Assessment Tool (NCGG-FAT) [21,22] and the Touch panel-type Dementia Assessment Scale (TDAS) [23,24] based on the Alzheimer's disease assessment scale (ADAS) [25]. To clarify the characteristics of rural older adults with MCI, we focused on three points as follows. First, we mainly compared cognitive function, physical performance, and depressive symptoms in MCI individuals with those in non-MCI individuals. Second, we examined correlations between physical performance and cognitive and mental function in each group (i.e., non-MCI group and MCI group). Finally, a binomial logistic regression model was estimated to determine predictive factors for MCI in rural older adults in Japan.

2. Experimental Section

2.1. Participants and Study Design

The participants were recruited in a rural area in Akita with a small population (total 32,440) with a super-aged rate of 38.7% according to public information, from 2018 to 2019. The inclusion criteria were age 65 years and over, having walking ability without personal assistance, and living at home. The exclusion criteria were dementia, major depression, severe hearing or visual impairment, stroke, Parkinson's disease, other neurological disease, intellectual disability, need for support or care as certified by the Japanese public long-term care insurance system due to disability, and inability to complete cognitive tests at the baseline assessment. The study was approved by the ethics committee of the Faculty of Medicine, Akita University (approval No. 1649) and was performed in accordance

with the Declaration of Helsinki II. Informed consent was obtained from all participants. According to sample size calculations using G*Power for unpaired *t* test [26], we estimated a sample size of 64 participants per group to detect a clinically significant effect with $\alpha = 0.05$, power = 80%, and effect size = 0.50.

2.2. Assessment and Outcome

After obtaining informed consent from each participant, demographics (age, gender, and education) and health variables (body mass index (BMI), medical history of hypertension and diabetes, frail phenotype, medication and Geriatric Depression Scale-15 (GDS)) were collected according to the ORANGE protocol. A questionnaire sent in advance by mail was self-described by each participant, including age, gender, educational duration, presence of hypertension and diabetes (e.g., yes or no), amount of medications, and GDS (e.g., score range from 0 to 15, as indexed more depressive symptoms in higher scores). Height and weight to calculate BMI were measured by public health nurses. Five components of the National Center for Geriatrics and Gerontology-Study of Geriatric Syndromes (NCGG-SGS) [27] based on the Fried frailty index [15] were applied to assess frailty: (i) self-reported unintentional weight loss (i.e., a decrease of 2–3 kg over six months [28]), (ii) self-reported exhaustion (i.e., presence of fatigue for two weeks [28]), (iii) self-reported low physical activity (i.e., no exercise habit for a week [29]), (iv) weakness (i.e., grip strength (GS) less than 26/18 kg for male/female [30]): GS was measured using a Smedley-type handheld dynamometer (GRIP-D; Takei Ltd., Niigata, Japan), and (v) slow walking speed (WS) (i.e., less than 1.0 m/s in 5 m walking test [29]): walking time was measured over a 2.4-m distance in seconds using infrared sensors and participants' WS (m/s) was calculated. They were used to define robust (score of zero), pre-frail (score of 1 to 2), and frail (score of 3 to 5). The frail index of NCGG-SGS is almost equal to the original index of Fried's study [15] except the modified cut-off values for slowness and weakness are appropriate criteria for physical frailty assessments in the Japanese older population [31,32]. The present study also applied NCGG-FAT and TDAS based on ADAS to assess cognitive function in the participants and to divide the participants into non-MCI and MCI groups. All the variables of five frail components, the NCGG-FAT and TDAS were evaluated by trained public health nurses throughout a comprehensive health checkup in a local spot.

2.3. Components of NCGG-FAT

The computerized multidimensional neurocognitive test was performed on an iPad (Apple, Cupertino, CA, USA) with a 9.7-inch touch display. The task instructions were presented with a letter size of at least 1.0×1.0 cm^2 on the display. For this study, a trained operator supported each participant by setting up the tablet PC and running each test. Participants completed the NCGG-FAT subtests as follows.

2.3.1. Tablet Version of Word Recognition (WR)

This test is comprised of two computerized tasks of immediate recognition and delayed recall. In the first task of immediate recognition, participants were instructed to memorize 10 words, each of which was displayed for 2 s on the tablet PC. After that, a total of 30 words including 10 target and 20 distracter words were shown to participants, and they were required to select the 10 target words immediately. This task was repeated for three trials. The average number of correct answers was recorded as a score ranging from 0 to 10. In another task, participants were asked to correctly recall the 10 target words after 20 min. The number of correctly recalled target words was scored ranging from 0 to 10. Finally, we calculated the sum score of the two tasks of immediate recognition and delayed recall.

2.3.2. Tablet Version of Trail Making Test Version A (TMT-A) and Version B (TMT-B)

In the Trail Making Test Version A (TMT-A) task, participants were instructed to touch the target numbers in a sequence as rapidly as possible. Target numbers from 1 to 15 were randomly displayed on the tablet panel. In addition, the Trail Making Test Version B (TMT-B) instructions required participants

to touch target numbers (e.g., 1–15) and letters in turn. The required time (seconds) to complete each task was recorded, within a maximum time of 90 s.

2.3.3. Tablet Version of Symbol Digit Substitution Task (SDST)

In the Symbol Digit Substitution Task (SDST), nine pairs of numbers and symbols were shown in the upper part of the tablet display. A target symbol was shown in the center of the tablet panel, and selectable numbers were displayed at the bottom. Participants were asked to touch the number corresponding to the target symbol shown in the central part of the tablet display as rapidly as possible. The number of correct numbers within 90 s was recorded.

2.4. Components of TDAS

The TDAS test was presented on a 14-inch touch panel display. The TDAS subtests consisted of seven of the ADAS-cog test items (11 test items) and two other tasks. Participants were instructed verbally or visually by the computer to complete the TDAS subtests as follows.

2.4.1. WR

The WR test was a computerized test based on the WR task of ADAS-cog. At the start of instructions for this task, 12 target words were individually presented on the display for 3 s each at 2 s intervals. After demonstrating the target words, the computer randomly displayed 24 words consisting of 12 target words and 12 non-target words. Participants were then instructed to respond by touching the displayed button of 'yes', 'no', or 'unknown' in response to the question regarding whether the word had been shown previously. Participants completed the trial three times. The total number of incorrect responses for three trials was recorded, with a maximum score of 72.

2.4.2. Following a Command

This task was modified from the command task of ADAS-cog. The computer presented 10 selectable icons labelled from 0 to 9 and then required participants to touch the number specified. The number of incorrect responses in two trials was scored with a maximum score of 2.

2.4.3. Orientation

This task was based on the orientation task of ADAS-cog. The computer displayed four screens in sequence. On each screen, participants were asked to touch selectable icons and answer what year, month, day, and weekday it is. The number of incorrect responses was scored with a maximum score of 4.

2.4.4. Visual-Spatial Perception

This task was modified from the constructional praxis task of ADAS-cog to evaluate visual-spatial perception. ADAS-cog requires subjects to copy the geometric forms presented. The computer first presented four screens displaying a target geometric form (i.e., a square, rhombus, cube, or triangular prism) for 5 s each. Participants were then required to correctly select the target form in response to a question task including the target form and four non-target forms. The number of incorrect responses was scored with a maximum score of 4.

2.4.5. Naming Fingers

This test assessed whether participants can name the fingers correctly, using the protocol of ADAS-cog. Participants were asked to correctly respond to a picture question of a hand marked with a red circle, by touching an icon labelled with the five finger names. An incorrect response was scored as one point, with a maximum score of 5.

2.4.6. Object Recognition

This task was based on the naming objects task of ADAS-cog. Participants were instructed to touch the correct usage icon (e.g., a pair of scissors, comb or broom) of five selectable icons labelled with the purpose of usage. Three trials were completed, and an incorrect response was scored as one point (maximum score = 3).

2.4.7. Accuracy of Order of a Process

This task was modified from the ideational praxis of ADAS-cog. The computer displayed seven icons labelled randomly with seven actions. Participants were asked to correctly touch the icons in order. The number of incorrect responses was recorded, with a maximum score of 5.

2.4.8. Money Calculation

This task assessed the money calculation ability of each participant. Participants needed to combine coins equal to an amount of money from various denominations of coins displayed on the screen. Three trials were completed, and an incorrect response was scored as one point (maximum score = 3).

2.4.9. Clock Time Recognition

This task included three kinds of question regarding clock time recognition. Participants were instructed to correctly state the time shown on a clock displayed on the screen. The number of incorrect responses was recorded, with a maximum score of 3.

2.5. MCI Classification by NCGG-FAT and TDAS

According to Petersen's report [2] in which individuals who showed cognitive impairment but were independent in activities of daily living were defined as having MCI, we applied MCI classification according to the cutoff point of NCGG-FAT or TDAS. For all cognitive subtests of NCGG-FAT, the standardized threshold in each corresponding domain for defining impairment in Japanese population-based cohorts consisting of older community-dwellers is a score more than 1.5 standard deviations (SD) below the age- and education-specific mean [21]. In TDAS, decreasing scores indicate cognitive improvement (range of scores from 0 to 101), and total scores ranging from 7 to 13 were classified as MCI [23].

3. Analyses

According to results of the normalization test (Kolmogorov–Smirnov test), Age, Height, Weight, and BMI were used by the unpaired *t* test. Gender (% female), Hypertension (% Yes), Diabetes (% Yes), Weight loss (% Yes), Poor energy (% Yes), and Low physical activity level (% Yes) were analyzed by chi-squared test for 2×2 contingency, except for Pearson's chi-square test for Frail phenotype (%, robust/pre-frail/frail) for 2×3 contingency. Mann–Whitney test was applied for GS (kg), WS (m/s), Amount of medications (*n*), Education (years), GDS-15 (score), and cognitive measurements of NCGG-FAT and TDAS (Table 1).

Table 1. Characteristics of participants with and without mild cognitive impairment (MCI).

Variables	Non-MCI Group		MCI Group		*p* Value	
	n = 54		*n* = 49			
	Mean	SD	Mean	SD		
Age (years)	74.1	6.1	74.4	5.7	0.84	
Gender (% female)	53.7%		55.1%		0.89	
Height (cm)	155.0	8.2	156.0	8.5	0.53	
Weight (kg)	57.9	11.4	60.0	10.1	0.34	
Body Mass Index (kg/m^2)	24.0	3.7	24.6	3.5	0.41	
	%		%		***p* Value**	
Hypertension (% Yes)	63.0%		61.2%		0.86	
Diabetes (% Yes)	20.4%		26.5%		0.46	
Frail five components						
Frail phenotype (%, robust/pre-frail/frail)	50%/50%/0%		43%/47%/10%		0.054	
(i) Weight loss (% Yes)	11.1%		14.3%		0.63	
(ii) Poor energy (% Yes)	16.7%		26.5%		0.22	
(iii) Low physical activity level (% Yes)	13.0%		18.4%		0.45	
	Median	**IQR**	**Median**	**IQR**	***p* Value**	
(iv) Grip strength (kg)	25.2	12.0	22.9	9.0	0.25	
(v) Walking speed (m/s)	1.3	0.4	1.2	0.3	0.03	*
Amount of medications (*n*)	3.0	3.0	4.0	4.0	0.17	
Education (years)	12.0	3.0	12.0	3.0	0.18	
GDS-15 total score (score)	2.0	2.0	3.0	4.0	0.046	*
NCGG-FAT						
Word recognition (score)	11.7	3.8	8.0	4.2	0.000	***
Tablet version of TMT-A (s)	19.0	6.0	27.0	11.0	0.000	***
Tablet version of TMT-B (s)	33.5	18.0	46.0	44.5	0.000	***
Tablet version of SDST (score)	42.0	12.0	33.0	12.5	0.000	***
TDAS						
TDAS total score (score)	2.0	3.0	7.0	9.0	0.000	***

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, Mann–Whitney test was applied for Education (years), Amount of medications (*n*), GDS-15 total score (score), Grip strength (kg), Walking speed (m/s), and cognitive measurements of NCGG-FAT and TDAS. Age, height, weight, and BMI were analyzed by unpaired *t* test, and gender (% female), hypertension (% Yes), diabetes (% Yes), weight loss (% Yes), poor energy (% Yes), and low physical activity level (% Yes) were analyzed by chi-squared test, except for Pearson's chi-square test for frail phenotype (%, robust/pre-frail/frail). SD, standard deviation; IQR, interquartile range; Loss weight, Loss weight more than 3 kg in six months; TMT-A, Trail Making Test A version; TMT-B, Trail Making Test B version; SDST, Symbol Digit Substitution Task; TDAS, Touch Panel-type Dementia Assessment Scale; GDS-15, Geriatric Depression Scale.

As the variables of WS, GS, subtests of NCGG-FAT, TDAS, and GDS-15 total score were not statistically normalized from the Kolmogorov–Smirnov test, Spearman correlation analysis for interval scales was applied to analyze the relationship among Age, GS, WS, subtests of NCGG-FAT, TDAS, and GDS total score for each group (Table 2).

The values of $p_{in} = 0.2$ and $p_{out} = 0.25$ were set up to select independent variables from Tables 1 and 2 for input into a binominal logistic regression model. The regression model was performed by a method of likelihood ratio, and set up the MCI classification as the dependent variable and predictors (i.e., independent variables) according to the following regression models; (i) 11 predictors of Model I include Age, GS, WS, Amount of medications, Education, WR, TMT-A, B, and SDST of NCGG-FAT, TDAS, and GDS-15 total score. (ii) Ten predictors of Model II except for TDAS score include Age, GS, WS, Amount of medications, Education, WR, TMT-A, B, and SDST of NCGG-FAT, and GDS-15 total score. Finally, (iii) six predictors of Model III except for all cognitive variables included Age, GS, WS, Amount of medications, Education, and GDS-15 total score. The model adaptation was examined by Hosmer–Lemeshow test (Table 3). SPSS Version 26.0 for Windows (SPSS Inc., Chicago. IL, USA) was used for analysis, and the level of significance was set at $p = 0.05$.

Table 2. Correlations for each group (non-MCI and MCI).

Variables	Non-MCI Group (n = 54)							MCI Group (n = 49)						
	WS	**GS**	**WR**	**TMT-A**	**TMT-B**	**SDST**	**TDAS**	**WS**	**GS**	**WR**	**TMT-A**	**TMT-B**	**SDST**	**TDAS**
Age (years)	−0.37 **	−0.21	−0.51 **	0.55 **	0.66 **	−0.66 **	0.05	−0.37 *	−0.04	−0.33 *	0.43 **	0.49 **	−0.47 **	0.25
BMI (kg/m^2)	0.08	0.24	0.18	0.05	0.05	0.02	0.00	−0.19	0.08	0.19	0.01	−0.01	0.03	−0.04
Education (years)	0.19	−0.01	0.03	−0.36 **	−0.35 **	0.15	−0.13	−0.02	0.28	0.09	−0.13	−0.10	0.29	−0.45 **
Medications (n)	0.06	−0.22	−0.11	−0.02	0.22	0.01	−0.16	−0.24	0.25	0.04	−0.03	0.07	−0.07	0.06
WS (m/s)	1.00	0.26	0.42 **	−0.31 **	−0.35 **	0.44 **	−0.05	1.00	−0.08	−0.08	−0.18	−0.07	0.12	−0.06
GS (kg)	0.26	1.00	0.18	−0.09	−0.02	0.22	0.31 *	−0.08	1.00	0.06	0.16	−0.12	0.16	−0.11
GDS-15 (score)	0.16	0.01	0.03	0.00	−0.01	−0.04	0.00	0.03	−0.16	0.12	−0.13	−0.28	0.12	−0.14

* $p < 0.05$, ** $p < 0.01$, Statistics represent Spearman r correlations for each parameter. BMI, body mass index; WS, walking speed; GS, grip strength; WR, word recognition; TMT-A, Trail Making Test A version; TMT-B, Trail Making Test B version; SDST, Symbol Digit Substitution Task; TDAS, Touch Panel-type Dementia Assessment Scale; GDS-15, Geriatric Depression Scale-15.

Table 3. Multiple comparison among binomial logistic regression models depending on MCI classification with odds ratio.

Model	Coefficient (β)	Odds Ratio	95% CI	*p* Value
Model I				
Age (years)	−0.27	0.77	0.66, 0.89	0.000
TMT-B (s)	0.08	1.09	1.04, 1.14	0.001
SDST (score)	−0.12	0.88	0.80, 0.97	0.012
TDAS total score (score)	0.65	1.91	1.37, 2.68	0.000
Model II				
Age (years)	−0.26	0.77	0.67, 0.88	0.000
WR (score)	−0.55	0.58	0.44, 0.76	0.000
TMT-A (s)	0.17	1.19	1.07, 1.33	0.001
TMT-B (s)	0.05	1.05	1.01, 1.10	0.025
GDS-15 total score (score)	0.32	1.37	1.07, 1.77	0.014
Model III				
Walking speed (m/s)	−2.29	0.10	0.02, 0.69	0.020
GDS-15 total score (score)	0.20	1.22	1.04, 1.43	0.015

Reference group for analysis was non-MCI group. Model I: Model χ^2 test, $p < 0.0001$; The Hosmer–Lemeshow test, $p = 0.12$; Percentage of correct classifications = 87.4%. Model II: Model χ^2 test, $p < 0.0001$; The Hosmer–Lemeshow test, $p = 0.84$; Percentage of correct classifications = 84.5%. Model III: Model χ^2 test, $p = 0.002$; The Hosmer–Lemeshow test, $p = 0.02$; Percentage of correct classifications = 59.2%. CI, confidence interval; WR, word recognition; TMT-A, Trail Making Test A version; TMT-B, Trail Making Test B version; SDST, Symbol Digit Substitution Task; TDAS, Touch Panel-type Dementia Assessment Scale; GDS, Geriatric Depression Scale-15.

4. Results

Our samples consisted of 103 older participants divided into 54 non-MCI people and 49 MCI people. We confirmed that the MCI group had significantly lower scores or longer required times of all cognitive items including WR test, TMT-A, B, SDST and TDAS scores than the non-MCI group ($p < 0.0001$) (Table 1). Demographic and health data including Age, Gender, BMI, presence of Hypertension or Diabetes, Frail phenotype, presence of Weight loss, Poor energy, Low physical activity level, Amount of medications, and Education showed no significant difference between the non-MCI group and MCI group. Of physical assessments, WS was significantly different between the groups ($p = 0.03$), whereas GS was not different ($p = 0.25$). Moreover, the MCI group showed a worse score of GDS ($p = 0.046 < 0.05$). Next, we examined correlations between physical performance, cognitive and mental function in each group (Table 2). According to the results of Spearman correlation analysis, WS was associated with some items of cognitive subtests including WR, TMT-A, B, and SDST in the non-MCI group ($|r| > 0.30$, $p < 0.01$), but these were not significant in the MCI group except for correlations between cognitive items and Age or Education. Finally, we performed an analysis to determine explanatory variables for MCI with reference to non-MCI by binomial logistic regression analysis (Table 3). According to a result of Phi coefficient of association, all the nominal scales including Gender (Phi coefficient = 0.01, $p = 0.89$), presence of Hypertension (Phi coefficient = 0.02, $p = 0.86$) and Diabetes (Phi coefficient = 0.07, $p = 0.46$), Weight loss (Phi coefficient = 0.05, $p = 0.63$), Poor energy (Phi coefficient = 0.12, $p = 0.22$), Low physical activity level (Phi coefficient = 0.08, $p = 0.45$) were not significantly associated with MCI classification, and they were not included into predictors for the regression model. Three regression models were estimated according to the predictors of Age, GS, WS, Amount of medications, Education, WR, TMT-A, B, SDST, TDAS, and GDS-15 total score. Model I that included them demonstrated that the classification of MCI had a significant association with Age, TMT-B, SDST, and TDAS. Next, Model II except for T-DAS score from Model I was applied to estimate a specific cognitive profile in MCI. Model II demonstrated that the classification of MCI had a significant association with Age, WR, TMT A, B, and GDS-15 total score. Finally, considering the self-explanatory effect of cognitive items, Model III except for all cognitive variables from Model II was applied to clarify the classification of MCI.

As shown in Model III, WS and GDS-15 total score were extracted as explanatory variables of MCI (Table S1). In the three estimated models, the results of Hosmer–Lameshow test showed adaptability of 87.4% ($p = 0.12$) in Model I, 84.5% ($p = 0.84$) in Model II, and 59.2% ($p = 0.02$) in Model III.

5. Discussion

In this study, we found characteristics of MCI in northern Japanese community-dwellers of super-aging society had slower WS and tendency to depression. Aging continues in the subjects of our survey area, and the population ratio 65 years or older reached 38.7% (July, 2019). Actually, the prevalence of MCI in this study was higher (47.6%) compared with other rural areas which were previously reported to be about 10%–30% [29,33]. Additionally, some wealthy urban areas different from our rural area showed that characteristics of MCI were greater with older age and less education than non-MCI [34,35]. Although this high prevalence and multifactorial approach may be due to different methods, it could also be because our community-dwellers living in an area of heavy snowfall in northern Japan experience a more negative impact on gait performance [36] and a potentially high incidence of depressive symptoms [37] because of fewer opportunities to go out and participate in social activities. In fact, we showed an association between cognitive function and demographic and health data including age, gender, BMI, medical history, medication, frailty phenotype, education, physical performance, and GDS in older adults living in a super-aging society (Table 1). We found that recognized risk factors for MCI including age, gender, BMI, presence of hypertension or diabetes, frailty phenotype, education, and amount of medications were not different, but WS and GDS were significantly different between the groups. We also found that WS was significantly associated with some cognitive items including SDST and TMT in the non-MCI group, but not in the MCI group (Table 2). The regression models demonstrated that MCI had a significant association with age, executive function, information coping speed, and composite cognitive performance, indicating that these are predictive variables for the presence of MCI. However, because of the effect of variables on these cognitive scores (Model I), we applied Models II and III (Table 3). Model II excluding composite cognitive performance, as indexed in the TDAS score, demonstrated that MCI had a significant association with age, WR, attention, executive function, and GDS. Compatible with the results of Reinvang et al. [38], attention and executive dysfunction in neuropsychological tests could be early symptoms of MCI. Especially, the variables of SDST and TMT are recognized to reflect psychomotor processing and executive function [39], and several studies have reported that they are rapidly altered in MCI subjects [40,41]. Although they justify its use for the detection of cognitive impairment in older adults, most of these tests have numerous limitations (the problem of novelty, lack of sensitivity and specificity, patient cognitive reserve, etc.) [42,43]. This recent observation underscores the need to find new detection indicators for cognitive impairment. With this in perspective, a new approach associates WS of older adults with the presence of cognitive impairment.

Interestingly, in Model III excluding all cognitive domains, WS and GDS were selected as explanatory variables although the percentage of correct classifications was not so good in the Hosmer–Lameshow test. These findings indicate that the variables WS and GDS can potentially distinguish the presence or absence of MCI; therefore, they provide suggestive information on the presence of MCI. Recently, some studies have focused on both cognition and locomotor performance as predictors of adverse outcomes in community-dwellers with MCI [44,45]. In particular, slow gait speed at usual pace has been implicated in the onset of adverse outcomes, such as disability [46], cognitive impairment [47], institutionalization, falls [48,49], and mortality [50]. As previously reported, the association between slowing of walking and MCI is supported by shared neurological findings that include a smaller right hippocampus [51]. This finding underscores walking–brain behavior relationships and the value of WS as an early indicator of dementia risk. However, thus far, there is insufficient information to state that WS can potentially predict adverse outcomes in older community-dwellers, and more specific investigations need to be performed. Moreover, we showed that GS was no different between the groups ($p = 0.25$), suggesting that reinforcement of

lower, but not upper, limb muscular strength may be a critical target in rehabilitation. Likewise, recent studies have indicated that lower extremity motor dysfunction may be a feature of MCI [52], but little is known about the nature and biological mechanism such as myokines of lower extremity motor dysfunction associated with MCI. Regarding WS and a cognitive function, the concept of frailty has become a geriatric topic recently. Although we could not include frailty as global score in the correlation analysis or binomial regression analysis because the distribution of a frail group according to the frailty phenotype was greatly biased (e.g., % of robust/pre-frail/frail, 50%/50%/0% in the non-MCI group, 43%/47%/10% in the MCI group) (Table 1), some studies have reported that a physical frailty is associated with MCI and a reduction of WS in five items of the Fried index mostly reflect the occurrence of MCI and disability [31,53]. MCI with concomitant physical frailty may be considered to fulfil the criteria for cognitive frailty [54]. In this regard, we believe that the cognitive frailty concept has potential advantages in better stratifying the risk profiles of older adults with MCI. In a comparison between the groups, MCI also showed significantly higher depressive scores as indexed in the GDS. Concerning geriatric depression in MCI, cross-sectional research has shown that the association between depressive symptoms, as indexed in the Korean version of GDS, and memory or executive function was significantly greater in individuals with MCI than in those with AD [55]. Additionally, survival analysis followed for 6.28 years on average, indicating that the presence of MCI is a poor predictive factor in individuals with depressive symptoms as indexed in the GDS [56]. Thus, geriatric depressive symptoms in individuals with MCI need to be carefully screened in rural community-dwellers.

The limitations of our research need to be considered in developing our future research. First, the NCGG-FAT and TDAS used to classify individuals with MCI in this study were a tablet PC version of cognitive measurement tools based on the MCI criteria reported by Petersen [2], and evaluation of the accuracy of MCI's classification is essential for worldwide research. Second, our cohort was comprised of a localized group of individuals in one rural area of northern Japan, whose actual sample size (n = 103) did not reach the calculated required sample size (n = 128) due to difficulty sampling and recruiting in a depopulated, small rural area. Third, considering younger age was associated with MCI, we could not take the association into consideration. Fourth, although focusing this study on frailty concept was important, we guessed it was difficult to analyze frail status in detail due to bias of frail samples between the groups (e.g., 0% of the non-MCI group, 10% of the MCI group). Further examination concerning frailty is warranted in future research. Finally, we hypothesize that cognitive domains, gait performance, and tendency to depression might be associated with MCI status. For the three regression models in this study, WS and GDS were selected as explanatory variables in Model III. However, further research with sufficient adaptability should be carried out with a large sample size in multiple rural districts. These limitations need to be considered when interpreting this study's findings.

6. Conclusions

In conclusion, WS and GDS were shown to be potential predictive variables of MCI in our study, and we consider they provide important information about characteristics of MCI in rural community-dwellers. It is suggested that older individuals living in a super-aging society should work on training lower limb muscular strength, and avoiding depression in older adults by interaction of community-dwellers may contribute to prevention of the onset of MCI.

Supplementary Materials: The following are available online at http://www.mdpi.com/2077-0383/8/11/1937/s1, Table S1: Methodology of the binomial logistic regression models.

Author Contributions: Conceived the trial and participated in the study design: H.S., H.M., S.L., and H.O. Recruited and collected data: T.T., Y.I., T.O., and H.O. Analyzed data: Y.K. and H.O. All authors participated in interpretation of the results. Y.K. and H.O. drafted the manuscript, and all authors contributed to critical review and revision of the manuscript. H.O. takes responsibility for the manuscript as a whole.

Funding: This work was supported by the Japan Agency for Medical Research and Development (AMED) (Grant: 18dk0207027h0003).

Acknowledgments: The authors would like to thank all of participants for their enthusiasm in contributing to this project and making this study possible. We would also like to thank Soichi Kagaya (Akita University) for his management, and all staff at Yokote Hall and Omori Hospital who provided assistance in performing the assessments, and AMED for financial support.

Conflicts of Interest: All authors declare that they have 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. Petersen, R.C.; Stevens, J.C.; Ganguli, M.; Tangalos, E.G.; Cummings, J.L.; DeKosky, S.T. Practice parameter: Early detection of dementia: Mild cognitive impairment (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology* **2001**, *56*, 1133–1142. [CrossRef] [PubMed]
2. Petersen, R.C. Mild cognitive impairment as a diagnostic entity. *J. Intern. Med.* **2004**, *256*, 183–194. [CrossRef] [PubMed]
3. Öhman, H.; Savikko, N.; Strandberg, T.E.; Pitkälä, K.H. Effect of physical exercise on cognitive performance in older adults with mild cognitive impairment or dementia: A systematic review. *Dement. Geriatr. Cogn. Disord.* **2014**, *38*, 347–365. [CrossRef] [PubMed]
4. McKetton, L.; Cohn, M.; Tang-Wai, D.F.; Sobczyk, O.; Duffin, J.; Holmes, K.R.; Poublanc, J.; Sam, K.; Crawley, A.P.; Venkatraghavan, L.; et al. Cerebrovascular resistance in healthy aging and mild cognitive impairment. *Front. Aging Neurosci.* **2019**, *11*, 79. [CrossRef]
5. Mirza, S.S.; Ikram, M.A.; Bos, D.; Mihaescu, R.; Hofman, A.; Tiemeier, H. Mild cognitive impairment and risk of depression and anxiety: A population-based study. *Alzheimers Dement.* **2017**, *13*, 130–139. [CrossRef]
6. Panza, F.; Seripa, D.; Solfrizzi, V.; Tortelli, R.; Greco, A.; Pilotto, A.; Logroscino, G. Targeting cognitive frailty: Clinical and neurobiological roadmap for a single complex phenotype. *J. Alzheimers Dis.* **2015**, *47*, 793–813. [CrossRef]
7. Kluger, A.; Gianutsos, J.G.; Golomb, J.; Ferris, S.H.; George, A.E.; Franssen, E.; Reisberg, B. Patterns of motor impairment in normal aging, mild cognitive decline, and early Alzheimer's disease. *J. Gerontol. B Psychol. Sci. Soc. Sci.* **1997**, *52*, 28–39. [CrossRef]
8. McGough, E.L.; Kelly, V.E.; Logsdon, R.G.; McCurry, S.M.; Cochrane, B.B.; Engel, J.M.; Teri, L. Associations between physical performance and executive function in older adults with mild cognitive impairment: Gait speed and the timed "up & go" test. *Phys. Ther.* **2011**, *91*, 1198–1207.
9. Aggarwal, N.T.; Wilson, R.S.; Beck, T.L.; Bienias, J.L.; Bennett, D.A. Motor dysfunction in mild cognitive impairment and the risk of incident Alzheimer disease. *Arch. Neurol.* **2006**, *63*, 1763–1769. [CrossRef]
10. McGough, E.L.; Cochrane, B.B.; Pike, K.C.; Logsdon, R.G.; McCurry, S.M.; Teri, L. Dimensions of physical frailty and cognitive function in older adults with amnestic mild cognitive impairment. *Ann. Phys. Rehabil. Med.* **2013**, *56*, 329–341. [CrossRef]
11. Fitzpatrick, A.L.; Buchanan, C.K.; Nahin, R.L.; DeKosky, S.T.; Atkinson, H.H.; Carlson, M.C.; Williamson, J.D. Associations of gait speed and other measures of physical function with cognition in a healthy cohort of elderly persons. *J. Gerontol. A Biol. Sci. Med. Sci.* **2007**, *62*, 1244–1251. [CrossRef]
12. Montero-Odasso, M.; Muir, S.W.; Speechley, M. Dual-task complexity affects gait in people with mild cognitive impairment: The interplay between gait variability, dual tasking, and risk of falls. *Arch. Phys. Med. Rehabil.* **2012**, *93*, 293–299. [CrossRef] [PubMed]
13. Doi, T.; Makizako, H.; Shimada, H.; Yoshida, D.; Ito, K.; Kato, T.; Ando, H.; Suzuki, T. Brain atrophy and trunk stability during dual-task walking among older adults. *J. Gerontol. A Biol. Sci. Med. Sci.* **2012**, *67*, 790–795. [CrossRef] [PubMed]
14. Hooghiemstra, A.M.; Ramakers, I.H.G.B.; Sistermans, N.; Pijnenburg, Y.A.L.; Aalten, P.; Hamel, R.E.G.; Melis, R.J.; Verhey, F.R.; Olde Rikkert, M.G.; Scheltens, P.; et al. Gait speed and grip strength reflect cognitive impairment and are modestly related to incident cognitive decline in memory clinic patients with subjective cognitive decline and mild cognitive impairment: Findings from the 4C Study. *J. Gerontol. A Biol. Sci. Med. Sci.* **2017**, *72*, 846–854. [CrossRef] [PubMed]

15. Fried, L.P.; Tangen, C.M.; Walston, J.; Newman, A.B.; Hirsch, C.; Gottdiener, J.; Seeman, T.; Tracy, R.; Kop, W.J.; Burke, G.; et al. Cardiovascular Health Study Collaborative Research G. Frailty in older adults: Evidence for a phenotype. *J. Gerontol. A Biol. Sci. Med. Sci.* **2001**, *56*, M146–M156. [CrossRef] [PubMed]
16. Buchman, A.S.; Boyle, P.A.; Wilson, R.S.; Tang, Y.; Bennett, D.A. Frailty is associated with incident Alzheimer's disease and cognitive decline in the elderly. *Psychosom. Med.* **2007**, *69*, 483–489. [CrossRef] [PubMed]
17. Saji, N.; Sakurai, T.; Suzuki, K.; Mizusawa, H.; Toba, K. ORANGE investigators. ORANGE's challenge: Developing wide-ranging dementia research in Japan. *Lancet Neurol.* **2016**, *15*, 661–662. [CrossRef]
18. Japanese Ministry of Health; Labour and Welfare. Health and Welfare Bureau for the Elderly. Long-Term Care Insurance System of Japan. 2016. Available online: https://www.mhlw.go.jp/english/policy/care-welfare/care-welfare-elderly/dl/ltcisj_e.pdf (accessed on 17 June 2019).
19. Wada-Isoe, K.; Uemura, Y.; Nakashita, S.; Yamawaki, M.; Tanaka, K.; Yamamoto, M.; Shimokata, H.; Nakashima, K. Prevalence of dementia and mild cognitive impairment in the rural island Town of Ama-cho, Japan. *Dement. Geriatr. Cogn. Dis. Extra* **2012**, *2*, 190–199. [CrossRef]
20. Ishikawa, T.; Ikeda, M.; Matsumoto, N.; Shigenobu, K.; Brayne, C.; Tanabe, H. A longitudinal study regarding conversion from mild memory impairment to dementia in a Japanese community. *Int. J. Geriatr. Psychiatry* **2006**, *21*, 134–139. [CrossRef]
21. Makizako, H.; Shimada, H.; Park, H.; Doi, T.; Yoshida, D.; Uemura, K.; Tsutsumimoto, K.; Suzuki, T. Evaluation of multidimensional neurocognitive function using a tablet personal computer: Test-retest reliability and validity in community-dwelling older adults. *Geriatr. Gerontol. Int.* **2013**, *13*, 860–866. [CrossRef]
22. Shimada, H.; Makizako, H.; Doi, T.; Lee, S.; Lee, S. Conversion and reversion rates in Japanese older people with mild cognitive impairment. *J. Am. Med. Dir. Assoc.* **2017**, *18*, e801–e808. [CrossRef] [PubMed]
23. Inoue, M.; Jinbo, D.; Nakamura, Y.; Taniguchi, M.; Urakami, K. Development and evaluation of a computerized test battery for Alzheimer's disease screening in community-based settings. *Am. J. Alzheimers Dis. Other Dement.* **2009**, *24*, 129–135. [CrossRef] [PubMed]
24. Inoue, M.; Jimbo, D.; Taniguchi, M.; Urakami, K. Touch Panel-type Dementia Assessment Scale: A new computer-based rating scale for Alzheimer's disease. *Psychogeriatrics* **2011**, *11*, 28–33. [CrossRef] [PubMed]
25. Mohs, R.C.; Cohen, L. Alzheimer's Disease Assessment Scale (ADAS). *Psychopharmacol. Bull.* **1988**, *24*, 627–628.
26. Faul, F.; Erdfelder, E.; Buchner, A.; Lang, A.G. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behav. Res. Methods* **2009**, *41*, 1149–1160. [CrossRef]
27. Shimada, H.; Makizako, H.; Lee, S.; Doi, T.; Lee, S.; Tsutsumimoto, K.; Harada, K.; Hotta, R.; Bae, S.; Nakakubo, S.; et al. Impact of cognitive frailty on daily activities in older persons. *J. Nutr. Health Aging* **2016**, *20*, 729–735. [CrossRef]
28. Fukutomi, E.; Okumiya, K.; Wada, T.; Sakamoto, R.; Ishimoto, Y.; Kimura, Y.; Chen, W.L.; Imai, H.; Kasahara, Y.; Fujisawa, M.; et al. Relationships between each category of 25-item frailty risk assessment (Kihon Checklist) and newly certified older adults under Long-Term Care Insurance: A 24-month follow-up study in a rural community in Japan. *Geriatr. Gerontol. Int.* **2015**, *15*, 864–871. [CrossRef]
29. Shimada, H.; Makizako, H.; Doi, T.; Yoshida, D.; Tsutsumimoto, K.; Anan, Y.; Uemura, K.; Ito, T.; Lee, S.; Park, H.; et al. Combined Prevalence of Frailty and Mild Cognitive Impairment in a Population of Elderly Japanese People. *J. Am. Med. Dir. Assoc.* **2013**, *147*, 518–524. [CrossRef]
30. Chen, L.K.; Liu, L.K.; Woo, J.; Assantachai, P.; Auyeung, T.W.; Bahyah, K.S.; Chou, M.Y.; Chen, L.Y.; Hsu, P.S.; Krairit, O.; et al. Sarcopenia in Asia: Consensus Report of the Asian Working Group for Sarcopenia. *J. Am. Med. Dir. Assoc.* **2014**, *15*, 95–101. [CrossRef]
31. Makizako, H.; Shimada, H.; Doi, T.; Tsutsumimoto, K.; Suzuki, T. Impact of physical frailty on disability in community-dwelling older adults: A prospective cohort study. *BMJ Open* **2015**, *5*, e008462. [CrossRef]
32. Tsutsumimoto, K.; Doi, T.; Makizako, H.; Hotta, R.; Nakakubo, S.; Makino, K.; Suzuki, T.; Shimada, H. The association between anorexia of aging and physical frailty: Results from the national center for geriatrics and gerontology's study of geriatric syndromes. *Maturitas* **2017**, *97*, 32–37. [CrossRef] [PubMed]
33. Shimada, H.; Makizako, H.; Doi, T.; Tsutsumimoto, K.; Lee, S.; Suzuki, T. Cognitive impairment and disability in older Japanese adults. *PLoS ONE* **2016**, *11*, e0158720. [CrossRef] [PubMed]
34. Kluger, A.; Gianutsos, J.G.; Golomb, J.; Wagner, A., Jr.; Wagner, D.; Scheurich, S. Clinical features of MCI: Motor changes. *Int. Psychogeriatr.* **2008**, *20*, 32–39. [CrossRef] [PubMed]

35. Ravaglia, G.; Forti, P.; Montesi, F.; Lucicesare, A.; Pisacane, N.; Rietti, E.; Dalmonte, E.; Bianchin, M.; Mecocci, P. Mild cognitive impairment: Epidemiology and dementia risk in an elderly Italian population. *J. Am. Geriatr. Soc.* **2008**, *56*, 51–58. [CrossRef]
36. Mizumoto, A.; Ihira, H.; Makino, K.; Saitoh, S.; Ohnishi, H.; Furuna, T. Physical activity changes in the winter in older persons living in northern Japan: A prospective study. *BMC Geriatr.* **2015**, *15*, 43. [CrossRef]
37. Lewy, A.J.; Emens, J.S.; Songer, J.B.; Sims, N.; Laurie, A.L.; Fiala, S.C.; Buti, A.L. Winter depression: Integrating mood, circadian rhythms, and the sleep/wake and light/dark cycles into a bio-psycho-social-environmental model. *Sleep Med. Clin.* **2009**, *4*, 285–299. [CrossRef]
38. Reinvang, I.; Grambaite, R.; Espeseth, T. Executive dysfunction in MCI: Subtype or early symptom. *Int. J. Alzheimers Dis.* **2012**, *2012*, 936272. [CrossRef]
39. Lowry, K.A.; Brach, J.S.; Nebes, R.D.; Studenski, S.A.; VanSwearingen, J.M. Contributions of cognitive function to straight- and curved-path walking in older adults. *Arch. Phys. Med. Rehabil.* **2012**, *93*, 802–807. [CrossRef]
40. Chapman, R.M.; Mapstone, M.; McCrary, J.W.; Gardner, M.N.; Porsteinsson, A.; Sandoval, T.C.; Guillily, M.D.; DeGrush, E.; Reilly, L.A. Predicting conversion from mild cognitive impairment to Alzheimer's disease using neuropsychological tests and multivariate methods. *J. Clin. Exp. Neuropsychol.* **2011**, *33*, 187–199. [CrossRef]
41. Clark, L.R.; Schiehser, D.M.; Weissberger, G.H.; Salmon, D.P.; Delis, D.C.; Bondi, M.W. Specific measures of executive function predict cognitive decline in older adults. *J. Int. Neuropsychol. Soc.* **2012**, *18*, 118–127. [CrossRef]
42. Brooks, B.L.; Strauss, E.; Sherman, E.M.S.; Iverson, G.L.; Slick, D.J. Developments in neuropsychological assessment: Refining psychometric and clinical interpretive methods. *Can. Psychol.* **2009**, *50*, 196–209. [CrossRef]
43. Weston, A.L.; Weinstein, A.M.; Barton, C.; Yaffe, K. Potentially inappropriate medication use in older adults with mild cognitive impairment. *J. Gerontol. A Biol. Sci. Med. Sci.* **2010**, *65*, 318–321. [CrossRef] [PubMed]
44. Perrochon, A.; Kemoun, G. The Walking Trail-Making Test is an early detection tool for mild cognitive impairment. *Clin. Interv. Aging* **2014**, *9*, 111–119. [CrossRef] [PubMed]
45. Mielke, M.M.; Roberts, R.O.; Savica, R.; Cha, R.; Drubach, D.I.; Christianson, T.; Pankratz, V.S.; Geda, Y.E.; Machulda, M.M.; Ivnik, R.J.; et al. Assessing the temporal relationship between cognition and gait: Slow gait predicts cognitive decline in the Mayo Clinic Study of Aging. *J. Gerontol. A Biol. Sci. Med. Sci.* **2013**, *68*, 929–937. [CrossRef] [PubMed]
46. Abe, T.; Kitamura, A.; Taniguchi, Y.; Amano, H.; Seino, S.; Yokoyama, Y.; Nishi, M.; Narita, M.; Ikeuchi, T.; Fujiwara, Y.; et al. Pathway from gait speed to incidence of disability and mortality in older adults: A mediating role of physical activity. *Maturitas* **2019**, *123*, 32–36. [CrossRef] [PubMed]
47. Abellan van Kan, G.; Rolland, Y.; Andrieu, S.; Bauer, J.; Beauchet, O.; Bonnefoy, M.; Cesari, M.; Donini, L.M.; Gillette Guyonnet, S.; Inzitari, M.; et al. Gait speed at usual pace as a predictor of adverse outcomes in community-dwelling older people an International Academy on Nutrition and Aging (IANA) Task Force. *J. Nutr. Health Aging* **2009**, *13*, 881–889. [CrossRef]
48. Kyrdalen, I.L.; Thingstad, P.; Sandvik, L.; Ormstad, H. Associations between gait speed and well-known fall risk factors among community-dwelling older adults. *Physiother. Res. Int.* **2019**, *24*, e1743. [CrossRef]
49. Nakakubo, S.; Doi, T.; Makizako, H.; Tsutsumimoto, K.; Hotta, R.; Kurita, S.; Kim, M.; Suzuki, T.; Shimada, H. Association of walk ratio during normal gait speed and fall in community-dwelling elderly people. *Gait Posture* **2018**, *66*, 151–154. [CrossRef]
50. Veronese, N.; Stubbs, B.; Volpato, S.; Zuliani, G.; Maggi, S.; Cesari, M.; Lipnicki, D.M.; Smith, L.; Schofield, P.; Firth, J.; et al. Association between gait speed with mortality, cardiovascular disease and cancer: A systematic review and meta-analysis of prospective cohort studies. *J. Am. Med. Dir. Assoc.* **2018**, *19*, 981–988. [CrossRef]
51. Rosso, A.L.; Verghese, J.; Metti, A.L.; Boudreau, R.M.; Aizenstein, H.J.; Kritchevsky, S.; Harris, T.; Yaffe, K.; Satterfield, S.; Studenski, S.; et al. Slowing gait and risk for cognitive impairment: The hippocampus as a shared neural substrate. *Neurology* **2017**, *89*, 336–342. [CrossRef]
52. Boyle, P.A.; Wilson, R.S.; Buchman, A.S.; Aggarwal, N.T.; Tang, Y.; Arvanitakis, Z.; Kelly, J.; Bennett, D.A. Lower extremity motor function and disability in mild cognitive impairment. *Exp. Aging Res.* **2007**, *33*, 355–371. [CrossRef] [PubMed]

53. Park, E.Y.; Park, S.M.; Kim, J.H. Psychometric properties of the geriatric quality of life-dementia in older adults with dementia or mild cognitive impairment living in nursing homes. *BMC Geriatr.* **2019**, *19*, 281. [CrossRef] [PubMed]
54. Kelaiditi, E.; Cesari, M.; Canevelli, M.; Van Kan, G.A.; Ousset, P.J.; Gillette-Guyonnet, S.; Ritz, P.; Duveau, F.; Soto, M.E.; Provencher, V.; et al. Cognitive frailty: Rational and definition from an (IANA/IAGG) international consensus group. *J. Nutr. Health Aging* **2013**, *17*, 726–734. [CrossRef] [PubMed]
55. Lee, C.H.; Kim, D.H.; Moon, Y.S. Differential associations between depression and cognitive function in MCI and AD: A cross-sectional study. *Int. Psychogeriatr.* **2019**, *31*, 1151–1158. [CrossRef] [PubMed]
56. Steffens, D.C.; McQuoid, D.R.; Potter, G.G. Amnestic mild cognitive impairment and incident dementia and Alzheimer's disease in geriatric depression. *Int. Psychogeriatr.* **2014**, *26*, 2029–2036. [CrossRef]

Journal of
Clinical Medicine

MDPI

Article

Cerebral White Matter Hyperintensity as a Healthcare Quotient

Kaechang Park [1,*,†], Kiyotaka Nemoto [2,†], Yoshinori Yamakawa [3], Fumio Yamashita [4], Keitaro Yoshida [2], Masashi Tamura [2], Atsushi Kawaguchi [5], Tetsuaki Arai [2] and Makoto Sasaki [4]

1 Research Organization for Regional Alliance, Kochi University of Technology, Kochi 782-0003, Japan
2 Department of Psychiatry, Faculty of Medicine, University of Tsukuba, Tsukuba 305-8575, Japan; kiyotaka@nemotos.net (K.N.); keicima0154@gmail.com (K.Y.); tamuramas@gmail.com (M.T.); 4632tetsu@md.tsukuba.ac.jp (T.A.)
3 ImPACT Program of Council for Science, Technology and Innovation (Cabinet Office, Government of Japan), Tokyo 100-8974, Japan; yamakawa@bi-lab.org
4 Division of Ultrahigh Field MRI, Institute for Biomedical Sciences, Iwate Medical University, Morioka 028-3694, Japan; fyamashi@iwate-med.ac.jp (F.Y.); masasaki@iwate-med.ac.jp (M.S.)
5 Center for Comprehensive Community Medicine, Faculty of Medicine, Saga University, Saga 849-8501, Japan; akawa@cc.saga-u.ac.jp
* Correspondence: park.kaechang@kochi-tech.ac.jp; Tel.: +81-887-57-2025; Fax: +81-887-57-2026
† These authors contributed equally to this work.

Received: 18 August 2019; Accepted: 27 October 2019; Published: 1 November 2019

Abstract: To better understand the risk factors and optimal therapeutic strategies of cerebral white matter hyperintensity (WMH), we examined a large population of adults with and without various vascular risk factors (VRFs) or vascular risk conditions (VRCs), such as hypertension (HT), diabetes mellitus (DM), and dyslipidemia (DLP), including the comorbidities. We assessed two participant groups having no medical history of stroke or dementia that underwent brain checkup using magnetic resonance imaging (MRI): 5541 participants (2760 men, 2781 women) without VRCs and 1969 participants (1169 men, 800 women) who had received drug treatments for VRCs and the combination of comorbidities. For data analysis, we constructed WMH-brain healthcare quotient (WMH-BHQ) based on the percentile rank of WMH volume. This metric has an inverse relation to WMH. Multiple linear regression analysis of 5541 participants without VRCs revealed that age, systolic blood pressure (SBP), Brinkman index (BI), and female sex were significant factors lowering WMH-BHQ, whereas body mass index (BMI), male sex, fasting blood sugar, and triglyceride levels were increasing factors. The Kruskal–Wallis test and Dunn tests showed that WMH-BHQs significantly increased or decreased with BMI or SBP and with BI classification, respectively. Regarding the impact of impaired fasting glucose and abnormal lipid metabolism, there were almost no significant relationships. For 1969 participants who had HT, DM, and DLP, as well as their comorbidities, we found that DLP played a substantial role in increasing WMH-BHQ for some comorbidities, whereas the presence of HT and DM alone tended to decrease it. Cerebral WMH can be used as a healthcare quotient for quantitatively evaluating VRFs and VRCs and their comorbidities.

Keywords: white matter hyperintensity; MRI; healthcare quotient; chronic

1. Introduction

Cerebral vessel diseases are classified as large vessels diseases (LVDs) or small vessels diseases (SVDs) based on whether the diameters of the vessels involved are larger than a few millimeters or smaller than several hundred micrometers, respectively [1,2]. Both categories can be noninvasively diagnosed using magnetic resonance imaging (MRI) [3]. With regard to risk factors, many longitudinal

studies have reported that LVD can be responsible for stroke, cognitive decline, and dementia [1–4]. Hypertension (HT), diabetes mellitus (DM), and dyslipidemia (DLP), three primary vascular risk conditions (VRCs) in developed countries, are risk factors for LVD [5,6]. Although HT is an obvious risk factor for SVD, the roles of DM and DLP remain disputable [7,8]. Compared with LVD, SVD has not yet been sufficiently studied with regard to its onset and development. The pathological complexity of SVD, such as arteriosclerosis, hyalinosis, blood–brain barrier disruption, and venous collagenosis, have long complicated our ability to fully comprehend its many aspects [1,8]. For SVD studies using MRI, it is difficult to include large numbers of participants who have various conditions ranging from preclinical to chronic HT, DM, and DLP, partly because SVD is mostly asymptomatic and does not have hospital follow-up like LVD. To clarify the whole range of cerebral vessel damages, a large scale epidemiological study of SVD including preclinical or chronic HT, DM, and DLP is essential.

Brain MRIs show four major features of SVD: lacuna stroke, white matter hyperintensity (WMH), cerebral microbleeds, and visible perivascular spaces [9]. In our study, we focused on WMHs, also known as leukoaraiosis, which are commonly observed in the general population, particularly among individuals with preclinical or chronic HT, DM, and DLP. WMHs are recognized in >60% of people over 60 years old [10] and >30% of people with the age range from 40 to 50 years in Japan, where MRI examination is incorporated as part of health checkups in connection with a screening program called Brain Dock [11,12]. WMHs are regarded as disappearance of arterioles and capillary arteries caused by aging, HT, and reduced cerebral blood flow [9,13]. WMHs are also significantly associated with recurrent stroke, cognitive decline, and dementia [2,5].

Numerous efforts have been made to develop MRI-based measures of health status, such as the concept of "brain age", which reportedly reflects the mortality of an individual [14]. Our team earlier proposed brain healthcare quotients (BHQs) based on gray matter volume or fractional anisotropy and found significant associations between these proffered metrics and various physical factors, such as obesity, high blood pressure (BP), and daily personal schedules, as well as social factors, including subjective socioeconomic status, subjective well-being, and the adoption of a postmaterialism view of life [15]. In this cross-sectional study, we proposed another BHQ based on WMH, which begins to appear in early middle age and increases in frequency with age. Using an extensively large database obtained from 8921 participants who were examined through MRI as part of the Brain Dock component of a routine health checkups, we analyzed two groups of individuals with VRFs: those without VRCs and those with VRCs receiving drug treatment for high BP, impaired fasting glucose (IFG), or abnormal lipid metabolism (ALM), each of which chronically results in the onset of HT, DM, or ALM, respectively. In the drug treatment group, the BHQs of WMH were compared according to the comorbidity of HT, DM, or DLP because these VRCs commonly combine together. Nonetheless, the relationship between WMH and comorbidity remains remarkably unclear [9,10,13]. To help make progress in this area, we designed and executed a large scale, cross-sectional study covering healthy and non-healthy states ranging from preclinical to chronic HT, DM, and DLP to examine whether WMH can be used as a healthcare quotient to maintain a healthy state or prevent the onset and development of VRCs.

2. Materials and Methods

2.1. Participants

Data were collected between January 2013 and April 2017 from the brain dock center (BDC) affiliated with Kochi University of Technology. From BDC, we enrolled 8921 healthy participants without a history of cerebral stroke, who underwent the brain dock health checkups only once. Although we were interested in the WMH of individuals with various medical backgrounds, participants who had been clinically diagnosed with HT, DM, and DLP but had not been treated with drugs were excluded (n = 1411). Thus, 5541 participants (2760 males, 2781 females; age, 20–89 years; mean age ± SD, 51.38 ± 9.80 years; median age, 51 years) with no medical history for HT, DM, and DLP were selected for analysis (Table 1). Here, the term "medical history" refers to the drug treatment history before and

at the time of enrollment in the study. In addition, the following participants were also enrolled for analysis based on the examination results at BDC that compared WMH-BHQ of the participants with no medical history with that of those with a medical history of HT, DM, and DLP or their comorbidities (n = 1969).

Table 1. Number and age distribution of participants without and with hypertension (HT), diabetes mellitus (DM), and/or dyslipidemia (DLP).

	Total	Male	Female	Mean Age ± SD (Years)	Median Age (Years)
No medical history	5541	2760	2781	51.4 ± 9.8	51
HT only	1074	622	422	59.5 ± 9.3	59
DM only	150	114	36	59.3 ± 9.8	59
DLP only	299	124	175	56.9 ± 8.7	57
HT + DLP	220	124	96	60.0 ± 9.0	59
HT + DM	124	97	27	60.8 ± 9.0	60
DM + DLP	35	19	16	57.0 ± 7.6	57
HT + DM + DLP	67	39	28	60.4 ± 7.6	59

All participants lived in Kochi Prefecture, visited BDC, and underwent brain MRI as part of their routine health checkups. They also answered a questionnaire on their past and present medical history and lifestyles, such as smoking. Health checkups included systolic blood pressure (SBP), body mass index (BMI), Brinkman index (BI; multiplying the average number of cigarettes smoked per day by the number of years the person has smoked), and various blood chemistry test items, including hemoglobin A1c (HbA1c), fasting blood sugar (FBS), triglycerides (TG), and high- density lipoprotein (HDL) and low-density lipoprotein (LDL) cholesterol. Based on these tests, BMI, BI, High BP, IFG, and ALM were classified according to the criteria shown in Table 2.

Table 2. Classification of body mass index (BMI), Brinkman index (BI), high blood pressure (BP), and criteria of impaired fasting glucose (IFG) and abnormal lipid metabolism (ALM).

	Classification and Criteria
BMI	Underweight, BMI < 18.5; normal, 18.5 ≤ BMI < 25, overweight, 25 ≤ BMI< 30; and obese, BMI ≥ 30
BI	BI = 0, 0 < BI < 400, and BI ≥ 400
High BP	Systolic blood pressure (SBP) < 139, 140 ≤ SBP < 160, SBP ≥ 160
IFG	A (fasting blood sugar (FBS) < 100 and HbA1c < 5.6%) B (100 ≤ FBS < 100 or 5.6% ≤ HbA1c < 6.0%) C (110 ≤ FBS < 126 or 6.0% ≤ HbA1c < 6.5%) D (FBS ≥ 126 or HbA1c ≥ 6.5%)
ALM	Triglycerides (TG); TG < 29, 30 ≤ TG< 140, 140 ≤ TG< 400, and TG ≥ 400 Ratios of LDL to HDL (LH ratio): LH ratio < 1, 1 ≤ LH ratio < 1.5, 1.5 ≤ LH ratio < 2, 2 ≤ LH ratio < 2.5, and LH ratio ≥ 2.5

2.2. Automated Measurement of WMH Volume

A 1.5 Tesla MRI system (ECHELON Vega; Hitachi Medical Corporation, Tokyo, Japan) was used to perform MRI examinations for WMH diagnosis. The imaging protocol included T2-weighted spin-echo (repetition time/echo time (TR/TE) = 5800/96 ms), T1-weighted spin-echo (TR/TE = 520/14 ms), and fluid-attenuated inversion recovery (FLAIR; TR/TE = 8500/96 ms; inversion time = 2100 ms) images as described previously [16]. Images were obtained as 27 transaxial slices per scan. The slice thickness was 5 mm, with no interslice gap, as described previously [11,16]. Measurement of WMH volume was needed to evaluate the severity, especially for levels more than the maximum of the Fazekas scale. In our study, WMHs were automatically segmented and quantified for their volume using the following procedure. First, the FLAIR images were segmented into gray and white matter and cerebrospinal fluid space using SPM12, which also yielded the intensity inhomogeneity corrected image (IICI) [17,18]. Then, IICI was anatomically normalized into the template space using advanced

normalization tools. A region-of-interest delineating the middle cerebellar peduncle was applied to the anatomically normalized IICI to estimate the intensity distribution of normal white matter of each subject. IICI in native space was then normalized for its intensity in the brain region segmented by the gray and white matter. The intensity normalized IICI was thresholded using a 3.5 SD cutoff to segment WMH, with search regions limited to WMH mask. The WMH volume (WMHV) was calculated by multiplying the voxel size by slice thickness. Finally, the measured WMH was automatically colored red to be detected by the first author (K.P.) who were a neurosurgeon trained enough to confirm the presence and location of WMH.

2.3. WMH-BHQ, a Novel Quotient Based on WMHV

Based on the WMHV of each participant, we constructed a new metric, the WMH-BHQ in that higher values are better and the median value for a given set of subjects is 100. In other words, this new metric was devised to convert WMHV to the standardized scale so that one can easily understand whether the WMHV of a subject is more or less than the median. In the development of this metric, we realized that the distribution of WMHV is skewed, and therefore, we used the percentile rank to define WMH-BHQ, which is the percentage of scores in its frequency distribution that are equal to or lower than it. For example, a test score that is greater than 75% of the scores of people taking the test is said to be at the 75th percentile, where 75 is the percentile rank. From the raw WMHV we obtained the percentile ranking for each subject (WMHV percentile), where zero means the lowest WMHV and 100 percentile means that the participant has the highest WMHV in the group. From this ranking method, a cumulative probability curve was estimated so that the percentile rank with newer data could be calculated with this curve. This estimation was based on the nonparametric density estimation and implemented by using the polspline function in the "polspline" package with R 3.4.3. We then defined this new brain health metric, WMH-BHQ, to be:

$$\text{WMH-BHQ} = 100 + 15 \times (50 - \text{WMHV percentile})/24, \quad (1)$$

With this formula, the median value, which was equivalent to the 50th percentile, generates a WMH-BHQ equal to 100. Likewise, the 74th and 26th percentiles produce BHQ 85 and 115, respectively. Originally, we considered directly using interquartile range (i.e., 75th and 25th percentiles). However, if the newer data were beyond the range of original data, an error would be produced. To avoid this prospect, we used the value 24 instead of 25 in the denominator. As a result, the WM-BHQ of 95% of our subjects ranged from 70.31 to 130. A lower WMH-BHQ means a higher and more problematic level of WMHV. A histogram of the WMH-BHQ values of our subjects with upward sloping curve implies "well-being" of brain health and a downward slope shows "not well-being" in terms of WMH.

2.4. Statistical Analysis

WMHV and WMH-BHQ data were not normally distributed. Thus, Mann–Whitney *U* and Kruskal–Wallis tests were utilized to evaluate the associations between WMHs and VRCs or other possible risk factors by comparing the differences between group distributions. The groups were defined by the presence or absence of VRCs or by standard criteria and classifications of risk factors described in the upper paragraphs. When the null hypotheses were rejected in the Kruskal–Wallis tests, we used Dunn tests [18] for pairwise comparisons. The *p* values were then adjusted using the Benjamini–Hochberg procedure [19], which controls the false discovery rate for multiple comparisons. Multiple regression analyses were performed to examine complex associations among multiple variables while controlling for the effect of potential confounding factors [20]. All statistical analyses except for the Benjamini–Hochberg procedure were performed using the Statistical Package for the Social Sciences software version 22 (IBM Corp., Armonk, NY, USA) [20]. Adjusted *p* values based on

the Benjamini–Hochberg procedure were calculated using a Microsoft Excel (Microsoft Inc, Redmond, WA, USA) spreadsheet [21].

2.5. Standard Protocol Approvals, Registrations, and Participant Consents

Written informed consent was received from all participants and this study was reviewed and approved by the institutional review board of Kochi University of Technology.

2.6. Data Availability

Anonymized data might be shared by request.

3. Results

3.1. WMH-BHQ of Participants with no Medical History According to Age Decades

As shown in Figure 1a, the contour of the histogram for our metric WMH-BHQ changed remarkably across the different ages of the participants in this study. For subjects in their 40s, it was up right, while for those in their 50s, it was symmetric like a rainbow curve, and those in their 60s, 70s, and 80s generated plots shifting up and to the left as the age decades increased. In terms of WMH-BHQ, brain health obviously declined with age, as shown by the box plot in Figure 1b. The Kruskal–Wallis test and Dunn tests showed that all pairwise comparisons were significant except for those between 70s and 80s (p = 0.076). The WMHV histogram was asymptotic with a peak volume of <5 mL; therefore, WMH-BHQ was clearly superior to WMHV with regard to visualization and understanding of changes occurring over the age decades.

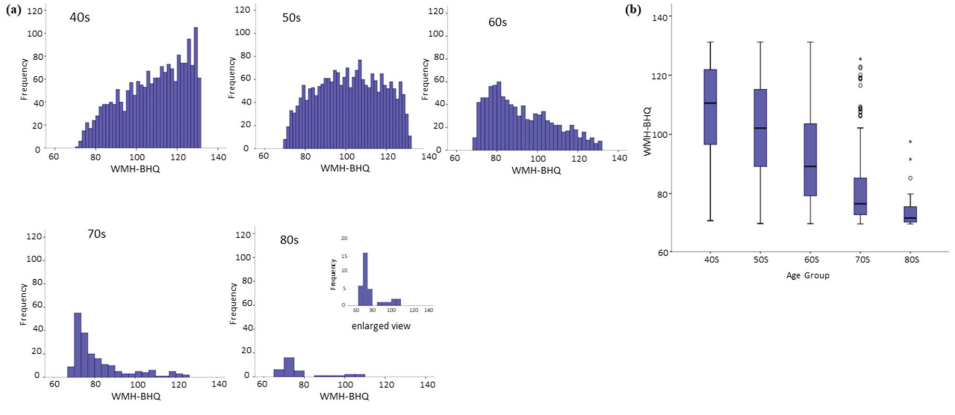


Figure 1. Histograms (**a**) and box plots (**b**) of WMH-BHQ with no medical history according to age decades of 40s, 50s, 60s, 70s, and 80s.

3.2. WMH-BHQ Histograms of Sex

There was a distinct difference in WMH-BHQ between males and females (Figure 2). The histogram of males was a trapezoid with an upward slope, while that of females showed a plateau, suggesting that females are more susceptive to WMH than males.

Compared with results across the various age decades, the Mann–Whitney U test showed a clear sex difference for participants younger than their 50s but not for participants in their 60s and beyond (Table 3).

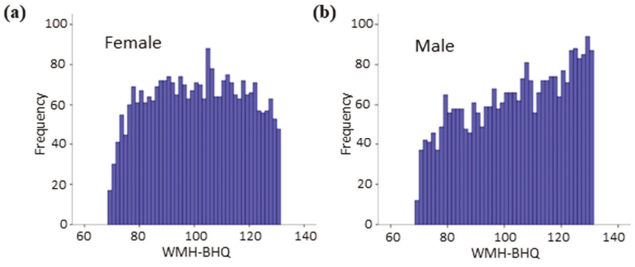


Figure 2. WMH-BHQ Histograms of females (**a**) and males (**b**).

Table 3. Mann-Whitney U test of white matter hyperintensity brain healthcare quotient (WMH-BHQ) without medical history according to genders and age decades.

Mann-Whitney U (M-W U) Test							
	Male		Female				
Age Decade	N	Mean Rank	N	Mean Rank	M-W U	Z	P
30s	351	276.7	239	323.1	35361	−3.24	0.001
40s	955	855.0	865	971.8	359991	−4.74	0.001
50s	920	971.9	1117	1057.8	470485.5	−3.28	0.001
60s	437	449.5	460	448.5	100299.5	−0.05	0.957
70s	89	88.0	82	83.9	3473	−0.54	0.586
80s	8	14.9	18	12.9	61	−0.61	0.541

3.3. Analysis of WMH-BHQ Risk Factors: No VRCs

Multiple linear regression analysis of participants without VRCs was performed using age, sex, BMI, BI, SBP, HbA1c, FBS, TG, HDL, and LDL as independent variables and WMH-BHQ as a dependent variable (Table 4). A stepwise model was adapted for variable selection procedure. Female or male sex was a significant risk factor lowering or raising WMH-BHQ, respectively. The increases in age, SBP, and BI were significantly associated with the decrease in WMH-BHQ, whereas the increases in BMI, FBS, and TG were significantly associated with the increase in WMH-BHQ. HbA1c, HDL, and LDL were excluded after the stepwise regressions.

Table 4. Multiple linear regression analysis for white matter hyperintensity brain healthcare quotients (WMH-BHQ) risk factors.

Model	Unstandardized Coefficients		Standardized Coefficients	t	*p*	95% Confidence Interval for B	
	B	Standard Error	Beta			Lower Bound	Upper Bound
(Constant)	125.170	2.256		55.49	<0.001	120.750	129.595
Age	−0.064	0.022	−0.377	−29.53	<0.001	−0.708	−0.620
Body mass index (BMI)	0.693	0.070	0.135	9.86	<0.001	0.556	0.831
Systolic blood pressure (SBP)	−0.079	0.015	−0.074	−5.41	<0.001	−0.108	−0.051
Brinkman index (BI)	−0.003	0.001	−0.049	−3.88	<0.001	−0.004	−0.001
Fasting blood glucose (FBG)	0.045	0.015	0.038	2.96	0.003	0.015	0.074
Triglycerides (TG)	0.008	0.003	0.036	2.75	0.006	0.002	0.013

3.4. WMH-BHQ without VRCs: the Effect of Three Classifications and Two Criteria

We explored in detail the impact of differing values for BMI, BI, high BP and the criteria of IFG and ALM. Regarding BMI, the box plots in Figure 3a showed that WMH-BHQ significantly increased as the classification of BMI became larger. The Kruskal–Wallis test and the following Dunn tests showed all pairwise comparisons to be significantly different. BMI was also positively associated with WMH-BHQ. The effect of cigarette smoking, as measured by the BI, was that lower levels were associated with higher WMH-BHQ values. In particular, a BI value of 0 (Dunn test; $p < 0.001$) and 0–400 ($p < 0.001$) revealed significantly better bran brain health than levels above 400 (Figure 3b).

Also, from Figure 4a, it is apparent that WMH-BHQ declined as SBP increased. Regarding high BP, WMH-BHQs of ≥400 ($p < 0.010$) and 0–400 ($p < 0.001$) revealed significant decreases compared with levels of ≥400 (Figure 4a). For IFG, a significant relationship existed only between (FBS < 100 and HbA1c < 5.6%) and (FBS ≥ 100 and <110 or HbA1c ≥ 5.6% and < 6.0%) ($p < 0.001$), although the other pair matches showed no significance (Figure 4b). TG showed a significant relationship only between 30 ≤ TG < 149 and 150 ≤ TG < 399 ($p = 0.001$), although the other pair matches showed no significance (Figure 4c). The LH ratio showed significant differences between 1 ≤ LH ratio < 1.5 and 2.5 ≤ LH ratio ($p = 0.010$) and between 1.5 ≤ LH ratio < 2.0 and 2.5 ≤ LH ratio ($p = 0.018$), although the other pair matches showed no significance (Figure 4d).

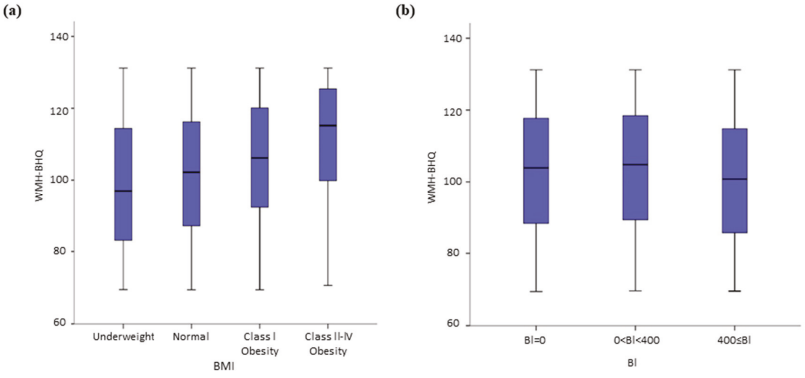


Figure 3. Box plots of WMH-BHQ with no medical history according to (**a**) Body Mass Index (BMI) and (**b**) Brinkman Index (BI).

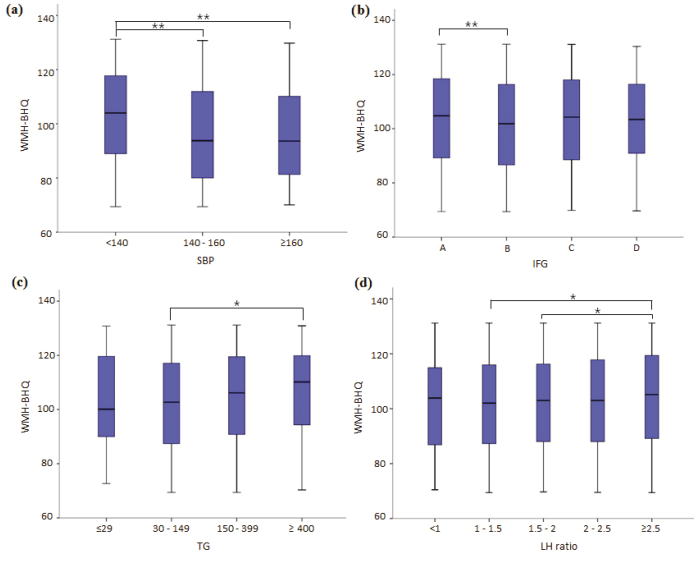


Figure 4. Box plots of WMH-BHQ with no medical history according to (**a**) systolic BP (SBP), (**b**) triglyceride (TG), (**c**) impaired fasting glucose (IFG) criteria, and d) the ratios of LDL to HDL (LH ratio). (**d**) A: fasting blood sugar (FBS) < 100 and HbA1c < 5.6; B: 100 ≤ FBS < 110 or 5.6 ≤ HbA1c < 6.0; C: 110 ≤ FBS < 126 or 6.0 ≤ HbA1c < 6.5; D: FBS ≥ 126 or HbA1c ≥ 6.5. * $p < 0.05$; ** $p < 0.001$

3.5. WMH-BHQ with VRCs and Their Comorbidities

WMH-BHQ histograms showed a downward slope for HT, while those of DM and DLP are almost plateaued compared with HT (Figure 5a). We also analyzed WMH-BHQ patterns regarding various multimorbidities, specifically, HT+DM, HT+DLP, DM+DLP, and HT+DM+DLP. Somewhat surprisingly, DM+DLP and HT+DM+DLP showed no downward slopes. Box plots showed that "no medical history" had the highest and HT+DM had the lowest median WMH-BHQs (Figure 5b). The Kruskal–Wallis test and the following Dunn test showed that all but DLP+DM had a significant difference in mean rank compared with no medical history (Table 5).

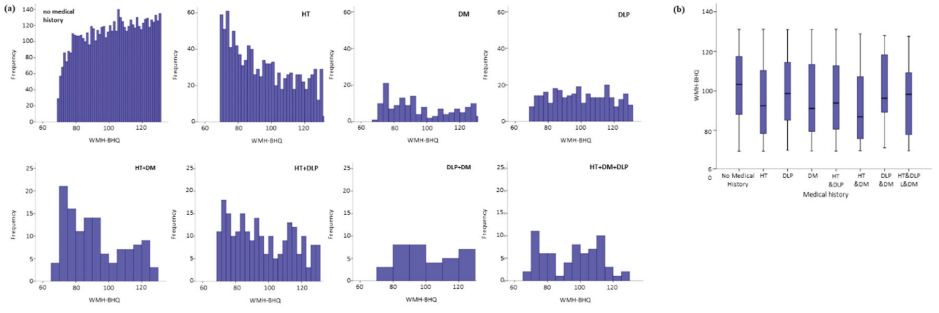


Figure 5. Histograms (**a**) and box plots (**b**) of WMH-BHQ according to no medical history, single morbidity, and multiple comorbidities. HT: Hypertension, DM: Diabetes mellitus, DLP: Dyslipidemia.

Table 5. Dunn test of WMH-BHQ among no medical history, hypertenstion (HT), diabetes mellitus (DM), dyslipidemia (DLP), and the comorbidities.

Pairwise Comparison by Dunn Test						
Comparison 1	**Comparison 2**	**Test Statistic**	**SD**	**Z**	**Unadjusted *p*-Value**	**Adjusted *p*-Value**
No medical history	HT	911.74	72.29	12.61	<0.001	<0.001
	DLP	368.32	128.72	2.86	0.004	0.015
	DM	793.29	179.40	4.42	<0.001	<0.001
	HT + DLP	765.91	149.05	5.14	<0.001	<0.001
	HT + DM	1307.35	196.87	6.64	<0.001	<0.001
	DLP + DM	188.32	367.63	0.51	0.608	0.710
	HT + DLP + DM	928.24	266.47	3.48	<0.001	0.002
HT	DLP	−543.42	141.77	−3.83	<0.001	0.001
	DM	−118.45	188.98	−0.63	0.531	0.676
	HT + DLP	−145.83	160.45	−0.91	0.363	0.485
	HT + DM	395.61	205.63	1.92	0.054	0.098
	DLP + DM	−723.42	372.40	−1.94	0.052	0.098
	HT + DLP + DM	16.49	273.01	0.06	0.952	0.952
DLP	DM	424.97	216.93	1.96	0.050	0.098
	HT + DLP	397.59	192.58	2.07	0.039	0.098
	HT + DM	939.03	231.58	4.06	<0.001	<0.001
	DLP + DM	−180.00	387.33	−0.47	0.642	0.719
	HT + DLP + DM	559.91	293.05	1.91	0056	0.098
DM	HT + DLP	−27.38	229.57	−0.12	0.905	0.939
	HT + DM	514.06	263.15	1.95	0.051	0.098
	DLP + DM	−604.97	406.99	−1.49	0.137	0.211
	HT + DLP + DM	134.94	318.59	0.42	0.672	0.724
HT + DLP	HT + DM	541.44	243.46	2.22	0.026	0.073
	DLP + DM	−577.59	394.55	−1.46	0.143	0.211
	HT + DLP + DM	162.32	302.53	0.54	0.592	0.710
HT+DM	DLP + DM	−1119.08	414.99	−2.70	0.007	0.022
	HT + DLP + DM	−379.12	328.74	−1.15	0.249	0.348
DLP+DM	HT + DLP + DM	739.91	452.18	1.64	0.102	0.168

In addition, there was a statistically significant difference in the mean rank of WMH-BHQ between DLP and HT ($p = 0.001$), and a marginally significant difference between DLP and DM ($p = 0.098$). Furthermore, there were significant differences in the mean rank of WMH-BHQ between DLP and HT+DM ($p < 0.001$) and between HT+DM and DM+DLP ($p = 0.022$). These results show that DLP may positively affect WMH-BHQ, or at least, DLP appears unlikely to be a negative factor in this regard.

4. Discussion

Age was the strongest risk factor positively related to WMH on multiple regression analysis. In addition to age, SBP and BI were positive factors for WMH, which is consistent with prior reports [4,5]. Conversely, female sex, FBS, TG, and BMI were negative factors for WMH. The reason for sex differences in WMH remains unclear although the risk factors for LVD include male sex [22,23]. Recently, a population-based cognitively unimpaired cohort study with participants aged >70 years demonstrated that females had significantly greater WMHV than males [24]. Although our study examined the percentile of WMHV, the female susceptibility to WMH was significantly observed in participants aged <60 years but not in those aged >60 years (Table 3). The average menopausal age in Japan is around 50 years, but the differences across individuals are also regarded as substantial. Further study of the exact menopausal age is needed to elucidate the influence of sex hormones. If anything, sex differences according to age may contribute to the development of sex-specific preventive strategies against WMH progression linking to stroke and dementia.

Regarding fasting blood glucose and triglycerides, both β values were relatively low and the Kruskal–Wallis test showed no significant relationships according to IGF and ALM criteria. Fasting blood glucose and triglycerides may not be so heavily involved in WMH onset and development. Conversely, BMI was a powerful negative factor that increased the strength according to BMI criteria. BMI is an obesity index, whereas the others indicate waist circumstance (WC) and hip-waist ratio (HWR) [25]. Obesity directly depends on the amount of adipose tissue that can increase without weight gain. Obesity can be evaluated more accurately by WC or HWR than BMI, especially in the elderly [25]. Instead of BMI, WC was used for multiple regression analysis and yielded similar result as that with BMI, a negative factor of WMH. That obesity appears to suppress WMH may be because some growth factors for vessels are reportedly secreted from adipose tissue [26]. For example, angiopoietin-like protein 4 (ANGPTL4) is a member of the angiopoietin family, which encodes a secretory glycoprotein highly expressed in adipose tissue and liver and placenta [27]. ANGPTL4 and/or other vascular trophic factors are delivered from adipose tissue to the brain small vessels and may prevent WMH onset and progression. Further study will be needed to validate the hypothesis.

The existence and progression of WMH in the brain can be visually recognized through WMH-BHQ before or after VRCs such as HT, DM, and DLP emerge, that is, at any progression of morbidity from preclinical to chronic stages. The visible changes of WMH in the brain readily force patients to make efforts to control VRCs, linking to the prevention of stroke or dementia, and can determine when to start drug treatments or how to promote nondrug therapies, such as low calorie and salt diets and/or physical exercise. For example, HT diagnosis based on BP values remains unreflective of organ damages caused by a high BP. If WMH was accurately grasped by means of the new metric WMH-BHQ and efficiently treated according to the change in this measure, the damages due to HT in the brain and entire body could be minimized. Thus, WMH-BHQ may be regarded as an effective indicator that can help us to visualize brain damages and estimate the health of the whole body in terms of cerebral small vessel damages. Another advantage of WMH-BHQ is that it is based on percentile and ranking order, which enables minimization of measurement bias compared with WMHV, which heavily depends on scanning conditions and the abilities of MRI equipment.

Regarding comorbidity, HT+DM yielded lower WMH-BHQs than the other double morbidities (HT+DLP and DM+DLP) as well as triple morbidities (HT+DM+DLP), although this may be due to the possible ceiling effect of the WMH volume associated with the VRF. Multimorbidity is becoming a global challenge to prevent stroke and extend lifespan [28,29]. A nationally representative cross-sectional

study of more than 1.4 million persons in Scotland showed that a diagnosis of stroke significantly became more common as the number of morbidities increased [30]. Additional comorbidities are widely considered to decrease health with a destructive metabolic domino effect [29]. In our study, however, DLP may have a suppressive or preventive effect on WMH in comorbidity, although the Scotland study did not describe DLP at all. According to the ALM criteria, there were no significant differences in WMH-BHQ. This evidence might imply some effect of statins, usual drugs for DLP, rather than DLP pathology. Several meta-analyses of placebo-controlled randomized trials suggest that statins may be beneficial in reducing the overall incidence of stroke [31,32]. It remains to be determined whether statins prevent onset and development of stroke through suppression of WMH. In our study, the numbers of DLP and DM patients were extremely smaller than those with HT. Further validation needs a larger number of participants with DLP and DM for the next follow-up study. In the near future, MRI parameters could be assessed through artificial intelligence pivoting on data mining [33,34]. Such approach together with the identification of biomarkers based on novel nanotechnology or biomedical engineering platforms would allow to propose new biosignatures for risk stratification in neurovascular patients [35].

Limitations

The Brain Dock program utilizes an MRI-based approach to preventive medicine that was uniquely developed in Japan, aiming at early detection of unruptured cerebral aneurysms. At Brain Dock, health checks are conducted for a vast number of participants, and therefore, a large database could be built for brain research. Our study covered approximately 9000 participants living in Kochi Prefecture, Japan, and a single MRI machine was used throughout the study. Thus, the selection and information bias in this study could be minimized. However, our study included a bias of socioeconomic state involved, whereas one-fourth of the participants belonged to the white-collar class, such as public officials with moderate yearly incomes. The socioeconomic impact of this proportion is likely considered significant on WMH and the onset as well as the progression of VRCs to no small extent. The usage of 1.5 T MRI yields lower measurement of WMH volume as compared with 3 T MRI [36]. Nevertheless, the difference in magnetic power may be minimized in case of WMH-BHQ using the percentile of WMH volume. Our study was designed as a cross-sectional approach and only referred to the associations with WMHs at preclinical and chronic stages of VRCs. The next step will involve a prospective cohort study to certify the causal validation of WMH-BHQ using participants undergoing Brain Dock examinations more than twice.

5. Conclusions

In this study, we showed that cerebral white matter hyperintensities can be used as a healthcare quotient for quantitatively evaluating vascular risk factors or vascular risk conditions. Because of easiness of interpretation (Higher WMH-BHQ is better for brain in terms of cerebral vascular risk), WMH-BHQ might be useful for both clinicians and patients/inidividuals to pay their attentions to reduce cerebral vascular risks.

Author Contributions: Conceptualization, K.P. and K.N.; methodology, K.N. and K.Y.; software, F.Y.; validation, M.T., T.A. and M.S.; formal analysis, K.Y.; investigation, M.S.; resources, K.P.; data curation, K.P.; writing—original draft preparation, K.P.; writing—review and editing, K.N.; visualization, A.K.; supervision, M.S.; project administration, K.P.; funding acquisition, Y.Y.

Funding: This work was partially supported by the fund of ImPACT Program of Council for Science, Technology and Innovation (Cabinet Office, Government of Japan).

Acknowledgments: We thank Miho Kawai for assistance with questionnaires in our Brain Dock.

Conflicts of Interest: None of the authors has a financial relationship with a commercial entity that has an interest in the subject of this manuscript.

References

1. Pantoni, L. Cerebral small vessel disease: From pathogenesis and clinical characteristics to therapeutic challenges. *Lancet Neurol.* **2010**, *9*, 689–701. [CrossRef]
2. Debette, S.; Markus, H.S. The clinical importance of white matter hyperintensities on brain magnetic resonance imaging: Systematic review and meta-analysis. *BMJ* **2010**, *341*, c3666. [CrossRef] [PubMed]
3. Salat, D.H. Imaging small vessel-associated white matter changes in aging. *Neuroscience* **2014**, *276*, 174–186. [CrossRef] [PubMed]
4. Zeestraten, E.A.; Lawrence, A.J.; Lambert, C.; Benjamin, P.; Brookes, R.L.; Mackinnon, A.D.; Morris, R.G.; Barrick, T.R.; Markus, H.S. Change in multimodal MRI markers predicts dementia risk in cerebral small vessel disease. *Neurology* **2017**, *89*, 1869–1876. [CrossRef]
5. Ding, J.; Sigurðsson, S.; Jonson, P.V.; Eiriksdottir, G.; Charidimou, A.; Lopez, O.L.; van Buchem, M.A.; Guðnason, V.; Launer, L.J. Large perivascular spaces visible on magnetic resonance imaging, cerebral small vessel disease progression, and risk of dementia: The age, gene/environment susceptibility-reykjavik study. *JAMA Neurol.* **2017**, *74*, 1105–1112. [CrossRef]
6. Baradaran, H.; Al-Dasuqi, K.; Knight-Greenfield, A.; Giambrone, A.; Delgado, D.; Ebani, E.J.; Kamel, H.; Gupta, A. Association between Carotid plaque features on CTA and cerebrovascular ischemia: A systematic review and meta-analysis. *AJNR Am. J. Neuroradiol* **2017**, *38*, 2321–2326. [CrossRef]
7. Staals, J.; Makin, S.D.; Doubal, F.N.; Dennis, M.S.; Wardlaw, J.M. Stroke subtype, vascular risk factors, and total MRI brain small-vessel disease burden. *Neurology* **2014**, *83*, 1228–1234. [CrossRef]
8. Pantoni, L.; Garcia, J.H. Pathogenesis of leukoaraiosis: A review. *Stroke* **1997**, *28*, 652–659. [CrossRef]
9. Lin, J.; Wang, D.; Lan, L.; Fan, Y. Multiple factors involved in the pathogenesis of white matter lesions. *Biomed. Res. Int.* **2017**, *2017*, 1–9. [CrossRef]
10. Habes, M.; Erus, G.; Toledo, J.B.; Zhang, T.; Bryan, N.; Launer, L.J.; Rosseel, Y.; Janowitz, D.; Doshi, J.; Van der Auwera, S.; et al. White matter hyperintensities and imaging patterns of brain ageing in the general population. *Brain* **2016**, *139*, 1164–1179. [CrossRef]
11. Park, K.; Yasuda, N.; Toyonaga, S.; Yamada, S.M.; Nakabayashi, H.; Nakasato, M.; Nakagomi, T.; Tsubosaki, E.; Shimizu, K. Significant association between leukoaraiosis and metabolic syndrome in healthy subjects. *Neurology* **2007**, *69*, 974–978. [CrossRef] [PubMed]
12. Bokura, H.; Yamaguchi, S.; Iijima, K.; Nagai, A.; Oguro, H. Metabolic syndrome is associated with silent ischemic brain lesions. *Stroke* **2008**, *39*, 1607–1609. [CrossRef] [PubMed]
13. Assayag, E.B.; Eldor, R.; Korczyn, A.D.; Kliper, E.; Shenhar-Tsarfaty, S.; Tene, O.; Molad, J.; Shapira, I.; Berliner, S.; Volfson, V.; et al. Type 2 diabetes mellitus and impaired renal function are associated with brain alterations and poststroke cognitive decline. *Stroke* **2017**, *48*, 2368–2374. [CrossRef] [PubMed]
14. Cole, J.H.; Ritchie, S.J.; Bastin, M.E.; Valdés Hernández, M.C.; Muñoz Maniega, S.; Royle, N.; Corley, J.; Pattie, A.; Harris, S.E.; Zhang, Q.; et al. Brain age predicts mortality. *Mol. Psychiatry* **2018**, *23*, 1385–1392. [CrossRef] [PubMed]
15. Nemoto, K.; Oka, H.; Fukuda, H.; Yamakawa, Y. MRI-based brain healthcare quotients: A bridge between neural and behavioral analyses for keeping the brain healthy. *PLoS ONE* **2017**, *12*, e0187137. [CrossRef] [PubMed]
16. Park, K.; Nakagawa, Y.; Kumagai, Y.; Nagahara, M. Leukoaraiosis, a common brain magnetic resonance imaging finding, as a predictor of traffic crashes. *PLoS ONE* **2013**, *8*, e57255. [CrossRef] [PubMed]
17. Ashburner, J. SPM: A history. *Neuroimage* **2012**, *62*, 791–800. [CrossRef]
18. Dunn, O.J. Multiple comparisons among means. *J. Am. Stat. Assoc.* **1961**, *56*, 52–64. [CrossRef]
19. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **1995**, *57*, 289–300. [CrossRef]
20. Darlington, R. Multiple regression in psychological research and practice. *Psychol. Bull.* **1968**, *69*, 161–182. [CrossRef]
21. IBM Corp. *IBM SPSS Statistics for Windows*; Version 22.0; IBM Corp: Armonk, NY, USA, 2013.
22. Schulz, U.G.; Rothwell, P.M. Differences in vascular risk factors between etiology subtypes of ischemic stroke: Importance of population-based studies. *Stroke* **2003**, *34*, 2050–2059. [CrossRef] [PubMed]

23. Verhaaren, B.F.; Vernooij, M.W.; de Boer, R.; Hofman, A.; Niessen, W.J.; van der Lugt, A.; Ikram, M.A. High blood pressure and cerebral white matter lesion progression in the general population. *Hypertension* **2013**, *61*, 1354–1359. [CrossRef] [PubMed]
24. Fatemi, F.; Kantarci, K.; Graff-Radford, J.; Preboske, G.M.; Weigand, S.D.; Przybelski, S.A.; Knopman, D.S.; Machulda, M.M.; Roberts, R.O.; Mielke, M.M. Sex differences in cerebrovascular pathologies on FLAIR in cognitively unimpaired elderly. *Neurology* **2018**, *90*, e466–e473. [CrossRef] [PubMed]
25. World Health Organization. *Obesity and Overweight [Fact Sheet]*; World Health Organization: Geneva, Switzerland, 2011.
26. Corvera, S.; Gealekman, O. Adipose tissue angiogenesis: Impact on obesity and type-2 diabetes. *Biochim. Biophys. Acta* **2014**, *1842*, 463–472. [CrossRef] [PubMed]
27. La Paglia, L.; Listì, A.; Caruso, S.; Amodeo, V.; Passiglia, F.; Bazan, V.; Fanale, D. Potential role of ANGPTL4 in the cross talk between metabolism and cancer through PPAR signaling pathway. *PPAR Res.* **2017**, *17*, 8187–8235. [CrossRef] [PubMed]
28. Kabboord, A.D.; van Eijk, M.; Fiocco, M.; van Balen, R.; Achterberg, W.P. Assessment of comorbidity burden and its association with functional rehabilitation outcome after stroke or hip fracture: A systematic review and meta-analysis. *J. Am. Med. Dir. Assoc.* **2016**, *17*, 1066.e13–1066.e21. [CrossRef]
29. Hole, B.; Salem, J. How long do patients with chronic disease expect to live? A systematic review of the literature. *BMJ Open* **2016**, *6*, e012248. [CrossRef]
30. Gallacher, K.I.; Batty, G.D.; McLean, G.; Mercer, S.W.; Guthrie, B.; May, C.R.; Langhorne, P.; Mair, F.S. Stroke, multimorbidity and polypharmacy in a nationally representative sample of 1,424,378 patients in Scotland: Implications for treatment burden. *BMC Med.* **2014**, *12*, 151. [CrossRef]
31. Chróinín, D.N.; Asplund, K.; Åsberg, S.; Callaly, E.; Cuadrado-Godia, E.; Díez-Tejedor, E.; Di Napoli, M.; Engelter, S.T.; Furie, K.L.; Giannopoulos, S.; et al. Statin therapy and outcome after ischemic stroke: Systematic review and meta-analysis of observational studies and randomized trials. *Stroke* **2013**, *44*, 448–456. [CrossRef]
32. Choi, J.Y.; Seo, W.K.; Kang, S.H.; Jung, J.M.; Cho, K.H.; Yu, S.; Oh, K. Statins improve survival in patients with cardioembolic stroke. *Stroke* **2014**, *45*, 1849–1852. [CrossRef]
33. Garg, R.; Oh, E.; Naidech, A.; Kording, K.; Prabhakaran, S. Automating ischemic stroke subtype classification using machine learning and natural language processing. *J. Stroke Cerebrovasc. Dis.* **2019**, *28*, 2045–2051. [CrossRef] [PubMed]
34. Schirmer, M.D.; Dalca, A.V.; Sridharan, R.; Giese, A.K.; Donahue, K.L.; Nardin, M.J.; Mocking, S.J.T.; McIntosh, E.C.; Frid, P.; Wasselius, J.; et al. White matter hyperintensity quantification in large scale clinical acute ischemic stroke cohort. The MRI-GENIE study. *Neuroimage Clin.* **2019**, *23*, 101884. [CrossRef] [PubMed]
35. Ganau, L.; Prisco, L.; Ligarotti, G.K.I.; Ambu, R.; Ganau, M. Understanding the pathological basis of neurological diseases through diagnostic platforms based on innovations in biomedical engineering: New concepts and theranostic perspectives. *Medicines (Basel)* **2018**, *5*, 22. [CrossRef] [PubMed]
36. Perri1, C.D.; Dwyer, M.G.; Wack, D.S.; Cox, J.L.; Hashmi, K.; Saluste, E.; Hussein, S.; Dalaker, T.O.; Schirda, C.; Bastianello, S. White Matter Hyperintensities on 1.5 and 3 Tesla Brain MRI in Healthy Individuals. *J. Biomed. Gr. Comput.* **2013**. [CrossRef]

Journal of
Clinical Medicine

MDPI

Article

Factors Associated with Health-Related Quality of Life in Community-Dwelling Older Adults: A Multinomial Logistic Analysis

Encarnación Blanco-Reina [1,*], Jenifer Valdellós [2], Ricardo Ocaña-Riola [3,4], María Rosa García-Merino [5], Lorena Aguilar-Cano [6], Gabriel Ariza-Zafra [7] and Inmaculada Bellido-Estévez [1]

1 Pharmacology and Therapeutics Department, School of Medicine, Instituto de Investigación Biomédica de Málaga-IBIMA, University of Málaga, 29016 Málaga, Spain; ibellido@uma.es
2 Health District of Málaga-Guadalhorce, 29009 Málaga, Spain; jenny_dok7@hotmail.com
3 Escuela Andaluza de Salud Pública, 18011 Granada, Spain; ricardo.ocana.easp@juntadeandalucia.es
4 Instituto de Investigación Biosanitaria ibs.GRANADA, 18012 Granada, Spain
5 Health District of Córdoba Sur, 14940 Córdoba, Spain; rosaballet@yahoo.es
6 Physical Medicine and Rehabilitation Department, Hospital Regional Universitario, 29010 Málaga, Spain; loagca2011@hotmail.com
7 Geriatrics Department, Complejo Hospitalario Universitario, 02006 Albacete, Spain; gariza@sescam.jccm.es
* Correspondence: eblanco@uma.es; Tel.: +34-9521-36648; Fax: +34-952131568

Received: 19 September 2019; Accepted: 25 October 2019; Published: 1 November 2019

Abstract: The main aim of this study was to determine the association of various clinical, functional and pharmacological factors with the physical (PCS) and mental (MCS) summary components of the health-related quality of life (HRQoL) of community-dwelling older adults. Design: Cross-sectional study. Patients and setting: Sample of 573 persons aged over 65 years, recruited at 12 primary healthcare centres in Málaga, Spain. Sociodemographic, clinical, functional, and comprehensive drug therapy data were collected. The main outcome was HRQoL assessed on the basis of the SF-12 questionnaire. A multinomial logistic regression model was constructed to study the relationship between independent variables and the HRQoL variable, divided into intervals. The average self-perceived HRQoL score was 43.2 (± 11.02) for the PCS and 48.5 (± 11.04) for the MCS. The factors associated with a poorer PCS were dependence for the instrumental activities of daily living (IADL), higher body mass index (BMI), number of medications, and presence of osteoarticular pathology. Female gender and the presence of a psychopathological disorder were associated with worse scores for the MCS. The condition that was most strongly associated with a poorer HRQoL (in both components, PCS and MCS) was that of frailty (odds ratio (OR) = 37.42, 95% confidence interval (CI) = 8.96–156.22, and OR = 20.95, 95% CI = 7.55–58.17, respectively). It is important to identify the determinant factors of a diminished HRQoL, especially if they are preventable or modifiable.

Keywords: health-related quality of life; older adults; frailty; medication; primary care

1. Introduction

The aging of the population is a global phenomenon that is producing dramatic sociodemographic transformations. In this respect, a recent study modelled life expectancy, all-cause mortality and cause of death forecasts for 250 causes of death from 2016 to 2040 in 195 countries and territories. For 2040, Japan, Singapore, Spain and Switzerland were forecast to have an average life expectancy exceeding 85 years for both genders, and another 59 countries, including China, were projected to surpass a life expectancy of 80 years. According to these forecasts, Spain will then have the greatest life expectancy in the world (85.8 years) [1]. This pattern of aging poses a major challenge to health and

social assistance services, as older people present more chronic conditions and generate higher per capita healthcare costs. Multimorbidity is strongly associated with adverse health outcomes, such as disability, dependence, mortality, increased need for health and social services and polymedication, and diminished health-related quality of life (HRQoL) [2].

Increased longevity should not be achieved at the expense of quality of life. A poor HRQoL has been associated with reduced activities of daily living, a higher frequency of hospitalisation and increased mortality [3,4]. Therefore, enhancing HRQoL should be a major consideration in the design and implementation of health care for older persons [5]. In addition, assigning a greater importance to the quality of life, in preference to disease-based outcomes, is consistent with the opinions expressed by such persons themselves [6,7]. As a consequence of the interest being generated by this topic, patient-perceived quality of life is now widely used as a measure of health care in clinical research and in health economics assessments [8].

HRQoL has been defined as "the subjective perception influenced by the current health status of the ability to perform activities important for the person" [9,10]. In response to heightened interest in assessing HRQoL and in determining the effectiveness of health care interventions in older patients, a range of instruments—both generic and specific—have been developed for these purposes.

Many factors can influence the HRQoL of older adults, including health status, social engagement and cognitive function [11]. In this area, studies have been undertaken to analyse the association between factors (such as female sex, age, functional impairment and comorbidities) and HRQoL, and in recent years, the role of medication as a determinant of overall health outcomes has emerged as an area of great interest. Various aspects of treatment regimens call for special attention, because they may have a negative influence on the HRQoL of older persons, such as polymedication, and potentially inappropriate or harmful medication [12–17]. It is important to recall that the benefits of medication should always outweigh the potential harm, and therefore individual circumstances should be taken into account. In recent studies, frailty, too, has been highlighted as a possible determinant of worsened quality of life [18–22]. Frailty is defined as an abnormal health state characterised by the loss of biological reserves, related to the aging process [23]. Therefore, in providing care for persons with frailty, special emphasis should be placed on quality of life considerations.

Evaluation of HRQoL can help clinicians to determine the needs of older patients and thus optimise their decision making. Therefore, we believe it is important to explore factors influencing HRQoL in this population in order to identify suitable intervention strategies. In this respect, we note that many previous studies have not addressed all the factors that might have a significant impact on HRQoL, and in some cases the results presented are contradictory. For all of these reasons, we consider it interesting to investigate HRQoL in older adults, resident in one of the countries where life expectancy is highest, to assess clinical, functional and pharmacological aspects of the question. In this study, we hypothesise that there are multiple predictive factors of a poor HRQoL and that they may affect its physical and mental components in different ways and to different degrees. Accordingly, our aim was to determine levels of HRQoL among community-dwelling older people and to analyse the factors associated with a poor HRQoL.

2. Patients and Methods

2.1. Study Design, Setting and Participants

In this cross-sectional investigation, the study population was composed of 89,615 community-dwelling residents aged 65 years or more, living in Málaga, Spain. Assuming a standard deviation for HRQoL of 11.0 [24], an absolute precision of $\delta = 0.9$ and a level of confidence of $1-\alpha = 0.95$, we calculated that the minimum sample size needed to estimate the average physical and mental HRQoL was 570. Finally, the total sample was composed of 573 persons. These patients were recruited from twelve primary care centres, by stratified random sampling designed to obtain a representative sample of the population, allocating the population in proportion to the size of each healthcare centre.

Participants were selected randomly within each healthcare centre from a general list of healthcare cards issued by the Spanish National Health Service. The inclusion criteria were people 65 years of age or older, included in the database of healthcare cards, belonging to the outpatient setting (not institutionalized), and giving their informed consent to participate in the study (people who did not give consent were excluded).

2.2. Data Collection and Measures

To obtain the study data for analysis, patients were interviewed using a structured questionnaire. Further data were obtained from medication packaging and digital medical records. The questionnaire was used to obtain detailed information on the patients' regular drug use, together with clinical, functional and sociodemographic data. Clinical diagnoses were examined, and the number of chronic conditions presented by each participant was determined. Patients' independence in performing instrumental activities of daily living (IADL) was assessed using the Lawton scale [25]. In addition, the body mass index was determined for all patients, and frailty was assessed according to Fried's criteria (as robust, pre-frail or frail) [26].

Medication assessment. Data were obtained for the medication prescribed (indication, dosage and duration of treatments during the last three months or more). The presence of polymedication was considered, defining this as the regular use of five or more medications, as was that of potentially inappropriate medication (PIM), according to the STOPP v2 criteria (Screening Tool of Older Person's Potentially Inappropriate Prescriptions, version 2) [27]. The latter variable was operationalised as the percentage of patients receiving at least one PIM.

Quality of life assessment. The main study outcome was HRQoL assessed by the SF-12 questionnaire, a widely used generic instrument. The SF-12 is an abbreviated version of the Short Form-36 Health Survey (SF-36), in which a subset of 12 items/questions are used to derive summary scores for physical health (PCS score) and mental health (MCS score) [28]. The response options form Likert-type scales that assess the intensity and/or frequency of people's health status. The final score obtained can be between 0 and 100, where lower scores indicate worse, and higher scores better HRQoL. Using only one-third of the SF-36 items, the SF-12 reproduces the two summary scores originally developed for the SF-36 with remarkable accuracy, but more quickly and requiring less effort from the respondent [29]. The SF-12 has been validated for use in the USA, the UK, Spain and many other European countries [24].

2.3. Statistical Analysis

Exploratory data analysis and frequency tables were used to describe the study variables. Using the standardised scores of the SF-12 questionnaire for the Spanish population [24], the HRQoL score of each participant was assigned to one of the following intervals, taking into account the population reference group according to age and sex:

- Very low: QoL ≤ 20th percentile of the Spanish population corresponding to their age group and sex;
- Low: QoL > 20th percentile and ≤50th percentile of the Spanish population corresponding to their age group and sex;
- High: QoL > 50th percentile and ≤80th percentile of the Spanish population corresponding to their age group and sex;
- Very high: QoL > 80th percentile of the Spanish population corresponding to their age group and sex.

A multinomial logistic regression model was used to study the relationship between the independent variables and the HRQoL variable, grouped into the corresponding intervals [30]. A 5% significance level was assumed to indicate statistical significance. Statistical data analysis was performed using SPSS version 23.0 (IBM SPSS Statistics, Armonk, NY, USA).

2.4. Ethical Considerations

This study was conducted in accordance with the provisions of the 1975 Declaration of Helsinki, revised in 2013. The Málaga Clinical Research Ethics Committee approved the study (PI-0234-14), and informed consent was obtained from all patients prior to their inclusion.

3. Results

3.1. Characteristics of the Study Population

The study sample was composed of 573 patients, with a mean age of 73.1 years (standard deviation 5.5, range 65–104) and of whom 57.2% were female. Most lived with their partner (62.4%) or family (16.5%), but 21.1% lived alone. On average, each patient presented 7.8 chronic conditions (standard deviation 3.3, range 0–20). The most prevalent diagnoses were bone and joint disorders (mainly osteoarthritis of the knee, hip, hand and shoulder) (75.2%), hypertension (70.5%) and dyslipidaemia (51.6%). Most of the patients presented with overweight (40.9%) or obesity (45.7%) and their mean body mass index was 30.2 (standard deviation 5.1, range 17–54.5). Half were independently capable of performing IADL, and the mean score on the Lawton scale was 6.6 (standard deviation 1.8, range 0–8). Frailty was present in 137 patients (23.9%; 95% confidence interval (CI) = 20.5–27.5), according to Fried's criteria. The main characteristics of the study population are detailed in Table 1.

The prevalence of polymedication was 68% (95% CI = 64.1–71.7), and on average each patient consumed 6.8 drugs (standard deviation 4.0; range 0–23). The most widely prescribed drugs were omeprazole and acetaminophen, followed by aspirin, simvastatin, metformin, metamizole, enalapril and bromazepam. The use of potentially inappropriate medication, according to the STOPP v2 criteria, was identified in 66.8% of patients (95% CI = 62.9–70.6). The number of PIMs per patient ranged from 0–10 (mean 2.1, standard deviation 2.2). The most frequent PIMs detected were benzodiazepines (61% of all PIMs).

3.2. Assessment of HRQoL and Analysis of Related Factors

The patients' perceptions of their HRQoL produced an average score of 43.2 (standard deviation 11.02, range 16.2–65.4) for the physical component summary (PCS) and a somewhat higher one, 48.5 (standard deviation 11.04, range 14.1–66.6), for the mental component (MCS). Males obtained higher scores than females in both the PCS and the MCS, with average values of 45.91 vs. 41.27 and 51.95 vs. 45.91, respectively. Table 2 shows the distribution within the global sample among the categories of perceived HRQoL (very low, low, high and very high). A notable feature of this distribution is that a higher proportion of patients perceived their HRQoL as very high in the mental component than in the physical one.

To further examine the impact of the independent variables on the HRQoL categories, a multinomial logistic regression analysis was performed (Tables 3–5). The factors related to having high vs. very high HRQoL (P_{50}–P_{80} and $\geq P_{80}$, respectively) were the level of dependence for IADL in the PCS, and BMI, respiratory disease and frailty in the MCS (Table 3). The presence of a low HRQoL (P_{20}–P_{50}) with respect to a very high score for the PCS was associated with the level of dependence for IADL, with accompanied living, with the presence of osteoarticular pathology and with frailty (Table 4). The odds of these older persons having a low HRQoL decrease by 30% for each additional point of independence on the Lawton scale (odds ratio (OR) = 0.70, 95% CI = 0.55–0.88). However, they double for those who do not live alone (OR = 2.13, 95% CI = 1.07–4.27), are 2.5 times greater for those with osteoarticular pathology compared to those without it (OR = 2.57, 95% CI = 1.35–4.85) and are seven times greater for those who are frail (OR = 7.43, 95% CI = 2.13–25.82) compared to those who are robust. In the case of the MCS, the presence of frailty was associated with a three times greater perception of low HRQoL (OR = 3.2, 95% CI = 1.21–8.46).

Table 1. Characteristics of study population (n = 573).

Quantitative Variables	Mean	Standard Deviation
Age (years)	73.1	5.5
Lawton (IADL)	6.6	1.8
BMI (Kg/m^2)	30.2	5.1
Number of comorbidities	7.8	3.3
Number of drugs per patient	6.8	4.0
Number of PIMs per patient	2.1	2.2
Qualitative Variables	**Subjects**	**Percentage**
Gender		
Male	245	42.8
Female	328	57.2
Living Arrangements		
Living alone	121	21.1
Accompanied living	452	78.9
Frailty		
Robust	124	21.6
Pre-frail	312	54.5
Frail	137	23.9
Most Frequent Comorbidities		
Bone and joint disorders	431	75.2
Hypertension	404	70.5
Dyslipidaemia	292	51.0
Insomnia	254	44.3
Gastrointestinal disease	241	42.0
Peripheral vascular disease	227	39.6
Psychopathology	207	36.1
Diabetes mellitus	172	30.0
Heart disease	139	24.3
Respiratory Disease	123	21.5
Polymedication	390	68.0
PIM prevalence	383	66.8

IADL: Instrumental Activities of Daily Living; BMI: Body Mass Index; PIM: Potentially Inappropriate Medication (according to STOPP v2 criteria).

Table 2. HRQoL assessment according to SF-12 questionnaire.

HRQoL Component	HRQoL Categories, n (%)			
Summary Score	**Very low ($\leq P_{20}$)**	**Low (P_{20}–P_{50})**	**High (P_{50}–P_{80})**	**Very high ($\geq P_{80}$)**
PCS	137 (23.9%)	170 (29.7%)	150 (26.2%)	116 (20.2%)
MCS	143 (25.0%)	117 (20.4%)	125 (21.8%)	188 (32.8%)

PCS: Physical Component Summary; MCS: Mental Component Summary; P: Percentile.

A very low perception of HRQoL ($\leq P_{20}$) in the PCS was also related to the level of dependence for IADL (OR = 0.62, 95% CI = 0.48–0.79), with the presence of osteoarticular pathology (OR = 4.38, 95% CI = 1.98–9.70) and with the states of prefrailty (OR = 4.19, 95% CI = 1.61–10.86) and frailty (OR = 37.42, 95% CI = 8.96–156.22). In addition, the odds of a very low HRQoL increase by 15% for each additional drug in the treatment regimen (OR = 1.15, 95% CI = 1.02–1.30) and by 8% for each unit increase in BMI (OR = 1.08, 95% CI = 1.01–1.15). However, age was inversely related to HRQoL. Thus, all other variables being equal, for each additional year of life, the odds of the patient perceiving a very low HRQoL decreased by 8% (Table 5). A similar pattern for the age effect was observed with respect to the mental component of HRQoL; thus, each additional year of life decreased the odds of a very low HRQoL by 7%. In the MCS, too, the states of prefrailty and frailty were associated with a very low

HRQoL. In this mental component, and with all other variables being equal, the female patients were 88% more likely than the males to have a very low HRQoL. Finally, the presence of a psychopathology (usually anxiety and/or depression) was related to a poorer HRQoL; the presence of any such disorder quadrupled the odds of the patient perceiving a very low HRQoL (OR = 4.69, 95% CI = 2.60–8.47).

Table 3. Factors related to HRQoL. Multinomial logistic regression for High HRQoL (with respect to Very High HRQoL) for physical and mental health components (PCS-MCS).

Independent Variable	PCS OR (95% CI)	MCS OR (95% CI)
Age	1.02 (0.96–1.07)	1.01 (0.95–1.05)
No. of comorbidities	1.06 (0.93–1.20)	1.02 (0.91–1.14)
BMI	0.99 (0.94–1.05)	0.94 (0.89–0.99) *
Independence (IADL)	0.74 (0.58–0.94) *	1.17 (0.98–1.39)
No. of medications	1.07 (0.96–1.19)	1.00 (0.91–1.10)
No. of PIMs	1.10 (0.85–1.43)	1.08 (0.87–1.34)
Female gender (ref. male)	0.72 (0.38–1.34)	1.45 (0.81–2.59)
Living accompanied (ref. alone)	1.26 (0.67–2.38)	1.14 (0.61–2.14)
Bone and joint disease	1.40 (0.79–2.49)	0.94 (0.53–1.66)
Heart disease	1.35 (0.62–2.92)	0.81 (0.42–1.54)
Respiratory disease	0.71 (0.33–1.53)	2.05 (1.10–3.84) *
Hypertension	1.20 (0.67–2.15)	1.20 (0.67–2.14)
Diabetes mellitus	1.22 (0.63–2.37)	0.76 (0.42–1.36)
Psychopathology	1.11 (0.58–2.12)	1.65 (0.91–2.98)
Insomnia	1.18 (0.64–2.16)	1.26 (0.73–2.16)
Pre–frail (ref. robust)	0.87 (0.48–1.57)	0.95 (0.53–1.69)
Frail (ref. robust)	0.46 (0.11–1.81)	3.39 (1.35–8.51) **

PIMs: Potentially Inappropriate Medications (according to STOPP v2 criteria). * $p < 0.05$; ** $p < 0.01$.

Table 4. Factors related to HRQoL. Multinomial Logistic Regression for Low HRQoL (with respect to Very High HRQoL) for physical and mental health components (PCS-MCS).

Independent Variable	PCS OR (95% CI)	MCS OR (95% CI)
Age	0.97 (0.92–1.03)	0.98 (0.93–1.03)
No. of comorbidities	1.08 (0.95–1.23)	1.11 (0.98–1.24)
BMI	1.04 (0.98–1.10)	0.95 (0.90–0.99) *
Independence (IADL)	0.70 (0.55–0.88) **	0.99 (0.84–1.18)
No. of medications	1.11 (0.99–1.24)	0.98 (0.89–1.08)
No. of PIMs	0.94 (0.72–1.22)	1.12 (0.89–1.39)
Female gender (ref. male)	1.01 (0.52–1.91)	1.67 (0.93–3.00)
Living accompanied (ref. alone)	2.13 (1.07–4.27) *	0.84 (0.45–1.57)
Bone and joint disease	2.57 (1.35–4.85) **	0.96 (0.53–1.72)
Heart disease	1.19 (0.54–2.62)	0.85 (0.45–1.62)
Respiratory disease	1.32 (0.64–2.71)	1.54 (0.81–2.92)
Hypertension	1.75 (0.94–3.27)	0.73 (0.41–1.31)
Diabetes mellitus	0.96 (0.48–1.92)	0.75 (0.42–1.36)
Psychopathology	1.20 (0.62–2.30)	1.36 (0.74–2.50)
Insomnia	1.33 (0.71–2.47)	1.01 (0.58–1.74)
Pre–frail (ref. robust)	1.48 (0.77–2.87)	1.50 (0.81–2.78)
Frail (ref. robust)	7.43 (2.13–25.82) **	3.20 (1.21–8.46) **

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

Table 5. Factors related to HRQoL. Multinomial Logistic Regression for Very Low HRQoL (with respect to Very High HRQoL) for physical and mental health components (PCS-MCS).

Independent Variable	PCS OR (95% CI)	MCS OR (95% CI)
Age	0.92 (0.86–0.98) *	0.93 (0.88–0.98) *
No. of comorbidities	1.11 (0.96–1.28)	0.98 (0.87–1.10)
BMI	1.08 (1.01–1.15) *	0.99 (0.95–1.05)
Independence (IADL)	0.62 (0.48–0.79) ***	0.94 (0.79–1.12)
No. of medications	1.15 (1.02–1.30) *	0.99 (0.90–1.10)
No. of PIMs	0.94 (0.70–1.24)	1.08 (0.87–1.35)
Female gender (ref. male)	1.58 (0.76–3.29)	1.88 (1.01–3.49) *
Living accompanied (ref. alone)	1.63 (0.76–3.50)	0.63 (0.33–1.17)
Bone and joint disease	4.38 (1.98–9.70) **	1.14 (0.60–2.18)
Heart disease	0.87 (0.36–2.08)	0.74 (0.37–1.49)
Respiratory disease	1.01 (0.45–2.22)	0.94 (0.47–1.89)
Hypertension	1.50 (0.73–3.05)	0.93 (0.51–1.72)
Diabetes mellitus	1.11 (0.52–2.36)	0.88 (0.47–1.63)
Psychopathology	1.53 (0.75–3.11)	4.69 (2.60–8.47) ***
Insomnia	1.24 (0.62–2.46)	1.66 (0.94–2.91)
Pre-frail (ref. robust)	4.19 (1.61–10.86) **	2.92 (1.36–6.26) **
Frail (ref. robust)	37.42 (8.96–156.22) ***	20.95 (7.55–58.17) ***

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

4. Discussion

Overall, our results for perceived HRQoL among community-dwelling older patients in Malaga (southern Spain) are consistent with those reported in similar studies conducted in other regions of Spain [31,32] and elsewhere in Europe [33]. With respect to the components of physical and mental health, the average PCS score (43.2) is slightly higher than that obtained in previous research [31–34], while the MCS (48.5) is somewhat lower than the average value for six European countries (54.3), also obtained using the SF-12 [33]. This lower score in the MCS could be due to cultural differences, and/or the non-negligible prevalence of psychopathological disorders (mainly anxiety and/or depression) in our region (these pathologies affect 36% of the older population). Regarding the influence of age on HRQoL, previous studies have produced conflicting results [32–35]. Our investigation revealed an inverse relationship between advanced age and the odds of a very low HRQoL. Therefore, other questions such as comorbidities, frailty, dependence and other possible confounders being equal, as the patient ages there is a lower probability of him/her perceiving a very low HRQoL, and this is true for both the physical and the mental components. These findings might be explained in terms of a better psychological adaptation to aging and perhaps to the fact that with age comes not only wisdom, but also the attribution of greater meaning or value to life. As concerns the influence of gender, previous studies have reported a poorer HRQoL among women than among men [31–34,36,37], and our own results corroborate this difference. Thus, women obtained poorer scores in both components of HRQoL, although the association was only significant for the mental component of a very low quality of life. In the physical component, female gender is probably not a determinant factor of a poorer HRQoL due to the adjustment made for confounding variables such as frailty or osteoarticular pathology, both of which are more prevalent among women.

One of the most consistent predictors of a poor quality of life is the level of dependence for IADL. According to the multinomial logistic regression model, with a higher score on the Lawton scale, i.e., with greater independence, the possibility of a poorer HRQoL decreased significantly (in all the categories considered: high, low and very low, with respect to very high) always with respect to the physical component. It seems logical that the higher the degree of independence among older persons, in activities such as handling economic affairs, using the telephone or managing different forms of transport, the better the quality of life they will perceive. According to a recent study,

functional dependence, together with the presence of depressive symptoms, would be an important factor mediating the well-known association between multimorbidity and a poor quality of life [38]. In this line, too, it has been shown that disability is one of the most significant conditions worsening the HRQoL [39]. Older adults who live in cohabitation are twice as likely to have a low physical HRQoL than are those living alone. This finding may be related to the level of functional dependence presented, but further study is needed to confirm this possibility.

Regarding medication, the number of drugs prescribed was positively associated with a very low HRQoL in the PCS. This finding is consistent with previous research, in which polymedication has been identified as a determinant factor of a poorer quality of life [12–14,40]. In our opinion, this association may be related to the side effects produced by the joint presence of several drugs within the patient. On the other hand, we found no evidence of any association between having at least one PIM and the perception of a poorer HRQoL. We speculate that such a relationship might not have been demonstrated because the STOPP v2 criteria contain a large number of items, of varying clinical significance, and therefore the impact produced on HRQoL by a single PIM might be slight. In other words, the sensitivity of this means of measuring the risk might be insufficient. Other authors, too, have failed to observe any significant association between the presence of a PIM and HRQoL, whether using the Beers [16] or the STOPP v2 criteria [41]. On the other hand, other previous studies do suggest that the prescription of drugs with a high anticholinergic load may be associated with a diminished quality of life [15–17,42].

Among this population, the most prevalent pathologies observed were osteoarthritis and hypertension. This is a common profile and similar to that reported in previous research in this field [39]. Our assessment of the impact on HRQoL of different conditions shows that physical disorders mostly affected the "physical" HRQoL while mental disorders mainly affected the "mental" aspect—as is only logical. The presence of bone and joint disease was significantly associated with a low or very low HRQoL, which corroborates previous findings in which this diagnosis was associated both with disability and with a diminished HRQoL [39]. We believe this relationship may be explained by the chronic pain which often accompanies these diseases, as well as a degree of physical limitation that is usually provoked. This association contrasts with its absence for other diagnoses, perhaps with a worse prognosis but of a more 'silent' nature, such as hypertension and diabetes mellitus. On the other hand, the presence of a psychopathology in an older person raises the possibility of his/her perceiving a very low HRQoL, in the mental component, by 400%, as has also been indicated in previous studies [12,31,32,43,44]. It would seem that the distress generated by these diseases has a negative impact on emotional regulation, motivation and other components of subjective perceptions of health and well-being. According to other researchers, and taking into account the significant impact on the patient's quality of life, we believe more screening for depression and anxiety among older populations should be performed, because these conditions tend to be under-diagnosed and also because a more active approach to this question would promote healthy aging [31]. In our sample, the BMI values observed presented two interesting aspects: on the one hand, at the mental level, a higher BMI was associated with a better HRQoL, but for the physical component it was significantly associated with a very low quality of life. It seems clear that overweight and obese patients perceive a poorer physical health [40,45–47]. This is corroborated by the fact that interventions aimed at achieving weight loss have been shown to improve the physical quality of life [48].

It is interesting to note that the factor most strongly associated with a diminished HRQoL was that of frailty, which severely affected both the MCS and the PCS, but especially the latter. The odds of a very low physical quality of life were 37 times greater among frail older persons than among those who were robust, while the MCS reflected 20 times greater odds of their presenting a poor mental quality of life. Previous studies, too, have reported this association [18–21,49], which was reinforced in a recent systematic review and meta-analysis that described it as clear and often substantial [22]. Although the growing numbers of frail older people pose a real challenge to health systems around the world, this state of pre-discapacity can in fact be prevented and treated. Furthermore, in our study sample,

the pre-frail persons also perceived a poorer quality of life than those who were relatively robust. Accordingly, we believe it necessary to design and incorporate care programmes specifically adapted to this emerging population of frail and pre-frail patients, to help them age with a better quality of life.

In our opinion, the type of study described in this paper is useful for identifying the characteristics and clinical conditions that are associated with a poorer HRQoL, with particular attention to those factors that may be preventable or treatable. As improving the quality of life and well-being of older persons is a priority objective, it is of major importance to extend our knowledge of these questions.

The strengths of our study lie in the analysis made of a representative sample of healthcare centres, the global approach taken, and the great variety of clinical, functional and treatment data compiled. We acknowledge that selecting a sample population from a single region or country may result in a certain lack of external validity. Nevertheless, the sample examined in this study may be representative of the population of older adults in the ambulatory setting, which is where the largest number of such patients are to be found. Another limitation to our study is its cross-sectional design, which does not allow causal relationships to be established, although it can detect factors related to HRQoL.

With regard to the analysis performed, since HRQoL is a quantitative variable, a multivariate linear regression model might, in principle, be considered appropriate to identify the factors related to it. However, in the present case the linear model considered did not meet the conditions of homoscedasticity and normality of the residuals necessary for the correct estimation of the study parameters. Previous statistical research has reached similar conclusions, i.e., that HRQoL cannot be analysed using linear regression models [50]. Neither nonlinear modelling, nor generalised least squares nor other more complex statistical techniques overcame the problem of achieving an adequate fit, and therefore we adopted the solution of treating QoL as a qualitative variable and using a multinomial logistic regression model. This approach provided a good fit to the data considered.

5. Conclusions

In conclusion, many factors may be predictive of a poor HRQoL and they affect its physical and mental components in different ways. For the PCS, the associated factors were dependence for IADL, a higher BMI, the number of medications and the presence of osteoarticular pathology. The main factors associated with a lower MCS score were female gender and the presence of a psychopathological disorder. Some factors may be preventable or modifiable, and so recognising them and optimising the response made are crucial to the priority objective of enhancing the quality of life among older people. Clearly and consistently, the factor that was most strongly associated with a poorer overall HRQoL was the state of frailty (and also, albeit to a lesser extent, that of pre-frailty). Frailty is a dynamic syndrome, in which transitions are possible between the states of normality, pre-frailty and frailty. Accordingly, detecting frailty and addressing it in a suitable way are questions of major importance.

Author Contributions: Conceptualization, E.B.-R., R.O.-R., G.A.-Z.; Methodology, E.B.-R., R.O.-R.; Formal analysis, E.B.-R., R.O.-R.; Investigation, E.B.-R., R.O.-R., G.A.-Z., J.V., L.A.-C., M.R.G.-M., I.B.-E.; Resources, E.B.-R., I.B.-E.; Data curation: J.V., L.A.-C., M.R.G.-M.; Writing—original draft preparation, E.B.-R., R.O.-R., J.V.; Writing—review & editing, E.B.-R., R.O.-R., J.V., I.B.-E.; Visualization, E.B.-R., R.O.-R., I.B.-E.; Supervision, E.B.-R., I.B.-E.; Project administration, E.B.-R.; Funding acquisition: E.B.-R.

Funding: This research has been supported by grant funding provided by the Fundación Pública Andaluza Progreso y Salud, Consejería de Salud, Junta de Andalucía, through the Programme Proyectos de Investigación Biomédica (Grant number PI 0234/14). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments: The authors wish to thank the Primary Care Management Team (Health District of Málaga) for providing access to the health centres and patient lists.

Conflicts of Interest: All authors declare that they have no conflict of interest.

References

1. Foreman, K.J.; Márquez, N.; Dolgert, A.; Fukutaki, L.; Fullman, N.; McGaughey, M.; Pletcher, M.A.; Smith, A.E.; Tang, K.; Yuan, C.W.; et al. Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: Reference and alternative scenarios for 2016-40 for 195 countries and territories. *Lancet* **2018**, *392*, 2052–2090. [CrossRef]
2. Marengoni, A.; Angleman, S.; Melis, R.; Mangialasche, F.; Karp, A.; Garmen, A.; Meinow, B.; Fratiglioni, L. Aging with multimorbidity: A systematic review of the literatura. *Ageing Res. Rev.* **2011**, *10*, 430–439. [CrossRef] [PubMed]
3. Lee, D.T.F.; Yu, D.S.F.; Kwong, A.N.L. Quality of life of older people in residential care home: A literature review. *J. Nurs. Healthc Chronic Illn.* **2009**, *1*, 116–125. [CrossRef]
4. Otero-Rodríguez, A.; León-Muñoz, L.M.; Balboa-Castillo, T.; Banegas, J.R.; Rodríguez-Artalejo, F.; Guallar-Castillón, P. Change in health-related quality of life as a predictor of mortality in the older adults. *Qual. Life Res.* **2010**, *19*, 15–23. [CrossRef] [PubMed]
5. World Health Organization. World Report on Ageing and Health. Geneva: World Health Organization. 2015. Available online: https://www.who.int/ageing/events/world-report-2015-launch/en/ (accessed on 29 July 2019).
6. Gabriel, Z.; Bowling, A. Quality of life from the perspectives of older people. *Ageing Soc.* **2004**, *24*, 675–691. [CrossRef]
7. National Voices, Age, U. K. & UCL Partners. I'm Still Me: A Narrative for Coordinated Support for Older People: National Voices. 2014. Available online: https://www.nationalvoices.org.uk/sites/default/files/public/publications/im_still_me.pdf (accessed on 29 July 2019).
8. Makai, P.; Brouwer, W.B.; Koopmanschap, M.A.; Stolk, E.A.; Nieboer, A.P. Quality of life instruments for economic evaluations in health and social care for older people: A systematic review. *Soc. Sci. Med* **2014**, *102*, 83–93. [CrossRef] [PubMed]
9. Group WHOQoL. Study protocol for the World Health Organization project to develop a Quality of Life assessment instrument (WHOQOL). *Qual. Life Res.* **1993**, *2*, 153–159. [CrossRef]
10. Shumaker, S.A.; Berzon, R.A. *The International Assessment of Health-Related Quality of Life: Theory, Translation, Measurement, and Analysis*; Rapid Communications of Oxford Ltd.: Oxford, UK, 1995.
11. Borowiak, E.; Kostka, T. Predictors of quality of life in older people living at home and in institutions. *Aging Clin. Exp. Res.* **2004**, *16*, 212–220. [CrossRef]
12. Montiel-Luque, A.; Núñez-Montenegro, A.J.; Martín-Aurioles, E.; Canca-Sánchez, J.C.; Toro-Toro, M.C.; González-Correa, J.A.; Polipresact Research Group. Medication-related factors associated with health-related quality of life in patients older than 65 years with polypharmacy. *PLoS ONE.* **2017**, *11*, e0171320. [CrossRef]
13. Schenker, Y.; Park, S.Y.; Jeong, K.; Pruskowski, J.; Kavalieratos, D.; Resick, J.; Abernethy, A.; Kutner, J.S. Associations Between Polypharmacy, Symptom Burden, and Quality of Life in Patients with Advanced, Life-Limiting Illness. *J. Gen. Intern. Med.* **2019**, *34*, 559–566. [CrossRef]
14. Tegegn, H.G.; Erku, D.A.; Sebsibe, G.; Gizaw, B.; Seifu, D.; Tigabe, M.; Belachew, S.A.; Ayele, A.A. Medication-related quality of life among Ethiopian elderly patients with polypharmacy: A cross-sectional study in an Ethiopia university hospital. *PLoS ONE.* **2019**, *14*, e0214191. [CrossRef] [PubMed]
15. Juola, A.L.; Pylkkanen, S.; Kautiainen, H.; Bell, J.S.; Bjorkman, M.P.; Finne-Soveri, H.; Soini, H.; Pitkälä, K.H. Burden of Potentially Harmful Medications and the Association with Quality of Life and Mortality Among Institutionalized Older People. *J. Am. Med. Dir. Assoc.* **2016**, *17*, e9–e14. [CrossRef] [PubMed]
16. Ie, K.; Chou, E.; Boyce, R.D.; Albert, S.M. Potentially Harmful Medication Use and Decline in Health-Related Quality of Life among Community-Dwelling Older Adults. *Drugs Real World Outcomes.* **2017**, *4*, 257–264. [CrossRef] [PubMed]
17. Harrison, S.L.; Kouladjian O'Donnell, L.; Bradley, C.E.; Milte, R.; Dyer, S.M.; Gnanamanickam, E.S.; Liu, E.; Hilmer, S.N.; Crotty, M. Associations between the Drug Burden Index, Potentially Inappropriate Medications and Quality of Life in Residential Aged Care. *Drugs Aging* **2018**, *35*, 83–91. [CrossRef]
18. Rivera-Almaraz, A.; Manrique-Espinoza, B.; Ávila-Funes, J.A.; Chatterji, S.; Naidoo, N.; Kowal, P.; Salinas-Rodríguez, A. Disability, quality of life and all-cause mortality in older Mexican adults: Association with multimorbidity and frailty. *BMC Geriatr.* **2018**, *18*, 236. [CrossRef]

19. Godin, J.; Armstrong, J.J.; Wallace, L.; Rockwood, K.; Andrew, M.K. The impact of frailty and cognitive impairment on quality of life: Employment and social context matter. *Int. Psychogeriatr.* **2018**, *13*, 1–9. [CrossRef]
20. Kirkhus, L.; Šaltytė Benth, J.; Grønberg, B.H.; Hjermstad, M.J.; Rostoft, S.; Harneshaug, M.; Selbæk, G.; Wyller, T.B.; Jordhøy, M.S. Frailty identified by geriatric assessment is associated with poor functioning, high symptom burden and increased risk of physical decline in older cancer patients: Prospective observational study. *Palliat. Med.* **2019**, *33*, 312–322. [CrossRef]
21. Siriwardhana, D.D.; Weerasinghe, M.C.; Rait, G.; Scholes, S.; Walters, K.R. The association between frailty and quality of life among rural community-dwelling older adults in Kegalle district of Sri Lanka: A cross-sectional study. *Qual. Life Res.* **2019**, *28*, 2057–2068. [CrossRef]
22. Crocker, T.F.; Brown, L.; Clegg, A.; Farley, K.; Franklin, M.; Simpkins, S.; Young, J. Quality of life is substantially worse for community-dwelling older people living with frailty: Systematic review and meta-analysis. *Qual. Life Res.* **2019**, *28*, 2041–2056. [CrossRef]
23. Clegg, A.; Young, J.; Iliffe, S.; Rikkert, M.O.; Rockwood, K. Frailty in elderly people. *Lancet.* **2013**, *381*, 752–762. [CrossRef]
24. Vilagut, G.; Valderas, J.M.; Ferrer, M.; Garin, O.; López-García, E.; Alonso, J. Interpretation of SF-36 and SF-12 questionnaires in Spain: Physical and mental components. *Med. Clin.* **2008**, *130*, 726–735. [CrossRef] [PubMed]
25. Lawton, M.P.; Brody, E.M. Assessment of older people: Self-maintaining and instrumental activities of daily living. *Gerontologist* **1969**, *9*, 179–186. [CrossRef]
26. Fried, L.P.; Tangen, C.M.; Walston, J.; Newman, A.B.; Hirsch, C.; Gottdiener, J.; Seeman, T.; Tracy, R.; Kop, W.J.; Burke, G.; et al. Frailty in Older Adults: Evidence for a Phenotype. *J. Gerontol. A Biol. Sci. Med. Sci.* **2001**, *56*, M146–M156. [CrossRef] [PubMed]
27. O'Mahony, D.; O'Sullivan, D.; Byrne, S.; O'Connor, M.N.; Ryan, C.; Gallagher, P. STOPP/START criteria for potentially inappropriate prescribing in older people: Version 2. *Age Ageing* **2015**, *44*, 213–218.
28. Ware, J., Jr.; Kosinski, M.; Keller, S.D. A 12-Item Short-Form Health Survey: Construction of scales and preliminary tests of reliability and validity. *Med. Care* **1996**, *34*, 220–233. [CrossRef]
29. Jenkinson, C.; Layte, R.; Jenkinson, D.; Lawrence, K.; Petersen, S.; Paice, C.; Stradling, J. A shorter form health survey: Can the SF-12 replicate the results from the SF-36 in longitudinal studies? *J. Public. Health. Med.* **1997**, *19*, 179–186. [CrossRef]
30. Hosmer, D.W.; Lemeshow, S.; Sturdivant, R.X. *Applied Logistic Regression*, 3rd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2013.
31. Baladón, L.; Rubio-Valera, M.; Serrano-Blanco, A.; Palao, D.J.; Fernández, A. Gender differences in the impact of mental disorders and chronic physical conditions on health-related quality of life among non-demented primary care elderly patients. *Qual. Life Res.* **2016**, *25*, 1461–1474. [CrossRef]
32. Naveiro-Rilo, J.C.; Diez-Juárez, M.D.; Flores-Zurutuza, L.; Javierre Pérez, P.; Alberte Pérez, C.; Molina Mazo, R. Quality of life in the elderly on polymedication and with multiple morbidities. *Rev. Esp. Geriatr. Gerontol.* **2014**, *49*, 158–164. [CrossRef]
33. König, H.H.; Heider, D.; Lehnert, T.; Riedel-Heller, S.G.; Angermeyer, M.C.; Matschinger, H.; Vilagut, G.; Bruffaerts, R.; Haro, J.M.; de Girolamo, G.; et al. Health status of the advanced elderly in six European countries: Results from a representative survey using EQ-5D and SF-12. *Health Qual. Life Outcomes.* **2010**, *29*, 143.
34. Juste, M.P.; Barbosa, F.B.; Carreiro, J.P. Calidad de vida en personas mayores. Apuntes para un programa de educación para la salud. *Revista de Investigación en Educación.* **2009**, *6*, 70–78. (In Spanish)
35. Van der Vorst, A.; Zijlstra, G.A.R.; De Witte, N.; Vogel, R.G.M.; Schols, J.M.G.A.; Kempen, G.I.J.M.; D-SCOPE Consortium. Explaining discrepancies in self-reported quality of life in frail older people: A mixed-methods study. *BMC Geriatr.* **2017**, *17*, 251. [CrossRef] [PubMed]
36. Orfila, F.; Ferrer, M.; Lamarca, R.; Tebe, C.; Domingo-Salvany, A.; Alonso, J. Gender differences in health-related quality of life among the elderly: The role of objective functional capacity and chronic conditions. *Soc. Sci. Med.* **2006**, *63*, 2367–2380. [CrossRef] [PubMed]
37. Padua, L.; Pasqualetti, P.; Coraci, D.; Imbimbo, I.; Giordani, A.; Loreti, C.; Marra, C.; Molino-Lova, R.; Pasquini, G.; Simonelli, I.; et al. Gender effect on well-being of the oldest old: A survey of nonagenarians living in Tuscany: The Mugello study. *Neurol. Sci.* **2018**, *39*, 509–517. [CrossRef]

38. She, R.; Yan, Z.; Jiang, H.; Vetrano, D.L.; Lau, J.T.F.; Qiu, C. Multimorbidity and Health-Related Quality of Life in Old Age: Role of Functional Dependence and Depressive Symptoms. *J. Am. Med. Dir. Assoc.* **2019**, *20*, 1143–1149. [CrossRef] [PubMed]
39. Forjaz, M.J.; Rodríguez-Blazquez, C.; Ayala, A.; Rodríguez-Rodríguez, V.; de Pedro-Cuesta, J.; García-Gutiérrez, S.; Prados-Torres, A. Chronic conditions, disability, and quality of life in older adults with multimorbidity in Spain. *Eur. J. Intern. Med.* **2015**, *26*, 176–181. [CrossRef]
40. Acar Tek, N.; Karaçil-Ermumcu, M.S. Determinants of Health Related Quality of Life in Home Dwelling Elderly Population: Appetite and Nutritional Status. *J. Nutr. Health Aging* **2018**, *22*, 996–1002. [CrossRef]
41. Akkawi, M.E.; Nik Mohamed, M.H.; Md Aris, M.A. Does inappropriate prescribing affect elderly patients' quality of life? A study from a Malaysian tertiary hospital. *Qual. Life Res.* **2019**, *28*, 1913–1920. [CrossRef]
42. Cossette, B.; Bagna, M.; Sene, M.; Sirois, C.; Lefebvre, G.P.; Germain, O.; Morais, J.A.; Gaudreau, P.; Payette, H. Association Between Anticholinergic Drug Use and Health-Related Quality of Life in Community-Dwelling Older Adults. *Drugs Aging* **2017**, *34*, 785–792. [CrossRef]
43. 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]
44. Chang, Y.C.; Yao, G.; Hu, S.C.; Wang, J.D. Depression Affects the Scores of All Facets of the WHOQOL-BREF and May Mediate the Effects of Physical Disability among Community-Dwelling Older Adults. *PLoS ONE.* **2015**, *10*, e0128356. [CrossRef]
45. De Almeida Roediger, M.; de Fátima Nunes Marucci, M.; Duim, E.L.; Santo, J.L.F.; de Oliveira Duarte, Y.A.; de Oliveira, C. Inflammation and quality of life in later life: Findings from the health, well-being and aging study (SABE). *Health Qual. Life Outcomes.* **2019**, *17*, 26. [CrossRef] [PubMed]
46. Banegas, J.R.; López-García, E.; Graciani, A.; Guallar-Castillón, P.; Gutierrez-Fisac, J.L.; Alonso, J.; Rodríguez-Artalejo, F. Relationship between obesity, hypertension and diabetes, and health-related quality of life among the elderly. *Eur. J. Cardiovasc. Prev. Rehabil.* **2007**, *14*, 456–462. [CrossRef] [PubMed]
47. Giovannini, S.; Macchi, C.; Liperoti, R.; Laudisio, A.; Coraci, D.; Loreti, C.; Vannetti, F.; Onder, G.; Padua, L.; Mugello Study Working Group. Association of Body Fat with Health-Related Quality of Life and Depression in Nonagenarians: The Mugello Study. *J. Am. Med. Dir. Assoc* **2019**, *20*, 564–568. [CrossRef] [PubMed]
48. Payne, M.E.; Porter Starr, K.N.; Orenduff, M.; Mulder, H.S.; McDonald, S.R.; Spira, A.P.; Pieper, C.F.; Bales, C.W. Quality of Life and Mental Health in Older Adults with Obesity and Frailty: Associations with a Weight Loss Intervention. *J. Nutr. Health Aging* **2018**, *22*, 1259–1265. [CrossRef]
49. Masel, M.C.; Graham, J.E.; Reistetter, T.A.; Markides, K.S.; Ottenbacher, K.J. Frailty and health related quality of life in older Mexican Americans. *Health Qual. Life Outcomes.* **2009**, *7*. [CrossRef]
50. Madariaga, I.A.; Antón, V.A.N. Aspectos estadísticos del cuestionario de calidad de vida relacionada con salud Short Form-36 (SF-36). *Estadística Española.* **2008**, *50*, 147–192. (In Spanish)

Journal of
Clinical Medicine

MDPI

Article

Multidiscipline Stroke Post-Acute Care Transfer System: Propensity-Score-Based Comparison of Functional Status

Chung-Yuan Wang [1], Hong-Hsi Hsien [2], Kuo-Wei Hung [3], Hsiu-Fen Lin [4], Hung-Yi Chiou [5,6], Shu-Chuan Jennifer Yeh [7,8], Yu-Jo Yeh [7] and Hon-Yi Shi [7,8,9,10,*]

1 Department of Physical Medicine and Rehabilitation, Pingtung Christian Hospital, Pingtung 90059, Taiwan
2 Department of Internal Medicine, St. Joseph Hospital, Kaohsiung 80760, Taiwan
3 Division of Neurology, Department of internal medicine, Yuan's General Hospital, Kaohsiung 80249, Taiwan
4 Department of Neurology, Kaohsiung Medical University Hospital, Kaohsiung 80756, Taiwan
5 School of Public Health, Taipei Medical University, Taipei 11031, Taiwan
6 Center for Neurotrauma and Neuroregeneration, Taipei Medical University, Taipei 11031, Taiwan
7 Department of Healthcare Administration and Medical Informatics, Kaohsiung Medical University, Kaohsiung 80708, Taiwan
8 Department of Business Management, National Sun Yat-sen University, Kaohsiung 80424, Taiwan
9 Department of Medical Research, Kaohsiung Medical University Hospital, Kaohsiung 80756, Taiwan
10 Department of Medical Research, China Medical University Hospital, China Medical University, Taichung 40447, Taiwan
* Correspondence: hshi@kmu.edu.tw; Fax: +886-7-3137487

Received: 2 July 2019; Accepted: 15 August 2019; Published: 16 August 2019

Abstract: Few studies have investigated the characteristics of stroke inpatients after post-acute care (PAC) rehabilitation, and few studies have applied propensity score matching (PSM) in a natural experimental design to examine the longitudinal impacts of a medical referral system on functional status. This study coupled a natural experimental design with PSM to assess the impact of a medical referral system in stroke patients and to examine the longitudinal effects of the system on functional status. The intervention was a hospital-based, function oriented, 12-week to 1-year rehabilitative PAC intervention for patients with cerebrovascular diseases. The average duration of PAC in the intra-hospital transfer group (31.52 days) was significantly shorter than that in the inter-hospital transfer group (37.1 days) ($p < 0.001$). The intra-hospital transfer group also had better functional outcomes. The training effect was larger in patients with moderate disability (Modified Rankin Scale, MRS = 3) and moderately severe disability (MRS = 4) compared to patients with slight disability (MRS = 2). Intensive post-stroke rehabilitative care delivered by per-diem payment is effective in terms of improving functional status. To construct a vertically integrated medical system, strengthening the qualified local hospitals with PAC wards, accelerating the inter-hospital transfer, and offering sufficient intensive rehabilitative PAC days are the most essential requirements.

Keywords: stroke; post-acute care; medical referral system; propensity score matching

1. Introduction

Acute stroke is a major cause of mortality and disability [1,2]. Stroke patients can incur considerable costs for medical care, including nursing, rehabilitative, and long term care. Therefore, many countries are attempting to establish comprehensive and integrated healthcare systems for stroke patients [3,4]. Post-acute care (PAC), which refers to medical care services that support the individual patient in recovery from illness or management of chronic disability, is aimed at enhancing the functional status of patients discharged from acute hospitalization [2–5]. Discharges to PAC facilities have increased

nearly 50% during the past 15 years, and PAC is a major contributor to hospitalization costs in the United States [5]. To control medical expenses, reform the medical referral system, and improve continuity of care, the Taiwan National Health Insurance Administration (NHIA) focused on stroke for its first national PAC project in 2014—post-acute care for cerebrovascular disease (PAC-CVD).

In Taiwan, beneficiaries are free to visit their preferred physicians and are not required to follow strict referral rules [6]. An efficient PAC system for stroke patients is needed to reduce unnecessary utilization of hospital resources and to ensure seamless care for these patients [7]. Stroke patients and their families expect that local hospitals can deliver PAC at a lower cost and with greater efficiency, effectiveness, and convenience. Our review of studies published in international journals, however, shows that most studies of PAC have analyzed a limited number of patients in a single hospital [8,9]. Additionally, few studies have used longitudinal data exceeding 1 year, and few studies have applied propensity score matching (PSM) in a natural experimental design to examine the longitudinal impacts of a medical referral system on functional status. Therefore, this study coupled a natural experimental design with PSM to assess the impact of the medical referral system in stroke patients and to examine the longitudinal effects of the medical referral system on functional status.

2. Materials and Methods

2.1. The PAC Program

In Taiwan, the multidisciplinary PAC stroke team consisted of neurologists, physiatrists, physical therapists, occupational therapists, speech therapists, and nurses. The PAC rehabilitation program was prescribed by the physiatrist, and it consisted of a complex program of universal activities that were performed at least three times per day. One hour of physical therapy, occupational therapy, or speech and swallowing therapy was carried out at each time. Notably, the fiscal incentive for a medical center to transfer a patient to a regional or a district hospital is mitigated by several factors, including the PAC-CVD transfer policy, the willingness of the stroke patient (or family) to accept further post-acute care in local hospitals, and whether the physician agrees to the transfer. These factors should cause health providers to reconsider the manner in which patient stays are controlled and to be mindful that the shortest lengths of stay (LOS) may not obtain the best outcome for the patient [10,11].

2.2. Study Design and Sample

The study population included all stroke patients admitted to PAC wards in four Taiwan hospitals between March 2014 and March 2018 (defined as ICD-9-CM codes 433.x, 434.x, and 436.x for ischemic stroke, and codes 430 and 431 for hemorrhagic stroke). The inclusion criteria were acute stroke and admission to PAC ward within 40 days after day of stroke onset. Another inclusion criterion was Modified Rankin Scale (MRS) level 2 to 4. Instead of focusing on patients who had received intensive in-patient rehabilitation, this national PAC project focused on the prevention of complications (e.g., pressure sores) in stroke patients who were bed-ridden (MRS 5). Patients who did not have major disability (MRS 0–1) were assumed to have undergone out-patient rehabilitation or were assumed to have resumed their pre-morbidity activities of daily living.

In observational studies, non-comparability between the intervention group and the comparison group can distort the estimation of the treatment effect [12,13]. The propensity score is a balancing score that can be used to compare groups that do not systematically differ. This study used PSM at the patient level to compare the baseline characteristics of the two groups, which increased the robustness of the analysis. A generalized estimating equation (GEE) model was used to cluster stroke patients treated by the same physician and to generate propensity scores for predicting the probability of the medical referral system. The covariates included patient demographics (age and gender), clinical attributes (stroke type, hypertension, hyperlipidemia, diabetes mellitus, atrial fibrillation, and previous stroke), quality of medical care (acute care LOS and PAC LOS), and pre-rehabilitation functional status. The caliper matching method ("greedy algorithm") was used for 1:1 PSM between the inter-hospital

transfer group and the intra-hospital transfer group. Thus, 483 patients in the inter-hospital transfer group were compared with an "all participants matched set" of 483 patients in the intra-hospital transfer group (Figure 1). These PAC stroke patients completed the pre-rehabilitation and the 12th week and first year post-rehabilitation assessments.

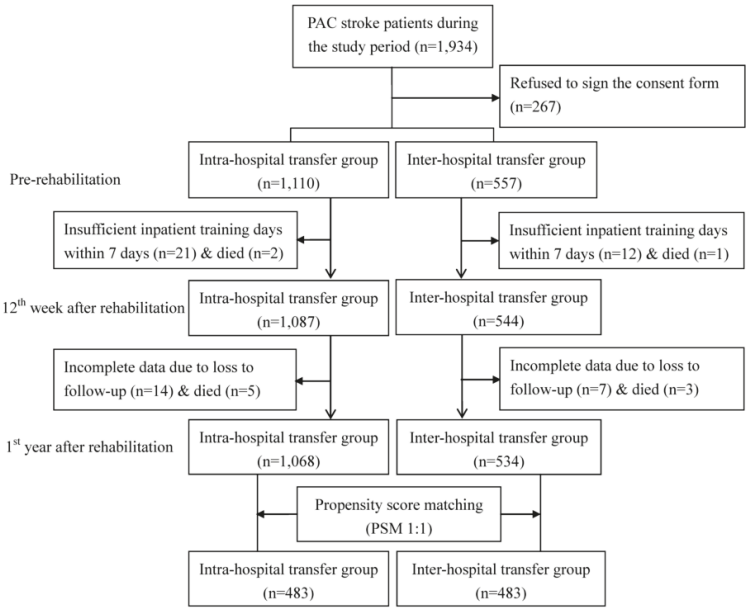


Figure 1. Flow chart of recruitment and study procedure.

2.3. Functional Status Instruments

The MRS scores of 0, 1, 2, 3, 4, and 5 are interpreted as no symptoms, no significant disability, slight disability, moderate disability, moderately severe disability, and severe disability, respectively [14]. The Barthel Index (BI) score was used to measure functional disability in daily life activities (e.g., eating, grooming, bathing, dressing, walking, transferring, staring, and controlling bladder and bowel) [15]. The score is for the 10-item BI ranges from 0 (totally dependent) to 100 (independent). The Functional Oral Intake Scale (FOIS) was used to assess functional oral intake in stroke patients with dysphagia [16]. The FOIS classifies swallowing function from level 1 (nothing by mouth) to level 7 (total oral diet with no restrictions). The Lawton-Brody Instrumental Activities of Daily Living Scale (IADL) is used to evaluate performance in daily life activities, including making telephone calls, shopping, preparing food, housekeeping, laundering, taking medicine, using transportation, and performing financial activities [17]. In the conventional use of the scale, women are scored in all eight domains, while men are not scored in the domains of preparing food, housekeeping, and laundering. The rationale for excluding these three domains in males is that performance of these tasks is subject to cultural differences in gender roles, which could compromise comparisons of the incidence of disability between men and women [18]. The EuroQoL five-dimensional (EQ-5D) measure is a self-assessment of mobility, self-care, usual activities, pain or discomfort, and anxiety or depression as part of a total health state [19]. The subject is required to score each item from 1 to 3 (no problem, some problem, or extreme problem, respectively). The Berg Balance Scale (BBS) is a scale of functional balance, including static and dynamic balance [20]. Each item on this 14-item scale is rated from 0 (poor balance) to 4 (good balance). The maximum score is 56. The Mini-Mental State Examination (MMSE) is the best-known short screening tool for cognitive impairment [21]. The MMSE tests orientation, attention, memory,

language, and visual–spatial skills. The maximum score is 30 points. An MMSE score below an education-adjusted cut-off score indicates cognitive impairment. The Taiwan version of these measures has been validated as a reliable and valid tool for measuring functional status in both clinical practice and research [22].

2.4. Statistical Analysis

The unit of analysis in this study was the individual stroke patient. Descriptive statistics were tabulated to depict the stroke patient demographics. For clarification, the values predicted by the regression models were used to illustrate the results, starting from before initiation of PAC until completion of 12 weeks to 1 year of follow up in the two matched study groups. Thus, the GEE models were used to estimate difference-in-differences models used to examine the effectiveness of the medical referral system. For the predicted values, standard errors in differences and standard errors in difference-in-differences were estimated using a bootstrap technique involving 1000 replications, with sample sizes equivalent to that of the original sample [23].

Hierarchical linear regression models were used to examine the roles of the MRS after accounting for demographic characteristics, clinical attributes, quality of care, and pre-rehabilitative function status. Five-step hierarchical linear regression models were used to analyze differences between explanatory factors. In Model 1, explanatory factors included age, gender, stroke type, hypertension, hyperlipidemia, diabetes mellitus, atrial fibrillation, previous stroke, length of stay in acute care, length of stay in post-acute care, and pre-rehabilitation functional status. Model 2 included Model 1 and MRS = 2; Model 3 included Model 1 and MRS = 3; Model 4 included Model 1 and MRS = 4; Model 5 included Model 1, MRS, and medical referral system. In each model, the adjusted *R*-square was estimated while adjusting for covariates. For each model, the variance inflation factor (VIF) was used to assess multicollinearity. No models showed multicollinearity.

Statistical analyses were performed using Stata Statistical Package, version 13.0 (Stata Corp, College Station, TX, USA). All tests were two-sided, and p values less than 0.05 were considered statistically significant.

3. Results

Table 1 compares the inter-hospital transfer group and the intra-hospital transfer group before and after PSM. Before PSM, all assessed characteristics significantly differed between the two groups ($p < 0.05$). After PSM, no variables significantly differed between the two groups.

All PAC stroke patients had significantly improved scores for functional measures at the 1-year follow-up survey ($p < 0.001$) (Table 2). When the 12th week post-rehabilitation scores were used as the baseline, functional status showed significant improvements at the first year post-rehabilitation ($p < 0.001$). All subscale scores continued to improve throughout the follow-up period. Additionally, in both groups, functional status scores at the first year post-rehabilitation were significantly higher than the functional status scores at the 12th week post-rehabilitation and the functional status scores at pre-rehabilitation ($p < 0.001$). Throughout the follow-up period, the intra-hospital transfer group also had significantly higher scores for functional status measures compared to the inter-hospital transfer group ($p < 0.001$).

Table 3 shows that Model 1 revealed a significant association between patient demographics and scores for functional status measures at the first year post-rehabilitation during the study period ($p < 0.05$). In Models 2–4, patients with MRS = 2 showed very little functional status improvement after adjustment for patient demographics; however, patients with MRS = 3 and MRS = 4 showed improvements. That is, rehabilitative PAC improved quality of life in patients with moderate disability (MRS = 3) or moderately severe disability (MRS = 4) but not in patients with slight disability (MRS = 2). After adjustment for all relevant influential factors, the intra-hospital transfer group had greater improvements in functional status scores compared to the inter-hospital transfer group.

Table 1. Stroke patient characteristics *.

Variables		Before Propensity Score Matching			After Propensity Score Matching		
		Intra-Hospital Transfer Group (n = 1068)	Inter-Hospital Transfer Group (n = 534)	p Value	Intra-Hospital Transfer Group (n = 483)	Inter-Hospital Transfer Group (n = 483)	p Value
Demographics							
Age, years *		65.67 ± 12.38	63.96 ± 13.50	0.024	63.61 ± 13.10	63.96 ± 13.24	0.834
Gender	Female	450(42.1%)	186(34.8%)	0.014	171(35.4%)	169(35.0%)	0.784
	Male	618(57.9%)	348(65.2%)		312(64.6%)	314(65.0%)	
Clinical Attributes							
Stroke type	Ischemic	942(88.2%)	396(74.2%)	<0.001	363(75.0%)	360(74.5%)	0.880
	Hemorrhagic	126(11.8%)	138(25.8%)		120(25.0%)	123(25.5%)	
Hypertension	Yes	698(65.4%)	405(75.8%)	<0.001	372(77.0%)	367(76.0%)	0.221
Hyperlipidemia	Yes	332(31.1%)	227(42.5%)	<0.001	202(41.8%)	206(42.6%)	0.507
Diabetes mellitus	Yes	419(39.2)	200(37.5%)	0.570	182(37.7%)	181(37.5%)	0.990
Atrial fibrillation	Yes	67(6.3%)	55(10.3%)	0.013	47(9.7%)	49(10.1%)	0.887
Previous stroke	Yes	173(16.2%)	98(18.4%)	0.803	87(18.0%)	89(18.4%)	0.879
Quality of Medical Care							
Acute care LOS, days *		13.01 ± 27.83	24.45 ± 34.61	<0.001	23.75 ± 11.84	24.50 ± 11.56	0.356
PAC LOS, days *		31.52 ± 17.75	37.1 ± 12.59	<0.001	35.75 ± 12.34	36.50 ± 11.88	0.506
Pre-rehabilitation functional status							
BI *		41.91 ± 23.10	34.67 ± 23.48	<0.001	35.75 ± 20.11	34.00 ± 18.21	0.269
FOIS *		5.95 ± 3.04	5.38 ± 2.25	<0.001	5.53 ± 2.75	5.14 ± 2.84	0.927
EQ5D *		10.67 ± 1.86	10.40 ±1.78	0.015	10.80 ± 1.82	10.93 ± 2.05	0.261
IADL *		1.41 ± 1.20	1.15 ± 1.12	<0.001	1.32 ± 1.14	1.27 ± 1.05	0.694
BBS *		15.30 ± 14.99	16.91 ± 17.27	0.097	14.00 ± 17.26	15.50 ± 17.71	0.972
MMSE *		20.15 ± 7.90	18.50 ± 9.66	0.001	20.75 ± 11.15	19.75 ± 10.47	0.908

Note: LOS, length of stay; PAC, post-acute care; QOL, quality of life; BI, Barthel Index; FOIS, Functional Oral Intake Scale; EQ5D, EuroQoL five-dimensional; IADL, Instrumental Activities of Daily Living Scale; BBS, Berg Balance Scale; MMSE, Mini-Mental State Examination. * Values are expressed as mean ± standard deviation or n (%).

Table 2. Comparison of functional status between intra-hospital transfer group and inter-hospital transfer group, before and after rehabilitation (n = 1195).

Functional Status Measure	Group	Before Rehabilitation (T1)	12th Week After Rehabilitation (T2)	Difference †	p Value	First Year After Rehabilitation (T3)	Difference †	p Value
		Mean ± SD				Mean ± SD		
BI	Intra-hospital transfer	41.91 ± 23.10	51.50 ± 24.10	9.59	<0.001	68.84 ± 26.49	17.34	<0.001
	Inter-hospital transfer	34.67 ± 23.48	42.28 ± 25.96	7.61	<0.001	54.29 ± 27.20	12.01	0.002
	Difference ‡	7.24	9.22	1.98	<0.001	14.55	5.33	<0.001
FOIS	Intra-hospital transfer	5.95 ± 3.04	5.98 ± 1.82	0.03	0.037	6.41 ± 1.32	0.43	0.006
	Inter-hospital transfer	5.38 ± 2.25	5.68 ± 1.89	0.30	0.002	6.26 ± 1.47	0.58	0.044
	Difference ‡	0.57	0.30	−0.27	<0.001	0.15	−0.15	<0.001
EQ5D	Intra-hospital transfer	10.67 ± 1.86	9.81 ± 1.65	−0.86	<0.001	8.15 ± 2.23	−1.66	<0.001
	Inter-hospital transfer	10.40 ± 1.79	10.02 ± 1.70	−0.38	0.004	9.19 ± 1.88	−0.83	0.001
	Difference ‡	0.27	−0.21	−0.48	<0.001	−1.04	−0.83	<0.001
IADL	Intra-hospital transfer	1.41 ± 1.20	1.84 ± 1.32	0.43	<0.001	2.87 ± 1.75	1.03	<0.001
	Inter-hospital group	1.15 ± 1.13	1.36 ± 1.31	0.21	0.054	1.86 ± 1.57	0.5	0.023
	Difference ‡	0.26	0.48	0.22	<0.001	1.01	0.53	<0.001
BBS	Intra-hospital transfer	15.30 ± 14.99	26.32 ± 17.56	11.02	<0.001	34.58 ± 17.79	8.26	<0.001
	Inter-hospital transfer	16.91 ± 17.27	24.40 ± 19.25	7.49	<0.001	29.91 ± 19.35	5.51	0.022
	Difference ‡	−1.61	1.92	3.53	<0.001	4.67	2.75	<0.001
MMSE	Intra-hospital transfer	20.15 ± 7.90	21.62 ± 7.79	1.47	0.020	22.73 ± 7.39	1.11	0.078
	Inter-hospital transfer	18.50 ± 9.66	19.64 ± 10.80	1.14	0.090	21.25 ± 9.41	1.61	0.260
	Difference ‡	1.65	1.98	0.33	<0.001	1.48	−0.5	<0.001

Note: BI, Barthel Index; FOIS, Functional Oral Intake Scale; EQ5D, EuroQoL five-dimensional; IADL, Instrumental Activities of Daily Living Scale; BBS, Berg Balance Scale; MMSE, Mini-mental State Examination. † Difference indicates mean score for functional status at the 12th week after rehabilitation, mean score for functional status before rehabilitation, or mean score for functional status at the first year after rehabilitation. ‡ Difference indicates mean score for functional status in intra-hospital transfer group, or mean score for functional status in inter-hospital transfer group at each time point.

Table 3. Change in coefficient of multiple correlations associated with addition of subsequent variables to the model *.

Model	BI		FOIS		EQ5D		IADL		BBS		MMSE	
	R^2	R^2 Change	R^2	R^2 Change	R^2	R^2 Change	R^2	R^2 Change	R^2	R^2 Change	R^2	R^2 Change
1st level	0.64	-	0.28	-	0.38	-	0.51	-	0.48	-	0.79	-
2nd level	0.64	0.00	0.28	0.00	0.39	0.01	0.52	0.01	0.48	0.00	0.79	0.00
3rd level	0.64	0.00	0.29	0.01	0.44	0.06	0.55	0.04	0.51	0.03	0.79	0.00
4th level	0.64	0.00	0.29	0.01	0.44	0.06	0.55	0.04	0.51	0.03	0.79	0.00
5th level	0.67	0.03	0.29	0.01	0.49	0.11	0.58	0.07	0.53	0.05	0.79	0.00

Note: BI, Barthel Index; FOIS, Functional Oral Intake Scale; EQ5D, EuroQoL five-dimensional; IADL, Instrumental Activities of Daily Living Scale; BBS, Berg Balance Scale; MMSE, Mini-mental State Examination. * Model 1 included age, gender, stroke type, hypertension, hyperlipidemia, diabetes mellitus, atrial fibrillation, previous stroke, length of stay in acute care, length of stay in post-acute care, and pre-rehabilitation functional status; Model 2 included Model 1 and Modified Rankin Scale (MRS) = 2; Model 3 included Model 1 and MRS = 3; Model 4 included Model 1 and MRS = 4; Model 5 included Model 1, MRS, and medical referral system.

4. Discussion

Our data for the percentage of stroke patients with vascular risk factors were consistent with previous reports. After PSM, the percentages of stroke patients with hypertension, hyperlipidemia, diabetes mellitus, atrial fibrillation in our study were 77%, 42%, 37%, and 10%, respectively. Previous studies of stroke patients have reported hypertension in 63–80%, hyperlipidemia in 40–49%, diabetes mellitus in 34–42%, and atrial fibrillation in 7.3–11% [24,25]. A review of studies performed in Asia found that Taiwan has a higher prevalence of hypertension, diabetes mellitus, and hyperlipidemia compared to Japan, Korea and Singapore [26]. A previous Taiwan study of stroke incidence and recurrence during 2000–2011 also reported that, although the prevalence of diabetes mellitus, hyperlipidemia and atrial fibrillation increased during this period, the rates of primary ischemic stroke and 1-year recurrence of stroke decreased by 9% and 18% respectively [27]. Factors that can have important effects on stroke incidence and recurrence include medication control, early detection of diseases, diet control, body weight control, life style adjustment and exercise.

To achieve a vertically integrated medical system, post-stroke care should be patient-centered, and inter-hospital transfer of stroke patients must be seamless. In Australia, stroke patients treated at hospitals with stroke coordinators had lower LOS and higher quality of evidence-based care compared to those treated at hospitals without stroke coordinators [28]. In the Taiwan national PAC project reported here, the case manager assigned to each stroke patient ensured efficient transfer of the patient to a local hospital. Therefore, the mean LOS for patients in acute stage was much lower in the intra-hospital transfer group compared to the inter-hospital transfer group (13.01 and 24.45 days, respectively). Since minimizing the time from admission to inter-hospital transfer decreases total LOS and total cost, implementation of a case manager and an efficient inter-hospital transfer system are essential for an integrated medical system. Other possible reasons in the inter-hospital transfer group include geographical variations in the distribution of physicians and medical resources, and differences in care quality and expertise among hospitals and individual providers. For stroke patients in acute stage, those treated at teaching hospitals and certified primary stroke centers have a lower mortality rate, greater availability of rehabilitative care, and lower ADL dependence status [29]. A possible reason for the superior outcomes obtained by teaching hospitals and certified primary care centers for stroke is that they tend to have a high volume of stroke patients, and clinicians who treat a high volume of patients tend to achieve high skill levels.

Before PSM, compared to the intra-hospital transfer group, the inter-hospital transfer group in this study had a lower functional status before PAC and more comorbidity (hypertension, hyperlipidemia, atrial fibrillation). Most of the patients in inter-hospital transfer group had been treated at a medical center. Compared to stroke patients treated at medical centers tend to have a higher severity of disability, are more likely to require intubation (e.g., tracheostomy tube, nasogastric tube, and urinary catheter tube), and tend to require more time to stabilize. These differences might explain why the inter-hospital transfer group in our study had a longer mean LOS in acute stage compared to the intra-hospital transfer group. However, further studies are needed to compare the service path and treatment costs in patients with varying severity of stroke and in stroke patients treated at different hospital levels.

After the acute stage, local low-volume rehabilitation facilities can usually provide adequate care [30]. Several studies have reported that, compared to skilled nursing facilities, intensive inpatient rehabilitation facilities achieve higher functional outcomes in PAC for stroke [30,31]. Additionally, duration of hospital stay and in-hospital mortality are related to socioeconomic status [32]. Therefore, increasing socioeconomic inequality over time has markedly increased inequality in stroke survival. Direct non-healthcare costs (including informal care costs, paid care costs and transportation costs) and rehabilitation costs are the largest post-stroke care costs [33,34]. Stroke patients who receive PAC and rehabilitation at local hospitals can reduce the physical, mental and economic burdens on their families. Our study analyzed data obtained from four local southern Taiwan hospitals that had the

largest volumes of stroke inpatients in PAC. Therefore, the data obtained in this work are highly representative of hospitals throughout Taiwan and have immediate applications.

In Japan, the rehabilitation is 3 hours, and the average LOS in rehabilitation facilities after discharge from tertiary hospitals is approximately 74.7 days [35]. The Taiwan PAC-CVD program provides the maximum of 12 weeks of services. Reimbursements are similar regardless of the hospital accreditation level (regional or district) and the equipment costs of the hospital. The PAC-CVD project enables stroke patients to access rehabilitation programs that are more intensive (in terms of frequency and duration of treatment) compared to those available under current NHIA provisions. The PAC-CVD project was expected to offer a continuous care model to restore function and reduce disability in stroke patients [8,9]. Most PAC-CVD studies published thus far have been studies of a limited case number in a single hospital [8,9]. In the post-acute stage, the mean LOS in the intra-hospital transfer group and the inter-hospital transfer group was 31.52 and 37.1 days, respectively. Another study of a large population of stroke patients reported a mean PAC stay of only 15.1 days [36]. Notably, the LOS of stroke patients in different countries is related to differences in post-stroke policy. For example, the average LOS for inpatient rehabilitation after stroke is approximately 1 month in Ireland, Switzerland, and Thailand [37]. In contrast, the average LOS inpatient rehabilitation after stroke is only 15.8 days in the United States, which is much shorter than that in western countries (e.g., 32.7 days in Germany, 35.3 days in Canada) [38–40]. The main reasons for the short LOS in the United States are the high economic burden of inpatient care for stroke patients and the availability of various well-established PAC rehabilitation facilities (e.g., inpatient rehabilitation hospitals, skilled nursing facilities, home health agency services, long term care hospitals) [4]. According to our multi-center data, the LOS for stroke patients who undergo PAC in Taiwan is similar to that in other countries. Notably, our data indicated that the effectiveness of rehabilitation training during the 12th week to the first year is better than in the first 12 weeks. These data confirm that a continuous rehabilitative PAC program is essential.

A study by Dewilde concluded that MRS level is a major determinant of medical resource use [41]. For this national PAC-CVD project, the MRS level was the main criterion for participation. Initially, only patients with MRS levels of 2 to 4 were eligible for transfer to hospitals with PAC wards. However, some people have proposed excluding patients in MRS level 2 or including patients in MRS level 5. In 2019, the Taiwan NHIA excluded MRS level 2 patients from this national PAC-CVD project. This study found that after adjusting for all relevant influential factors, the quality of life improvement was smaller in patients at MRS level 2 compared to patients at other MRS levels. Therefore, these data support the decision made by the Taiwan NHIA. In terms of maximizing efficiency in the use of limited healthcare resources, the decision to limit inpatient rehabilitation to patients with MRS 3 to 4 was not only justifiable, but necessary.

Although all research questions were adequately and satisfactorily addressed, two limitations are noted. This study only collected data for acute stroke patients for 40 days after stroke onset. Furthermore, this study only analyzed patients treated at four hospitals in south Taiwan. However, the numbers of patients treated in the PAC-programs at these hospitals were among the four highest of all district hospitals in south Taiwan. Further studies are needed to compare a PAC group and a control group in other regions of Taiwan and under current NHI regulations. Additionally, the two groups in this study were matched for demographic characteristics, clinical attributes, quality of care, and pre-rehabilitative functional status. Future studies could consider the use of inverse probability weighting rather than PSM.

5. Conclusions

In conclusion, this study showed that rehabilitative PAC improved outcomes of stroke rehabilitation. To achieve a vertically integrated medical system for stroke rehabilitation, the key requirements are improving the PAC ward qualifications of local hospitals, accelerating inter-hospital transfer, and ensuring a sufficient duration of intensive rehabilitative PAC. Early rehabilitation is important for successful restoration of health, confidence, and self-care ability in these patients.

Author Contributions: Study concept and design, C.-Y.W. and H.-Y.S.; acquisition of data, H.-H.H., K.-W.H., H.-F.L., H.-Y.C., S.-C.J.Y., and Y.-J.Y.; analysis and interpretation of data, C.-Y.W. and H.-Y.S.; drafting of the manuscript, C.-Y.W. and H.-Y.S.; revision of the manuscript, H.-H.H., K.-W.H., H.-F.L., H.-Y.C., S.-C.J.Y., and Y.-J.Y.; approval of the manuscript, all authors.

Funding: This study was supported by funding from the Pingtung Christian Hospital (PS104001) and Ministry of Science and Technology (MOST 104-2410-H-037-006-SS2 and MOST 106-2410-B-037-076) in Taiwan.

Acknowledgments: We thank all the individuals who participated in this study, without which this study would not have been possible.

Conflicts of Interest: Authors declare that they have no competing interest.

References

1. Prabhakaran, S.; Ruff, I.; Bernstein, R.A. Acute stroke intervention: A systematic review. *JAMA* **2015**, *313*, 1451–1462. [CrossRef] [PubMed]
2. Sturm, J.W.; Dewey, H.M.; Donnan, G.A.; Macdonell, R.A.; McNeil, J.J. Handicap after stroke: How does it relate to disability, perception of recovery, and stroke subtype? The north North East Melbourne Stroke Incidence Study (NEMESIS). *Stroke* **2002**, *33*, 762–768. [CrossRef] [PubMed]
3. Burke, R.E.; Juarez-Colunga, E.; Levy, C.; Prochazka, A.V.; Coleman, E.A.; Ginde, A.A. Rise of post-acute care facilities as a discharge destination of US hospitalizations. *JAMA Intern. Med.* **2015**, *175*, 295–296. [CrossRef] [PubMed]
4. Huckfeldt, P.J.; Mehrotra, A.; Hussey, P.S. The Relative Importance of Post-Acute Care and Readmissions for Post-Discharge Spending. *Health Serv. Res.* **2016**, *51*, 1919–1938. [CrossRef] [PubMed]
5. Redberg, R.F. The role of post-acute care in variation in the Medicare program. *JAMA Intern. Med.* **2015**, *175*, 1058. [CrossRef] [PubMed]
6. Chen, C.C.; Cheng, S.H. Does pay-for-performance benefit patients with multiple chronic conditions? Evidence from a universal coverage health care system. *Health Policy Plan.* **2016**, *31*, 83–90. [CrossRef] [PubMed]
7. Lee, H.C.; Chang, K.C.; Lan, C.F. Factors associated with prolonged hospital stay for acute stroke in Taiwan. *Acta Neurol. Taiwan* **2008**, *17*, 17–25. [PubMed]
8. Wang, C.Y.; Chen, Y.R.; Hong, J.P.; Chan, C.C.; Chang, L.C.; Shi, H.Y. Rehabilitative post-acute care for stroke patients delivered by per-diem payment system in different hospitalization paths: A Taiwan pilot study. *Int. J. Qual. Health Care* **2017**, *29*, 779–784. [CrossRef] [PubMed]
9. Lai, C.L.; Tsai, M.M.; Luo, J.Y.; Liao, W.C.; Hsu, P.S.; Chen, H.Y. Post-acute care for stroke—A retrospective cohort study in Taiwan. *Patient Prefer. Adherence* **2017**, *11*, 1309–1315. [CrossRef]
10. Huang, S.T.; Yu, T.M.; Ke, T.Y.; Wu, M.J.; Chuang, Y.W.; Li, C.Y.; Chiu, C.W.; Lin, C.L.; Liang, W.M.; Chou, T.C.; et al. Intensive Periodontal Treatment Reduces Risks of Hospitalization for Cardiovascular Disease and All-Cause Mortality in the Hemodialysis Population. *J. Clin. Med.* **2018**, *7*, 344. [CrossRef]
11. Cheng, C.Y.; Hsu, C.Y.; Wang, T.C.; Jeng, Y.C.; Yang, W.H. Evaluation of Cardiac Complications Following Hemorrhagic Stroke Using 5-Year Centers for Disease Control and Prevention (CDC) Database. *J. Clin. Med.* **2018**, *7*, 519. [CrossRef] [PubMed]
12. Rosenbaum, P.R.; Rubin, D.B. The central role of the propensity score in observational studies for causal effects. *Biometrika* **1983**, *70*, 41–55. [CrossRef]
13. D'Agostino, R.B. Tutorial in biostatistics. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med.* **1998**, *17*, 2265–2281. [CrossRef]
14. Banks, J.L.; Marotta, C.A. Outcomes validity and reliability of the modified Rankin scale: Implications for stroke clinical trials: A literature review and synthesis. *Stroke* **2007**, *38*, 1091–1096. [CrossRef] [PubMed]
15. Wolfe, C.D.; Taub, N.A.; Woodrow, E.J.; Burney, P.G. Assessment of scales of disability and handicap for stroke patients. *Stroke* **1991**, *22*, 1242–1244. [CrossRef] [PubMed]
16. Crary, M.A.; Mann, G.D.; Groher, M.E. Initial psychometric assessment of a functional oral intake scale for dysphagia in stroke patients. *Arch. Phys. Med. Rehabil.* **2005**, *86*, 1516–1520. [CrossRef] [PubMed]
17. Lawton, M.P.; Brody, E.M. Assessment of older people: Self-maintaining and instrumental activities of daily living. *Gerontologist* **1969**, *9*, 179–186. [CrossRef]

18. Alexandre, T.S.; Corona, L.P.; Nunes, D.P.; Santos, J.L.; Duarte, Y.A.; Lebrão, M.L. Disability in instrumental activities of daily living among older adults: Gender differences. *Rev. Saude Publica* **2014**, *48*, 379–389. [CrossRef]
19. Rabin, R.; de Charro, F. EQ-5D: A measure of health status from the EuroQol group. *Ann. Med.* **2001**, *33*, 337–343. [CrossRef]
20. Chou, C.Y.; Chien, C.W.; Hsueh, I.P.; Sheu, C.F.; Wang, C.H.; Hsieh, C.L. Developing a short form of the Berg Balance Scale for people with stroke. *Phys. Ther.* **2006**, *86*, 195–204.
21. 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.* **1995**, *12*, 189–198. [CrossRef]
22. Guo, N.W.; Lui, H.C.; Wong, P.F.; Liao, K.K.; Yan, S.H.; Lin, K.P. Chinese version and norms of the Mini-Mental State Examination. *J. Rehabil. Med. Assoc. Taiwan* **1988**, *16*, 52–59.
23. Efron, B.; Tibshirani, R. *An Introduction to the Bootstrap*; Chapman & Hall: New York, NY, USA, 1993.
24. Hsieh, C.Y.; Wu, D.P.; Sung, S.F. Trends in vascular risk factors, stroke performance measures, and outcomes in patients with first-ever ischemic stroke in Taiwan between 2000 and 2012. *J. Neurol. Sci.* **2017**, *378*, 80–84. [CrossRef] [PubMed]
25. Cheng, J.H.; Zhang, Z.; Ye, Q.; Ye, Z.S.; Xia, N.G. Characteristics of the ischemic stroke patients whose seizures occur at stroke presentation at a single institution in Eastern China. *J. Neurol. Sci.* **2018**, *387*, 46–50. [CrossRef] [PubMed]
26. Hsieh, F.I.; Chiou, H.Y. Stroke: Morbidity, Risk Factors, and Care in Taiwan. *J. Stroke* **2014**, *16*, 59–64. [CrossRef] [PubMed]
27. Lee, M.; Wu, Y.L.; Ovbiagele, B. Trends in Incident and Recurrent Rates of First-Ever Ischemic Stroke in Taiwan between 2000 and 2011. *J. Stroke* **2016**, *18*, 60–65. [CrossRef] [PubMed]
28. Purvis, T.; Kilkenny, M.F.; Middleton, S.; Cadilhac, D.A. Influence of stroke coordinators on delivery of acute stroke care and hospital outcomes: An observational study. *Int. J. Stroke* **2018**, *13*, 585–591. [CrossRef]
29. Bettger, J.P.; Thomas, L.; Liang, L.; Xian, Y.; Bushnell, C.D.; Saver, J.L.; Fonarow, G.C.; Peterson, E.D. Hospital Variation in Functional Recovery After Stroke. *Circ. Cardiovasc. Qual. Outcomes* **2017**, *10*, e002391. [CrossRef]
30. Graham, J.E.; Deutsch, A.; O'Connell, A.A.; Karmarkar, A.M.; Granger, C.V.; Ottenbacher, K.J. Inpatient rehabilitation volume and functional outcomes in stroke, lower extremity fracture, and lower extremity joint replacement. *Med. Care* **2013**, *51*, 404–412. [CrossRef]
31. Alcusky, M.; Ulbricht, C.M.; Lapane, K.L. Postacute Care Setting, Facility Characteristics, and Poststroke Outcomes: A Systematic Review. *Arch. Phys. Med. Rehabil.* **2018**, *99*, 1124–1140. [CrossRef]
32. Blattner, M.; Price, J.; Holtkamp, M.D. Socioeconomic class and universal healthcare: Analysis of stroke cost and outcomes in US military healthcare. *J. Neurol. Sci.* **2018**, *386*, 64–68. [CrossRef] [PubMed]
33. Fattore, G.; Torbica, A.; Susi, A.; Giovanni, A.; Benelli, G.; Gozzo, M.; Toso, V. The social and economic burden of stroke survivors in Italy: A prospective, incidence-based, multi-centre cost of illness study. *BMC Neurol.* **2012**, *12*, 137. [CrossRef] [PubMed]
34. Rajsic, S.; Gothe, H.; Borba, H.H.; Sroczynski, G.; Vujicic, J.; Toell, T.; Siebert, U. Economic burden of stroke: A systematic review on post-stroke care. *Eur. J. Health Econ.* **2019**, *20*, 107–134. [CrossRef] [PubMed]
35. Miyai, I.; Sonoda, S.; Nagai, S.; Takayama, Y.; Inoue, Y.; Kakehi, A.; Kurihara, M.; Ishikawa, M. Results of new policies for inpatient rehabilitation coverage in Japan. *Neurorehabil. Neural. Repair.* **2011**, *25*, 540–547. [CrossRef] [PubMed]
36. Peng, L.N.; Lu, W.H.; Liang, C.K.; Chou, M.Y.; Chung, C.P.; Tsai, S.L.; Chen, Z.J.; Hsiao, F.Y.; Chen, L.K. Functional Outcomes, Subsequent Healthcare Utilization, and Mortality of Stroke Postacute Care Patients in Taiwan: A Nationwide Propensity Score-matched Study. *J. Am. Med. Dir. Assoc.* **2017**, *18*, 990.e7–990.e12. [CrossRef] [PubMed]
37. Kuptniratsaikul, V.; Kovindha, A.; Massakulpan, P.; Permsirivanich, W.; Kuptniratsaikul, P.S. Inpatient rehabilitation services for patients after stroke in Thailand: A multi-centre study. *J. Rehabil. Med.* **2009**, *41*, 684–686. [CrossRef] [PubMed]
38. Lewis, Z.H.; Hay, C.C.; Graham, J.E.; Lin, Y.L.; Karmarkar, A.M.; Ottenbacher, K.J. Social Support and Actual Versus Expected Length of Stay in Inpatient Rehabilitation Facilities. *Arch. Phys. Med. Rehabil.* **2016**, *97*, 2068–2075. [CrossRef] [PubMed]
39. Meyer, M.; Britt, E.; McHale, H.A.; Teasell, R. Length of stay benchmarks for inpatient rehabilitation after stroke. *Disabil. Rehabil.* **2012**, *34*, 1077–1081. [CrossRef] [PubMed]

40. Verpillat, P.; Dorey, J.; Guilhaume-Goulant, C.; Dabbous, F.; Brunet, J.; Aballéa, S. A chart review of management of ischemic stroke patients in Germany. *J. Mark. Access Health Policy* **2015**, *3*. [CrossRef]
41. Dewilde, S.; Annemans, L.; Peeters, A.; Hemelsoet, D.; Vandermeeren, Y.; Desfontaines, P.; Brouns, R.; Vanhooren, G.; Cras, P.; Michielsens, B.; et al. Modified Rankin scale as a determinant of direct medical costs after stroke. *Int. J. Stroke* **2017**, *12*, 392–400. [CrossRef]

Journal of
Clinical Medicine

MDPI

Article

Relationships between Multimorbidity and Suicidal Thoughts and Plans among Korean Adults

Youn Huh [1], Ga Eun Nam [2,*], Yang-Hyun Kim [2] and Jun Hyung Lee [1,*]

[1] Department of Family Medicine, Inje University Ilsan Paik Hospital, College of Medicine, Inje University, 170, Juhwa-ro, Ilsanseo-gu, Goyang-si, Gyeonggi-do, Seoul 10380, Korea
[2] Department of Family Medicine, Korea University Anam Hospital, Korea University College of Medicine, 73, Goryeodae-ro, Seongbuk-gu, Seoul 02841, Korea
* Correspondence: namgaaa@daum.net (G.E.N.); leeddoll@naver.com (J.H.L.); Tel.: +82-2-920-5104 (G.E.N.); +82-31-910-7115 (J.H.L.)

Received: 30 June 2019; Accepted: 22 July 2019; Published: 24 July 2019

Abstract: Multimorbidity and suicide rates are rising simultaneously among Korean adults. To address this issue, we assessed the association between multimorbidity and suicidal behavior among adults aged ≥19 years in Korea. We analyzed the data from the 2017 Korea National Health and Nutrition Examination Survey. Multimorbidity was defined as experiencing two or more chronic diseases. We compared the presence of suicidal thoughts and plans according to multimorbidity using chi-square test, and examined the associations between multimorbidity and suicidal thoughts and plans using multiple binary logistic regression analyses. Multimorbidity was found in 30.8% of total participants. As the number of chronic diseases increased, the percentage of thoughts and plans tended to increase ($p < 0.001$ and $p = 0.002$). Among participants with multimorbidities, 8.5% had suicidal thoughts, whereas only 3.4% without multimorbidity had such thoughts ($p < 0.001$). Participants with multimorbidity had significantly higher odds of suicidal thoughts (OR = 2.14; 95% CI = 1.54–2.97) and suicidal plans (OR = 2.01; 95% CI = 1.08–3.73) compared to those without multimorbidity after adjusting confounding variables. **Conclusion:** People with multimorbidity had a higher prevalence of suicidal thoughts and plans. Early detection of and intervention for suicidal thoughts and plans are critical for suicide prevention among people with multimorbidity.

Keywords: chronic disease; multimorbidity; suicidal thoughts; suicidal plans

1. Introduction

The National Commission on Chronic Illness in the United States defines chronic illness as permanent illnesses accompanied by disability due to complication or injury, and those requiring special training for rehabilitation, long-term protection, monitoring, and treatment [1]. These individuals require continuous and comprehensive medical intervention, and communication and self-care is necessary [2]. Multimorbidity refers to the state of having multiple chronic illnesses simultaneously, potentially mixed with acute illnesses as well [3]. Unlike the concept of co-morbidity, which refers to how the impact of indicator diseases are influenced by co-morbid illnesses, the concept of multimorbidity is firmly centered on individuals with multimorbidity [4].

There has been an increase of chronic illnesses following the extension of longevity in society. As expected, in South Korea the prevalence of chronic illnesses and multimorbidity has been shown to increase with age. An analysis of older adults with one or more chronic illnesses revealed that 70.9% of older adults with chronic illnesses have multimorbidity, reaching an average of 4.1 chronic illnesses [5]. Another study found that older patients with multimorbidity had an average of 5.1 chronic illnesses, whereas patients without multimorbidity tended to have an average of 1.6 [5]. There have been several studies on the definition of multimorbidity [3,6]. Because people with two or more chronic medical

conditions are more common than those with three or more chronic medical conditions, former status is more frequently accepted as the definition of multimorbidity [6]. Studies adopting this definition reported that multimorbidity better reflected poor quality of life, impaired functioning, and increased use of medical facilities such as emergency admissions, particularly with higher numbers of coexisting conditions [7–11].

Suicide is one of the leading causes of death worldwide [12]. Among the countries in the Organization for Economic Co-operation and Development (OECD) as of 2012, the suicide rate (per the OECD standard population of 100,000 people) in South Korea was 29.1, which is the highest out of all OECD countries (average is 12.5). The specific risk factors associated with suicide among adults include depression and other mental disorders, illnesses, loss of function, the death of a partner, and other traumatic incidents [13]. By comparison, according to a 2006 report by the World Health Organization (WHO), the general risk factors of suicide include low socioeconomic status, education level, loss of a job, social stress, mental disorders, illnesses, and chronic pain [14,15]. One study found that 34% of people with suicidal thoughts have plans for suicide, 72% of people who plan for suicide actually attempt suicide, and 26% of people who have suicidal thoughts attempt suicide impulsively without having elaborated a prior plan to do so [16]. Therefore, suicidal thoughts and suicidal plans are important when discussing suicide.

Multimorbidity is associated with suicide-related risk, and primary care and mental health clinics need to evaluate for suicide ideation for patients with multimorbidity [17]. To address the impact of these issues, we aimed to investigate the increase in suicide-related variables in multimorbidity among Korean adults. Specifically, we examined the prevalence of suicidal thoughts and plans among adults who participated in the Korea National Health and Nutrition Examination Survey (KNHANES) in 2017, depending on whether they had multimorbidity.

2. Methods

2.1. Participants and Data

Drawing on raw data from the second year of the 7th KNHANES of 2017, we initially selected a total of 8127 individuals. Among them, we excluded people aged ≤18 years (n = 1938) and those who had any missing variables (n = 395), because people under the age of 18 did not perform suicide-related questionnaire in the 7th KNHANES of 2017. Finally, 5794 adults were considered in our analyses.

The KNHANES is a cross-sectional health surveillance system that assesses the status of and trends in national health and nutrition of South Korea to identify vulnerable groups to be prioritized in health policies, as well as evaluate whether existing health policies and projects are effective. It provides data on smoking habits, drinking habits, physical activities, and obesity, as well as various other statistics required by the WHO and OECD. The KNHANES comprises a health interview survey, a nutrition survey, a health examination survey, and data on demographic characteristics, diet, and health collected through personal interviews. Physical examinations along with blood and urine sampling were carried out at a mobile examination center. A stratified, multistage probability sampling design was used to select the household units that participated in the survey [18].

Our study adhered to the principles of the Declaration of Helsinki and was approved by the Institutional Review Board of Ilsan Paik Hospital (No. ISPAIK IRB 12 February 2017). Requirement of informed consent was waived because anonymous and de-identified information was used.

2.2. Research Variables

The chronic illnesses surveyed in the 7th KNHANES included hypertension, dyslipidemia, cerebral infarction (stroke), myocardial infarction, angina, osteoarthritis, rheumatoid arthritis, osteoporosis, tuberculosis, asthma, allergic rhinitis, depression, renal failure, atopic dermatitis, diabetes mellitus, thyroid disease, stomach cancer, liver cancer, colon cancer, breast cancer, cervical cancer, lung cancer, thyroid cancer, other cancers, hepatitis B, hepatitis C, and cirrhosis. Participants who self-reported

not currently having a disease or a doctor's diagnosis of a disease were regarded as not having it; by contrast, participants who self-reported having a disease or a doctor's diagnosis of one were defined as having a chronic illness. Of these, stomach cancer, liver cancer, colon cancer, breast cancer, cervical cancer, lung cancer, thyroid cancer, and other cancers were combined into an overarching "cancer" category.

In the mental health section of surveys, participants were asked to answer the following questions (employing answer options of yes, no, or I do not know/no answer): "Have you seriously considered suicide in the last year?" and "Have you seriously planed suicide in the last year?"

People who had smoked at least 100 cigarettes in their lifetime and continued smoking at the time of survey were defined as current smokers. People who consumed at least one alcoholic beverage per month in a year prior to the survey were considered alcohol drinkers and those who did not were considered non-drinkers. Education level was categorized into two groups depending on whether participants had achieved a higher qualification than middle school graduation or not. Monthly personal income level was divided into two groups: the lowest quartile group and the second-lowest to highest quartile group.

2.3. Analysis Method

We combined the raw data from the 2017 KNHANES according to the KNHANES raw data analysis guidelines. Based also on a complex sample design, we conducted all analyses by assigning a dispersed stratification estimation, stratification variables, and weighted sample values. We divided participants who had 2 or more of the above-stated chronic illnesses as multimorbidity based on the definition used in the previous study [6]. Subsequently, a chi-square test was performed to evaluate proportions of suicidal thoughts and plan according to the number of chronic illness or the presence of multimorbidity. Then, to determine the association of suicidal thoughts and plans with multimorbidity, we conducted a multiple binary logistic regression analysis and calculated the odd ratio (OR) and 95% confidence intervals (CIs). Model 1 did not adjust. Model 2 adjusted for sex and age. Model 3 adjusted for sex, age, education, personal income, smoking status, and alcohol consumption. Since there are studies showing that income and education level are related the risk of suicide, we adjusted these variables [14,15]. Stratified analyses according to sex and age were also performed. All analyses were performed using SPSS Statistics 21.0 (IBM Corp., Armonk, NY, USA).

3. Results

3.1. Participants' General Characteristics

Participants' demographics are shown in Table 1. A total of 5794 adults were included. Of the entire sample, 30.8% had multimorbidity. The proportion of patients with multimorbidity increased with age. The proportion of women with multimorbidity was higher than that of men. In addition, the ratio of people with multimorbidity was higher when education level or income was low. Regarding smoking, the percentage of patients with multimorbidity was the highest among ex-smokers. The proportion of patients with multimorbidity was higher in non-drinkers than in drinkers.

Table 1. General characteristics of participants. Data are presented as % (standard error (SE)).

		Without Multimorbidity (n = 4007)		Multimorbidity (n = 1787)		p-Value
		n	% (SE)	n	% (SE)	
Age (years)	19–40	1531	91.7 (0.9)	135	8.3 (0.9)	<0.001
	41–64	1890	73.4 (1.0)	746	26.6 (1.0)	
	≥65	586	39.6 (1.5)	906	60.4 (1.5)	

Table 1. *Cont.*

		Without Multimorbidity (n = 4007)		Multimorbidity (n = 1787)		p-Value
		n	% (SE)	n	% (SE)	
Sex	Male	1836	77.2 (1.0)	723	22.8 (1.0)	<0.001
	Female	2171	72.4 (1.1)	1064	27.6 (1.1)	
Personal income	≤1st quartile	536	52.9 (1.7)	597	47.1 (1.7)	<0.001
	≥2nd quartile	3471	78.8 (0.8)	1190	21.2 (0.8)	
Education (years)	≤9	843	49.7 (1.4)	973	50.3 (1.4)	<0.001
	≥10	3164	82.4 (0.8)	814	17.6 (0.8)	
Smoking status	Never smoker	2491	74.0 (1.1)	1151	26.0 (1.1)	<0.001
	Former smoker	737	70.5 (1.6)	391	29.5 (1.6)	
	Current smoker	779	81.1 (1.4)	245	18.9 (1.4)	
Alcohol consumption	Non-drinker	1605	67.4 (1.2)	1019	32.6 (1.2)	<0.001
	Alcohol drinker	2402	79.9 (0.9)	768	20.1 (0.9)	

3.2. Suicidal Thoughts and Plans According to the Number of Chronic Diseases

Regarding suicidal ideation, 5.0% of participants had suicidal thoughts and 1.3% had suicidal plans. In the absence of chronic diseases, suicidal thoughts were present in 2.7% of the cases, and suicidal plans were present in 0.6%. If the patients had one chronic disease, suicidal thoughts were present in 4.5% of the cases, and suicidal plans were present in 1.6%. Suicidal thoughts and plans were present in 9.7%, and 2.2% of participants who had more than 3 chronic diseases, respectively. As the number of chronic diseases increased, the proportions of suicide thoughts and plans tended to increase ($p < 0.001$ and $p = 0.002$) (Table 2).

Table 2. Suicidal thoughts and plans according to the number of chronic diseases. Data are presented as % (standard error (SE)). * Obtained by using chi-square test.

		Number of Chronic Diseases				
		0 (n = 2348)	1 (n = 1659)	2 (n = 943)	≥3 (n = 844)	p-Value *
Suicidal thoughts (−)	n	2279	1589	877	761	<0.001
	% (SE)	97.3 (0.4)	95.5 (0.7)	92.5 (1.3)	90.3 (1.3)	
Suicidal thoughts (+)	n	69	70	66	83	
	% (SE)	2.7 (0.4)	4.5 (0.7)	7.5 (1.3)	9.7 (1.3)	
Suicidal plans (−)	n	2332	1637	925	824	0.002
	% (SE)	99.4 (0.2)	98.4 (0.4)	98.0 (0.5)	97.8 (0.5)	
Suicidal plans (+)	n	16	22	18	20	
	% (SE)	0.6 (0.2)	1.6 (0.4)	2.0 (0.5)	2.2 (0.5)	

3.3. Suicidal Thoughts and Plans According to the Presence of Multimorbidity

Table 3 shows the relationship between multimorbidity and suicidal thoughts and plans. While 3.4% of adults without multimorbidity were found to have suicidal thoughts, 8.5% of adults with multimorbidity had such thoughts ($p < 0.001$). As for suicide plans, 2.1% of adults with multimorbidity and 1.0% without multimorbidity were found to have them, respectively ($p = 0.003$).

Table 3. Suicidal thoughts and plans among participants with and without multimorbidities. Values are presented as weighted % (standard error (SE)). * Obtained by using chi-square test.

	Without Multimorbidity (n = 4007)		Multimorbidity (n = 1787)		p-Value *
	n	% (SE)	n	% (SE)	
Suicidal thoughts (−)	3868	96.6 (0.4)	1638	91.5 (1.0)	<0.001
Suicidal thoughts (+)	139	3.4 (0.4)	149	8.5 (1.0)	
Suicidal plans (−)	3969	99.0 (0.2)	1749	97.9 (0.4)	0.003
Suicidal plans (+)	38	1.0 (0.2)	38	2.1 (0.4)	

3.4. Multivariable Logistic Regression Analysis Between Multimorbidity and Suicidal Thoughts and Plans

Table 4 presents multivariable logistic regression analysis results regarding the association of multimorbidity with suicidal thoughts and plans. Compared to participants without multimorbidity, participants with multimorbidity had significantly higher odds of suicidal thoughts in all adjusted models (OR, 95% CI: 2.65, 1.95–3.60 in model 1; 2.49, 1.82–3.40 in model 2; 2.14, 1.54–2.97 in model 3). Stratified analysis with sex and age showed similar findings. In men, the odds of suicide thoughts increased by 3.1 times among people with multimorbidity compared to those without multimorbidity after adjusting for all potential variables (OR, 95% CI: 3.09, 1.84–5.20 in model 3). The odds of suicide thoughts increased by 3.2 times in younger people (19–40 years) with multimorbidity compared with those without it (OR, 95% CI: 3.15, 1.43–6.97 in model 3). For suicidal plans, compared to participants without multimorbidity, participants with it had higher chances of suicidal plans (OR, 95% CI: 2.14, 1.30–3.54 in model 1; 2.26, 1.22–4.20 in model 2; 2.01, 1.08–3.73 in model 3). Suicidal plans were 2.3 times more prevalent in women with multimorbidity than in those without multimorbidity (OR, 95% CI: 2.30, 1.07–4.97 in model 2); however, this association was attenuated after further adjustment in model 3. Suicidal plans were 2.5 times more prevalent in middle aged people (41–64 years) with multimorbidity than in those without multimorbidity (OR, 95% CI: 2.54, 1.12–5.77 in model 3). On the other hand, there was no statistically significant association between multimorbidity and suicidal plans among younger or older men.

Table 4. Associations of multimorbidity and suicidal thoughts and plans. OR, odds ratio; CI, confidence interval. Model 1 was not adjusted. Model 2 was adjusted for sex and age. Model 3 was adjusted for sex, age, education, personal income, smoking status, and alcohol consumption. Values were calculated by multivariable logistic regression analysis.

	Suicidal Thoughts		Suicidal Plans	
	OR (95% CI)	p	OR (95% CI)	p
Total				
Model 1	2.65 (1.95–3.60)	<0.001	2.14 (1.30–3.54)	0.003
Model 2	2.49 (1.82–3.40)	<0.001	2.26 (1.22–4.20)	0.010
Model 3	2.14 (1.54–2.97)	<0.001	2.01 (1.08–3.73)	0.028
Sex				
Men				
Model 1	3.30 (2.06–5.28)	<0.001	2.33 (1.02–5.30)	0.044
Model 2	3.40 (2.05–5.65)	<0.001	2.29 (0.85–6.20)	0.101
Model 3	3.09 (1.84–5.20)	<0.001	2.00 (0.74–5.45)	0.172
Women				
Model 1	2.18 (1.57–3.04)	<0.001	1.97 (1.06–3.67)	0.032
Model 2	1.92 (1.33–2.76)	0.001	2.30 (1.07–4.97)	0.034
Model 3	1.63 (1.12–2.36)	0.011	2.01 (0.94–4.33)	0.073

Table 4. *Cont.*

	Suicidal Thoughts		Suicidal Plans	
	OR (95% CI)	*p*	OR (95% CI)	*p*
Age (years)				
19–40				
Model 1	3.15 (1.50–6.60)	0.003	2.30 (0.61–8.68)	0.220
Model 2	3.14 (1.50–6.58)	0.003	2.34 (0.62–8.77)	0.207
Model 3	3.15 (1.43–6.97)	0.005	2.04 (0.45–9.18)	0.350
41–64				
Model 1	2.58 (1.65–4.04)	<0.001	3.37 (1.48–7.68)	0.004
Model 2	2.55 (1.63–4.00)	<0.001	3.34 (1.48–7.55)	0.004
Model 3	1.83 (1.17–2.85)	0.008	2.54 (1.12–5.77)	0.026
≥65				
Model 1	1.87 (1.21–2.90)	0.005	1.16 (0.47–2.84)	0.750
Model 2	1.89 (1.21–2.96)	0.005	1.20 (0.48–2.97)	0.697
Model 3	1.89 (1.20–2.98)	0.006	1.20 (0.47–3.10)	0.700

4. Discussion

In this study, it was shown that 30.8% of the total population of the 7th KNHANES has multimorbidity with more than two chronic diseases. As other studies have shown, the prevalence of multimorbidity increases with age, [19] with 60.4% of people aged 65 and older having multimorbidity. The proportion of women found with multimorbidity is overwhelmingly higher than that of men, which is the same as in previous studies [20]. Further, this study found evidence that level of education and smoking and drinking habits are related to multimorbidity.

We verified differences in suicidal thoughts and plans according to whether participants had multimorbidity. A comparative analysis was performed on these two groups drawing on data from the 2017 waves of the 7th KNHANES. Participants who had multimorbidity had a significantly greater prevalence of suicidal thoughts compared to those without such group of diseases. It was also found that the higher the number of chronic diseases, the higher the percentage of suicidal thoughts and suicidal plans.

As life extension continues to increase in South Korea, it has become necessary to change the perception of diseases among adults. This new understanding is essential for properly managing patients with multimorbidity [3]. Multimorbidity is highly likely to occur from polypharmacy, adverse drug side effects, or competing medical recommendations [21–23]. Moreover, if a person has multimorbidity, their risk of suicide appears to increase. The concurrence of chronic illnesses can lead to a decline in functional state and increase in mortality; therefore, more emphasis must be placed on the comprehensive impact of multimorbidity on patients' overall health and quality of life [24].

When a stratified analysis was conducted by age, younger people with multimorbidity were more at risk of suicide than those without multimorbidity. In this study, the associations of multimorbidity with suicidal thoughts in older adults aged over 65 were lower than in other age groups and did not affect suicidal plans. However, the suicide rate among older adults is increasing as the population of this group increases. According to South Korea's statistics for 2014, the suicide rate is approximately 37.5 per 100,000 people aged in their 60s, 57.6 per 100,000 people aged in their 70s, and 78.6 per 100,000 people aged in their 80s, which are all much higher than the rates of 17.8 and 27.9 of people in their 20s and 30s, respectively. These rates indicate that older adults have a higher risk of suicide compared to other age groups [25]. It is also relevant to mention that suicide attempts among older adults are more likely to lead to death compared to other age groups [25,26]. However, unlike suicide in other age groups, suicide among older adults is generally not impulsive, and diverse realistic factors tend to contribute to its occurrence [26,27].

When a stratified analysis was conducted by sex, there was a significant difference of suicidal thoughts according to multimorbidity in men and women. In this study, the associations of

multimorbidity with suicidal thoughts in men are higher than in women. This shows that there is a difference in suicide between women and men, corroborating other studies [28–30]. However, suicidal plans according to multimorbidity in men and women did not significantly differ.

Although this study provides reliable basic data on multimorbidity and suicidal thoughts and plans in a representative sample of Koreans, it does not go without limitations. First, because of their nature, cross-sectional studies cannot explain causal relationships between disease status and the various suicide-related relevant variables among adults. Furthermore, we evaluated only the presence of suicidal thoughts in the last year, without considering the frequency of these thoughts. Second, because participants' disease status was recorded based on self-reports, it is possible that biases affected the data. Third, we could not perform a detailed survey of assessed suicidal thoughts and plans, as we relied on previously written questions. In addition, while we might have considered some of the factors that influence suicidal thoughts, we could not account for all confounding variables. Finally, because older adults only accounted for a small proportion of all participants in the KNHANES, there was a major difference in the number of participants between disease groups. Therefore, there is the chance of false negatives and underestimation for some age-related results.

Previous studies have analyzed suicidal thoughts in relation to chronic illnesses; however, this study is unique in verifying whether participants had suicidal thoughts and plans according to whether they had multimorbidity. The presence of suicidal thoughts is a key risk factor for death by suicide. Suicidal behavior often recur in people who have previously attempted suicide, and suicide attempts often occur in people who frequently think about suicide [31,32]. Therefore, to prevent suicide, people with suicidal thoughts must be categorized as a high-risk group for suicide attempts and be cared for accordingly. By extension, it is necessary to consider multimorbidity in suicide prevention measures, as these individuals appear at risk of suicide. Interventions by medical professionals should, therefore, consider these thoughts as reflective of a person's mental health status [33]. Four weeks before suicide death, about 50% visited medical instructions such as outpatients, inpatients, and emergency room visits, and 83% of them visited within one year [34]. Therefore, it is necessary to screen suicidal risk for high-risk patients.

In conclusion, early detection and intervention regarding suicidal thoughts is essential in suicide prevention strategies for older adults with multimorbidity. Further research is needed to determine the suicide-related risk by chronic diseases and any clusters of chronic diseases.

Author Contributions: G.E.N., and Y.H. were the principal investigators. They contributed substantially to the study design, literature search, collection and assembly of data, data analyses and data interpretation. G.E.N., Y.H., and J.H.L. wrote all drafts and the final version of the report. Y.H., and J.H.L. analysed data and created all the tables. G.E.N., Y.H., Y.-H.K., and J.H.L. contributed to the conception and design of the study, the collection and assembly of data, data analyses and data interpretation. All authors contributed to preparation of the report and approved the final version. G.E.N., and J.H.L. had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis and all authors had final responsibility for the decision to submit for publication.

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

References

1. Choi, Y.S. *Development of a Private Demonstration Model for Those with Frequent Chronic Diseases*, Kungmin Kŏn'gang Pohŏm Kongdan Kŏn'gang Pohŏm Chŏngchaek Yon'guwŏn. 2009.
2. Chung, Y.H.; Ko, S.J.; Kim, E.J. *Research on An Effective Method of Caring for Chronic Diseases*; Korea Institute for Health and Social Affairs: Sejong, Korea, 2013.
3. Van den Akker, M.; Buntinx, F.; Roos, S.; Knottnerus, J.A. Problems in determining occurrence rates of multimorbidity. *J. Clin. Epidemiol.* **2001**, *54*, 675–679. [PubMed]
4. Marengoni, A.; Winblad, B.; Karp, A.; Fratiglioni, L. Prevalence of chronic diseases and multimorbidity among the elderly population in Sweden. *Am. J. Public Health* **2008**, *98*, 1198–1200. [PubMed]
5. Chung, Y.H. Analysis of multiple chronic diseases among the elderly. *Health Welf. Issue Focus* **2013**, *196*, 1–8.

6. Fortin, M.; Bravo, G.; Hudon, C.; Vanasse, A.; Lapointe, L. Prevalence of multimorbidity among adults seen in family practice. *Ann. Fam. Med.* **2005**, *3*, 223–228. [PubMed]
7. Fortin, M.; Bravo, G.; Hudon, C.; Lapointe, L.; Dubois, M.F.; Almirall, J. Psychological distress and multimorbidity in primary care. *Ann. Fam. Med.* **2006**, *4*, 417–422.
8. Fortin, M.; Laponite, L.; Hudon, C.; VAnasse, A.; Ntetu, A.L.; Maltais, D. Multimorbidity and quality of life in primary care: A systematic review. *Health Qual. Life Outcomes* **2004**, *2*, 51. [PubMed]
9. Condelius, A.; Edberg, A.K.; Jakobsson, U.; Hallberg, I.R. Hospital admissions among people 65+ related to multimorbidity, municipal and outpatient care. *Arch. Gerontol. Geriatr.* **2008**, *46*, 41–55.
10. Bayliss, E.A.; Bayliss, M.S.; Ware, J.E.; Steiner, J.F. Predicting declines in physical function in persons with multiple chronic medical conditions: What we can learn from the medical problem list. *Health Qual. Life Outcomes* **2004**, *2*, 47.
11. Marengoni, A.; Angleman, S.; Melis, R.; Mangialasche, F.; Karp, A.; Garmen, A.; Meinow, B.; Fratiglioni, L. Aging with multimorbidity: A systematic review of the literature. *Ageing Res. Rev.* **2011**, *10*, 430–439.
12. World Health Organization. *Prevention of Suicide: Guidelines for the Formulation and Implementation of National Strategies*; World Health Organization: Geneva, Switzerland, 1996.
13. Lapierre, S.; Erlangsen, A.; Waern, M.; De Leo, D.; Oyama, H.; Scocco, P.; Gallo, J.; Szanto, K.; Conwell, Y.; Draper, B.; et al. A systematic review of elderly suicide prevention programs. *Crisis* **2011**, *32*, 88–98. [PubMed]
14. World Health Organization. *Preventing Suicide: A Resource at Work*; World Health Organization: Geneva, Switzerland, 2006.
15. Nock, M.K.; Borges, G.; Bromet, E.J.; Alonso, J.; Angermeyer, M.; Beautrais, A. Cross-national prevalence and risk factors for suicidal ideation, plans and attempts. *Br. J. Psychiatry* **2008**, *192*, 98–105. [PubMed]
16. Kessler, R.C.; Borges, G.; Walters, E.E. Prevalence of and risk factors for lifetime suicide attempts in the National Comorbidity Survey. *Arch. Gen. Psychiatry* **1999**, *56*, 617–626. [PubMed]
17. Kavalidou, K.; Smith, D.J.; O'Connor, R.C. The role of physical and mental health multimorbidity in suicidal ideation. *J. Affect. Disord.* **2016**, *209*, 80–85. [PubMed]
18. Kweon, S.; Kim, Y.; Jang, M.-J.; Kim, Y.; Kim, K.; Choi, S.; Chun, C.; Khang, Y.; Oh, K. Data resource profile: The Korea national health and nutrition examination survey (KNHANES). *Int J. Epidemiol.* **2014**, *43*, 69–77. [PubMed]
19. Fortin, M.; Stewart, M.; Poitras, M.E.; Almirall, J.; Maddocks, H. A systematic review of prevalence studies on multimorbidity: Toward a more uniform methodology. *Ann. Fam. Med.* **2012**, *10*, 142–151. [PubMed]
20. Van den Akker, M.; Buntinx, F.; Metsemakers, J.F.; Roos, S.; Knottnerus, J.A. Multimorbidity in general practice: Prevalence, incidence, and determinants of co-occurring chronic and recurrent diseases. *J. Clin. Epidemiol.* **1998**, *51*, 367–375. [PubMed]
21. Van Weel, C.; Schellevis, F.G. Comorbidity and guidelines: Conflicting interests. *Lancet* **2006**, *367*, 550–551.
22. Smith, S.M.; O'Kelly, S.; O'Dowd, T. GPs' and pharmacists' experiences of managing multimorbidity: A 'Pandora's box'. *Br. J. Gen. Pract.* **2010**, *60*, e285–e294.
23. Guthrie, B.; Payne, K.; Alderson, P.; McMurdo, M.E.; Mercer, S. Adapting clinical guidelines to take account of multimorbidity. *Br. Med. J. BMJ* **2012**, *345*, e6341.
24. Vogeli, C.; Shields, A.E.; Lee, T.A.; Gibson, T.B.; Marder, W.D.; Weiss, K.B.; Blumenthal, D. Multiple chronic conditions: Prevalence, health consequences, and implications for quality, care management, and costs. *J. Gen. Intern. Med.* **2007**, *22*, 391–395.
25. Korea, S. *Cause of Death Statistics 2014*; Statics Korea: Daejeon, Korea, 2014.
26. McIntosh, J.L. Suicide prevention in the elderly (age 65–99). *Suicide Life Threat. Behav.* **1995**, *25*, 180–192.
27. Mościcki, E.K. Epidemiology of suicide. *Int. Psychogeriatr.* **1995**, *7*, 137–148. [PubMed]
28. Lee, S.U.; Oh, I.H.; Jeon, H.J.; Roh, S. Suicide rates across income levels:retrospective cohort data on 1 million participatnts collected between 2003 and 2013 in South Korea. *J. Epidemiol.* **2017**, *27*, 258–264. [PubMed]
29. Chang, S.S.; Gunnell, D.; Sterne, J.A.; Lu, T.H.; Cheng, A.T. Was the economic crisis 1997–1998 responsible for rising suicide rates in East/Southeast Asia? A time-trend analysis for Japan, Hong Kong, South Korea, Taiwan, Singapore and Thailand. *Soc. Sci. Med.* **2009**, *68*, 1322–1331.
30. Hawton, K.; Arensman, E.; Wasserman, D.; Hultén, A.; Bille-Brahe, U.; Bjerke, T.; Crepet, P.; Deisenhammer, E.; Kerkhof, A.; De Leo, D.; et al. Relation between attempted suicide and suicide rates among young people in Europe. *J. Epidemiol. Community Health* **1998**, *52*, 191–194. [PubMed]

31. Bonner, R.L.; Rich, A.R. Toward a predictive model of suicidal ideation and behavior: Some preliminary data in college students. *Suicide Life Threat. Behav.* **1987**, *17*, 50–63.
32. Lester, D. A depression paradox theory of suicide. *Perssonal. Individ. Differ.* **1989**, *10*, 1103–1104.
33. Kim, H.S. A study on epistemology of Korean elder's suicidal thought. *J. Korean Gerontol. Soc.* **2002**, *22*, 159–172.
34. Ahmedani, B.K.; Simon, G.E.; Stewart, C.; Beck, A.; Waitzfelder, B.E.; Rossom, R.; Lynch, F.; Owen-Smith, A.; Hunkeler, E.M.; Whiteside, U.; et al. Health care contacts in the year before suicide death. *J. Gen. Intern. Med.* **2014**, *29*, 870–877.

Journal of
Clinical Medicine

MDPI

Article

Two-Dimensional Laser-Align Device for Ultrasound-Guided Injection

Eric Chiwei Shiao [1] and Po-Ling Kuo [1,2,*]

[1] Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei 10617, Taiwan
[2] Department of Electrical Engineering, Department of Physical Medicine and Rehabilitation, College of Medicine, Health Science and Wellness Center, National Taiwan University, Taipei 10617, Taiwan
* Correspondence: poling@ntu.edu.tw

Received: 30 May 2019; Accepted: 15 July 2019; Published: 18 July 2019

Abstract: Ultrasound-guided injection is a widely used technique, however, it takes substantial amounts of time for novices to master the skill. The most critical issue to improve the accuracy of the injection is to align the needle with the scan plane of the ultrasound beam and orient the needle angle after piercing skin to aim at the targeted tissue. In the present study, we developed a two-dimensional laser align device to assist physicians to accurately position the needle in the scan plane and advance it at an angle correctly pointing to the target. The device is inexpensive, light-weighted, and easy to fabricate and accommodate for any types of ultrasound probe. Statistical analysis revealed that the assistance with our device significantly reduced the successful targeting time and times of retargeting in comparison with the traditional freehand approach or only with in-plane assistance for inexperience subjects. Our results indicate that the device exhibits great potential in effectively reducing the learning time to master the skill and speeding up the procedure for ultrasound-guided injection.

Keywords: ultrasound-guided injection; laser assisted; long-axis injection

1. Introduction

As point-of-care has gained increasing attention, ultrasound-guided injection becomes an important and popular technique for physician. The real-time feature of ultrasonography allows clinicians to target tissue for treatment more accurately and efficiently. However, it takes substantial amounts of time for novices to master the skill. For example, the in-plane approach requires well-trained coordination between hand and eyes to accurately position the needle in the scan plane and advance it at an angle correctly pointing to the target. It is of great importance to develop tools that assist the novices in accessing the target more easily and hence to improve their learning curve.

Several devices have been developed to assist clinicians in aiming the target more quickly and precisely [1]. However, limitations still exist. For examples, the Infiniti Plus^{TM} needle guidance system (CIVCO Medical Solutions, Kalona, IA, USA) lacks the flexibility for direct adjustment of needle angle once the needle is inserted into the tissue [2]; the electromagnetic tracking system [3], optical tracking system [4,5], and robotic assistance system [6] are too large and expensive to be used in regular clinics. In contrast, laser assisting approaches are relatively inexpensive and compact-sized. However, most of the recently developed laser-based systems for needle guidance [7–9] only focus on the assistance in positioning the needle in the scan plane, without aiding its insertion angle.

In this research, we developed a novel two-dimensional (2D) laser-align (LA) device to help clinicians position the needle in the scan plane and at the right angle to precisely access the targeted tissue in ultrasound-guided injection. The device is inexpensive, light-weighted, and easy to fabricate and accommodate for any types of transducer. The device is switchable among three modes: 2DLA

assisting, one-dimensional (1D) in-plane LA assisting without guidance for the inserting angle of the needle, and no LA assisting, namely the traditional freehand approach. We further compared the changes in the performance of ultrasound-guided targeting for novices and experienced physicians when assisted by different modes.

2. Materials and Methods

2.1. Device Fabrication

The 2DLA device consisted of a 5 V power module containing two coin batteries, a laser module, a clip, and a transparent film printed with slanting lines corresponding to various angles for needle advancement. The laser module included two laser units and on-off and brightness were independently controlled by two sets of switch and variable resistor, respectively. This design allowed users to switch the device function among the 2D, 1D, or no LA assisting mode. The variable resistors let the user change the brightness of lasers in an environment of various darkness.

The 2DLA assistance was accomplished by composing the laser module with two miniature laser diodes and their associated lenses, such that a laser line and a laser point were projected along and on a line aligned with the scan plane of the ultrasound beam, respectively. The pointing laser unit was used to guide the needle angle when entering the tissue. This was done by matching the projected laser point with a particular mark on the surface of the syringe body, while the marks corresponding to various needle angle were determined using the method described below. The lining laser unit was converted from a pointing laser unit with a half cylinder shaped lens, which refracted a point light source into a line one. It emitted a sector plane which aligned with the scan plane of the ultrasound beam. The sector plane projected a line on the skin surface next to the transducer and the point where the line began indicated the point where the needle tip was expected to enter the skin.

The clip was custom-made by 3D printing to fit a 10 MHz ultrasonic probe (BenQ, Taipei, Taiwan). It firmly mounted the laser units onto the probe at a specific angle, such that the pointing laser always projected on the syringe surface whatever the needle angle was and provided sharp laser contrast for needle and syringe positioning. The angle was determined as follows. Given a syringe of body length L and its attached needle of length l, the syringe body should be confined within the region between the interior and exterior of two concentric circles of radius L and l, respectively, as the needle angle to the skin changes. It is assumed that the circles are centered at point p, where the needle tip was placed, as shown in Figure 1a. To ensure that the spot emitted from the pointing laser unit always projects on the syringe body as the needle angle changes, the unit could be positioned either at position a or b as shown in Figure 1a, both of which are close to the two ends of the syringe body when the needle angle was 90° to the skin, respectively, and the laser emitted from both positions is tangent to the inner circle. These setups allowed the projection to cover the whole syringe body and hence provided better spatial resolution for angle guidance.

Let the angle of the laser light with respect to the line vertical to the skin, and that of the needle to the skin be φ and θ, respectively, and the distance from the needle tip to the point where the laser light projected on the syringe body be x (Figure 1b), we have

$$\varphi_a = \sin^{-1}(l/\overline{pa}) = \sin^{-1}(l/L), \tag{1}$$

$$\varphi_b = \cos^{-1}(l/L), \tag{2}$$

$$\overline{pb} = L \times \cot\varphi = l \times L/\sqrt{(L^2 - l^2)}, \tag{3}$$

$$x = h(\sin\theta + \tan(\varphi - \theta)\cos\theta) = h\sin\varphi\sec(\varphi - \theta), \tag{4}$$

$$x' \equiv dx/d\theta = -\mathrm{h} \times \sin\varphi\tan(\varphi - \theta)\sec(\varphi - \theta), \tag{5}$$

where φ_a and φ_b represents the angle φ when the unit was set at position a and b, respectively, h denotes $\overline{pa}$ or $\overline{pb}$ in each condition, and x' is the slope of x against various θ.

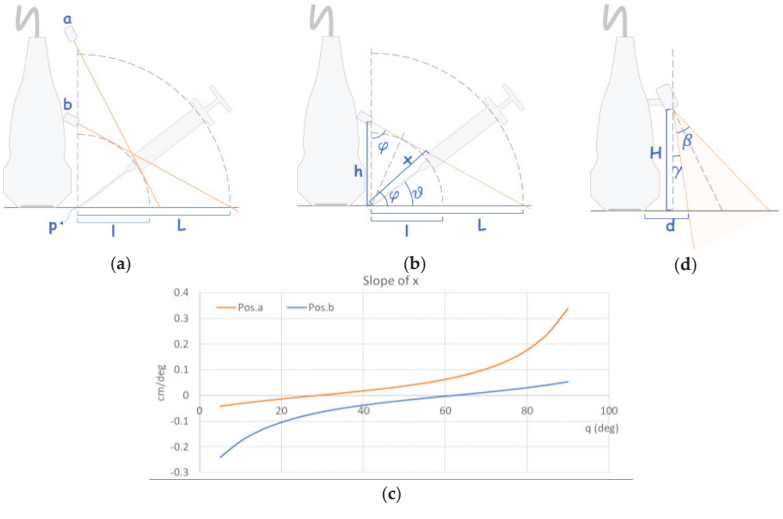


Figure 1. Geometric representation of the proposed positioning of the laser units. (**a**) Proposed setups of the pointing laser unit. (**b**) Geometric characteristics of the setup. (**c**) Profiles of slope of x against various needle angles. (**d**) Geometric layout of the lining laser unit.

Figure 1c depicts the change in x' as the needle angle varied. Note that larger x' offered finer spatial resolution for the guided points projected on the syringe body. Thus, the laser emitted from position a were projected on a wider range at larger needle angle, while that of position b was opposite. Since the needle angle is usually smaller than 45° in regular clinical practice, we chose position b to setup the pointing laser unit. The projected range of the pointing laser calculated using Equation (2) and Equation (3) would be the widest as the needle angle varied from 0° to 30°, if φ and $\overline{pb}$ were set to be 61.5° and 5.42 cm, respectively. An array of marks corresponding to various θ was determined using Equation (4) and labeled on the syringe. Thus, the needle was expected to enter the skin at a specific angle if the emitted laser spot was aligned with the mark corresponding to the angle.

The lining laser unit was set close to the pointing one to save the device size. Figure 1d demonstrates the geometric layout of the setup. The emitting angle of the laser sector plane is denoted by β, while the plane was aligned with the scan plane of the ultrasound beam and projected as a line on the surface of the skin or the syringe. The vertical distance between the unit and the skin surface is H, and d represents the distance between the end of the probe and the beginning of the projected line, which corresponds to point p in Figure 1a. Note that d is usually about 1 cm in clinical setting. Hence, the unit had to be tilted and γ represents the angle between the vertical line through the emitting point of the laser and the boundary of the sector plane, and we have

$$\gamma = \tan^{-1}(d/H). \tag{6}$$

Note that the unit was tilted by an angle of $\gamma + 0.5\beta$. We set β and γ as 52.7° and 9.46°, respectively, and H = 6 cm, such that the two units were separated by 0.58 cm to reduce the device size.

A transparent film printed with slanting lines was attached to the ultrasound monitor. Each line represented a particular angle for needle entering and had the corresponding mark labeled on the syringe. The labels included marks for various angles and a line aligned with the long axis of the syringe. One was expected to enter the needle in the scan plane at correct angle to the target; once the syringe was aligned with the projected laser line and the pointing laser was pointed to the right mark

that matching the angle determined by the slanting lines. Figure 2 illustrates the device prototype and the typical procedure of needle entering with the 2DLA assistance was demonstrated in Figure 3.

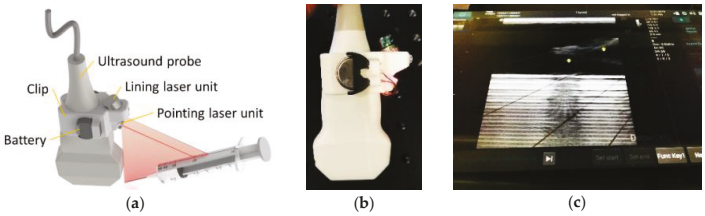


Figure 2. Schematic and photographs of the 2DLA device. (**a**) Diagram simulating the device in work. (**b**) A device assembled with an ultrasound probe. (**c**) The transparent film printed with the slanting lines for various needle entering angle and attached to an ultrasound screen.

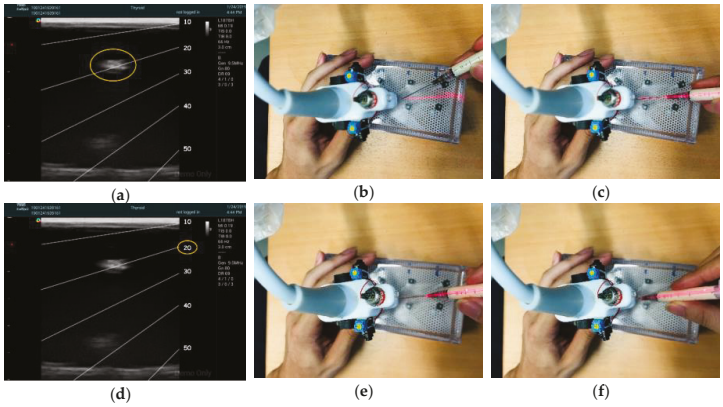
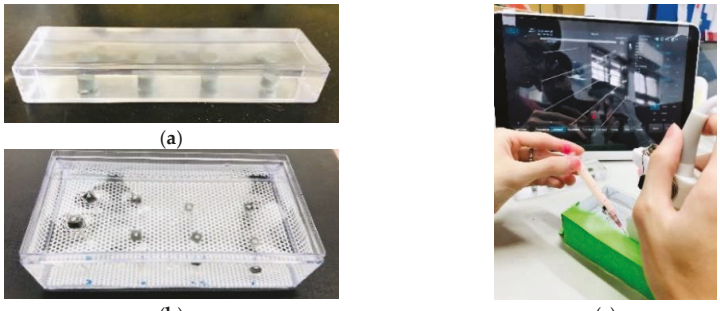


Figure 3. Procedure of needle entering guided with the 2DLA mode. (**a**) Position the probe to clearly visualize the target (highlighted by the yellow circle) on the screen. (**b**) Turn on the lasers and place the needle tip right at the beginning point of the projected laser line (the end of the laser line close to the probe). (**c**) Reorient the needle to match the projected laser line with the straight line attached on the syringe surface, which indicates the long axis of the needle. (**d**) Determine the insertion angle (highlighted by the yellow circle) using the slanting lines attached on the screen. (**e**) Adjust the entering angle to match the laser spot projected by the pointing laser unit with the mark labeled on the syringe surface corresponding to the insertion angle determined in (**d**). (**f**) Enter the needle into the phantom with persistent needle orientation until the needle tip touches the target on the screen.

2.2. Device Performance

To evaluate the effectiveness of our device, twenty subjects were enrolled to perform ultrasound-guided targeting in phantoms. Six of the subjects were physicians experienced in ultrasound-guided injection and the other fourteen subjects had little experience in ultrasound imaging before this study.

The phantoms were mainly composed of liquid paraffin and thermoplastic styrene-butadiene rubber which allowed the phantoms being formed and reshaped easily. The thermal-cured phantoms were transparent, excellent for ultrasound transmission, and possessed elasticity similar to that of soft tissues. The transparency of the phantoms allowed direct visualization of the needle advancement during practice and the performance test. Blocks made of poly (dimethylsiloxane) and iron powders with high echogenicity were embedded in the phantoms to simulate targets. There were two types of phantoms used in this study (Figure 4a,b); one was for practicing and had targets embedded at the same distance to the phantom surface; and the other was made for the performance test and the distance of the simulated targets to the phantom surface varied from 1 to 2 cm.



(a) (b) (c)

Figure 4. Photographs of the two types of phantom and its use in the performance test. Phantoms for (**a**) practicing, (**b**) performance test for ultrasound-guided targeting and (**c**) A subject performed ultrasound-guided targeting with a masked phantom.

The performance test was conducted by asking all participants to needle three targets of different depths and each needling was guided with one of the three assisting modes in randomized order. The order of the target depths was randomized as well and the variation of target depth was to prevent subjects from being familiar with the target depth from the preceding needling. In the 2DLA mode, both the lining and pointing lasers were turned on; the subject was instructed to locate the target for injection on the B-mode image, place the needle tip at the beginning end of the projected laser line, align the needle and overlap the long-axis line labeled on the syringe body with the laser line, estimate the needle angle to the target using the slanting lines on the transparent film, match the mark on the syringe body corresponding to the angle with the laser point, and advance the needle until the tip touched the target and complete the task. The procedure in the 1DLA condition was similar except that the pointing laser was turned off and the subjects had to estimate the needle angle by themselves, while the laser units were completely turned off in the no LA assisting, freehand (FH) condition. Before the performance test, the participants had unlimited time to practice targeting in the three modes. In the performance test, the phantom was masked on the side facing the subject to prevent direct visualization of the needle advancement by the subject, while the opposite side of the phantom remained transparent for video recording (Figure 4c). For each assigned target, the subject was allowed to withdraw the needle unlimited times to adjust the needle angle until successful targeting, which was considered as the needle tip penetrated the target with the presence of the entire inserted needle on the screen. Performance was evaluated by the total duration spent from entering the needle to the phantom to the moment when a successful targeting was achieved; the number of needle withdrawals until successful targeting; and the last duration, defined as the time spent from the last withdrawal to the successful targeting, which was equal to the total time if there was no needle withdrawal.

2.3. Statistics

Statistical analysis was performed using IBM SPSS Statistics 25. The differences of the three performance variables grouped with respect to different subjects, task order, and target depths were analyzed with Kruskal-Wallis one-way ANOVA test. The differences of the performance conducted in the three modes were examined using Mann-Whitney U test. Significance was set as $p < 0.05$.

3. Results

We first asked whether the performance of individuals of the same experience level was significantly different, and whether the performance was affected by the task order and the target depths. Table 1 summarized the statistical results of the performance data (Table A1) grouped against different subjects, task order, and target depths, while the data were categorized based on the level of experience. Given the subjects of the same experience level, it appears that there was no significant difference between the

averaged total durations they spent to finish the three tasks, the number of needle withdrawals, and the last durations to achieve successful targeting. Likewise, the performance was not significantly different whether the task was the first or last conducted, or the target was positioned more superficially or deeper. Detailed data were listed in the Appendix A.

Table 1. Significances of differences between the means of the performance data categorized with respect to different subjects, task orders, and target depths. * $p < 0.05$.

	Group	Subjects	Orders	Target depths
Total Duration	Inexperienced	35.83 ± 22.18, $p = 0.133$	35.83 ± 10.90, $p = 0.134$	35.83 ± 3.71, $p = 0.779$
	Experienced	25.56 ± 16.34, $p = 0.471$	25.56 ± 5.99, $p = 0.535$	25.56 ± 8.20, $p = 0.385$
Withdrawals	Inexperienced	3.31 ± 1.90, $p = 0.281$	3.31 ± 0.85, $p = 0.234$	3.31 ± 0.36, $p = 0.975$
	Experienced	1.67 ± 0.43, $p = 0.206$	1.67 ± 0.36, $p = 0.076$	1.67 ± 0.24, $p = 0.331$
Last Duration	Inexperienced	9.43 ± 3.72, $p = 0.120$	9.43 ± 0.78, $p = 0.582$	9.43 ± 0.53, $p = 0.883$
	Experienced	13.44 ± 4.75, $p = 0.925$	13.44 ± 2.12, $p = 0.804$	13.44 ± 5.69, $p = 0.162$

We next examined the performance differences in various assisting conditions. Figure 5 depicts the performance of the inexperienced and experienced subjects and the data were detailed in Table 2. As for the inexperienced subjects, they less frequently withdrew the needle for retargeting as more assistance was provided, while the last duration for successful targeting followed an opposite trend. As expected, they also spent significantly longer total duration in the FH condition than in the other two conditions. In contrast, the experienced subjects spent the shortest total duration and last duration in the FH condition than in the other two conditions. Detail *p*-values of comparisons between conditions were summarized in Table 3. As one may expect, the inexperienced subjects took significantly longer total duration and more frequent needle withdrawals than the experienced subjects in the FH condition. However, their performances were not significantly different when the LA assistances were employed, except that the experienced subjects spent significantly longer time in the last duration than the inexperienced ones in the 2DLA condition. Detail data of the two subject groups and the *p*-values of comparisons between them were summarized in Table 2.

Table 2. Means and standard deviations of inexperienced group and experienced group in FH, 1DLA, and 2DLA conditions with significance level of relation between two groups. * $p < 0.05$.

	Group	FH		1DLA		2DLA	
Total Duration	Inexperienced	56.21 ± 48.67	$p = 0.012$ *	29.36 ± 19.90	$p = 0.968$	21.93 ± 10.91	$p = 0.207$
	Experienced	9.33 ± 6.60		29.50 ± 23.24		37.83 ± 25.90	
Withdrawals	Inexperienced	5.79 ± 4.36	$p = 0.026$ *	2.64 ± 1.39	$p = 0.274$	1.50 ± 0.63	$p = 0.718$
	Experienced	1.83 ± 0.69		1.83 ± 0.69		1.33 ± 0.47	
Last Duration	Inexperienced	5.86 ± 3.60	$p = 0.179$	9.14 ± 4.73	$p = 0.312$	13.29 ± 5.23	$p = 0.050$ *
	Experienced	3.33 ± 1.25		13.50 ± 7.46		23.50 ± 11.54	

Table 3. Significance of difference in the performance of the inexperienced and experienced subjects compared between various assistance conditions. * $p < 0.05$, ** $p < 0.01$.

	Group	FH–1DLA	FH–2DLA	1DLA–2DLA
Total Duration	Inexperienced	$p = 0.210$	$p = 0.039$ *	$p = 0.541$
	Experienced	$p = 0.026$ *	$p = 0.015$ *	$p = 0.485$
Withdrawals	Inexperienced	$p = 0.019$ *	$p = 0.001$ **	$p = 0.039$ *
	Experienced	$p = 1.000$	$p = 0.310$	$p = 0.310$
Last Duration	Inexperienced	$p = 0.039$ *	$p < 0.001$ **	$p = 0.044$ *
	Experienced	$p = 0.002$ **	$p = 0.002$ **	$p = 0.093$

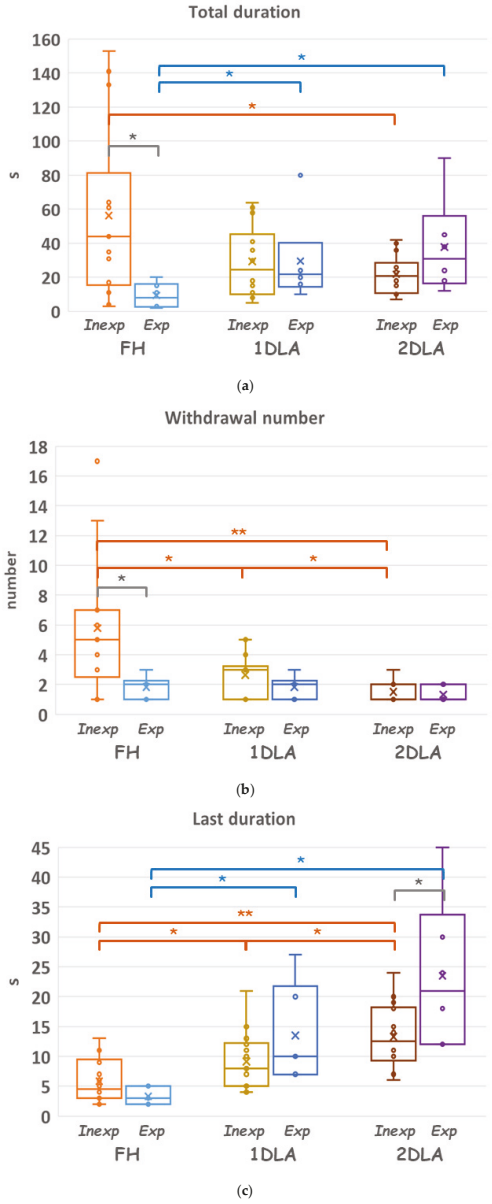


Figure 5. Comparisons of the performance of the subjects assisted by the three different modes. Data of (**a**) total duration, (**b**) withdrawal number, and (**c**) last duration of the inexperienced and experienced subjects performing in the freehand (FH), assisted with the lining laser (1DLA), and assisted with lining plus pointing laser (2DLA) condition. * $p < 0.05$, ** $p < 0.01$.

4. Discussion

In the present work, we showed that for people new to ultrasound-guided injection, our novel device did help their performance in reducing the spent time and number of needle withdrawals for successful targeting. These results agree well with those reported previously [7–9], and we further

demonstrated that our novel 2DLA design significantly helped the inexperienced operators to perform better than in the traditional freehand condition or simply assisted by a lining laser. By following the guidance provided by the pointing laser and the marks on the syringe, most of the inexperienced participants could pierce the target in one shot. This is reasonable since the 2DLA mode provided a vast amount of information for needle operation such that the subject's expertise in ultrasound-guided injection had a minor role. This may explain the fact that most of the performance of the inexperienced and experienced subjects was not significantly distinguishable when assisted in the two LA modes. Indeed, the experienced subjects even spent much longer time to achieve the successful targeting in the two LA modes than that in the FH condition, since they already had good eye-hand coordination and individual tempo for injection and may have been confused with the laser assistance. Furthermore, it may be cumbersome to check many parameters before needle advancement in the 2DLA mode, which may explain why it took longer time for the last successful targeting in both inexperienced and experienced groups. However, we believe this drawback can be overcome as the operators become familiar with the device. Nevertheless, this mode still improved the efficacy of the inexperienced operators as the total time was less, owing to the much fewer withdrawals. The results shown in Table 1 indicate that the randomly assigned targets of different depths were successful to prevent the participants from predicting the needle angle using the experience learned in preceding tests. Furthermore, these results suggest that the task difficulty perceived by the participants was not dependent on the target depth, nor the task order. Collectively, our results indicate that the 2DLA device has a high potential to speed up the learning curve of novices in ultrasound-guided injection.

With 5.5 × 5.0 × 2.8 cm in size, 32 g in weight, and a price of about $20, our 2DLA device was compact, lightweight, inexpensive, easy to mount onto the probe, and provided immediate guidance if the operators need to change the target suddenly. Unlike the device designed by Daehee et al. [10], there is no direct contact between our device and patients' skin or needle, which reduced the risk of infection. These features render our device competitive in clinical use or medial training when compared with commercially available products. For example, although the Infiniti Plus^{TM} needle guidance system provided by the CIVCO company has very affordable price (~$10 for each set), it prohibits repeated adjustment of the needle orientation for injection of multiple targets, or immediate replacement of syringe of various size in conditions such as the combined operation of ultrasound-guided aspiration and injection [11]. The eTRAX needle tip tracking guidance system, which is also from the CIVCO, embeds an electromagnetic sensor in the needle tip to provide needle tracking, but it costs more than 2 thousand dollars and is roughly forty times larger than ours in volume [12]. The SCENERGY system from the Clear Guide Medical attaches stereocameras to the ultrasound probe and integrates the ultrasound image with the CT/MRI image to improve needle navigation, but the assistive device is much larger than ours, and it requires attachment of labels on the subject's skin for recognition, which may increase the risk of infection [13].

However, there are limitations in the present study. Although nonparametric tests were applied to make our results statistically stricter, the enrollment size was small. Furthermore, the marks labeled on the syringe for guidance of injection angle were calculated based on a flat skin surface, which may well approximate conditions when the radius of curvature of the skin is much larger than the probe size, such as scanning along the long axis of an arm. However, the guidance will be erroneous when entering the needle through skin surface with smaller radius of curvature, holding the ultrasound probe askew or using a curvilinear array probe. Additional calculation is required to fix the scale distortion arising in these conditions. For example, when our device is applied to a curvilinear probe, the slanting lines printed on the transparent film should be curved to match the sector image generated by the probe. Likewise, since the needle entering point is usually out of the image view when the probe is applied on a curved surface with a large curvature, such as scanning along the transverse plane of an arm, it is hard to predict the needle orientation unless the geometric condition between the probe and the entering point is given.

5. Conclusions

In summary, we developed a novel 2DLA device that provided guidance for needle orientation both for the image plane and the angle for entering. Statistical analysis revealed that the 2DLA mode significantly reduced the time for successful targeting and frequency of retargeting in comparison with the FH and 1DLA mode. Injecting quickly and accurately is one of the most challenging tasks during the training for ultrasound-guided injection. Our device exhibits great potential in effectively reducing the learning time to master the skill and speeding up the procedure.

Author Contributions: Conceptualization, E.C.S. and P.L.K.; methodology, E.C.S.; validation, P.L.K.; formal analysis, E.C.S.; investigation, E.C.S.; data curation, E.C.S.; writing—original draft preparation, E.C.S.; writing—review and editing, P.L.K.; visualization, E.C.S. and P.L.K.; supervision, P.L.K.; project administration, P.L.K.; funding acquisition, P.L.K.

Funding: This research was funded by the Ministry of Science and Technology, Taiwan, grant number 106-3114-B-038-001, 107-2321-B-038-002, 108-2321-B-038-003, and 107-2221-E-002-040; and the Ministry of Education, Taiwan, grant number 108L893902.

Acknowledgments: The authors are grateful for the valuable input about the tips of ultrasound-guided injection in clinical practices by Chueh-Hung Wu, Ming-Yen Hsiao, Shaw-Gang Shyu, and Yi-Chian Wang in National Taiwan University Hospital, Taipei, Taiwan.

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

Appendix A

Table A1. Data of each subject in inexperienced group and experienced group.

	Subject	Task	Order	Target Depth (cm)	Total Duration (s)	Withdraw Times (#)	Duration of Last Insertion (s)
Inexperienced Group	1	FH	I	1	64	4	7
		1DLA	II	1.5	15	1	15
		2DLA	III	2	10	1	10
	2	FH	I	2	61	7	11
		1DLA	II	1.5	58	5	13
		2DLA	III	1	42	2	14
	3	FH	III	2	31	3	9
		1DLA	II	1	8	1	8
		2DLA	I	1.5	26	2	11
	4	FH	III	1	44	7	2
		1DLA	I	2	36	3	5
		2DLA	II	1.5	23	2	6
	5	FH	II	2	17	5	3
		1DLA	I	1	19	3	4
		2DLA	III	1.5	11	1	11
	6	FH	I	1.5	35	6	3
		1DLA	III	1	18	3	5
		2DLA	II	2	7	1	7
	7	FH	II	1	46	4	6
		1DLA	III	1.5	5	1	5
		2DLA	I	2	10	1	10
	8	FH	II	1.5	3	1	3
		1DLA	I	2	11	1	11
		2DLA	III	1	19	1	19
	9	FH	III	1.5	4	1	4
		1DLA	II	2	37	3	7
		2DLA	I	1	15	1	15
	10	FH	II	1	11	1	11
		1DLA	III	2	8	1	8
		2DLA	I	1.5	40	2	20
	11	FH	III	1.5	133	13	5
		1DLA	I	1	64	5	21
		2DLA	II	2	24	1	24
	12	FH	I	2	141	17	2
		1DLA	III	1.5	30	4	4
		2DLA	II	1	36	3	7
	13	FH	I	2	153	7	13
		1DLA	II	1.5	61	3	12
		2DLA	III	1	18	1	18
	14	FH	III	1	44	5	3
		1DLA	I	1.5	41	3	10
		2DLA	II	2	26	2	14

Table A1. *Cont.*

	Subject	Task	Order	Target Depth (cm)	Total Duration (s)	Withdraw Times (#)	Duration of Last Insertion (s)
Experienced Group	1	FH	III	1.5	11	2	3
		1DLA	II	2	80	3	20
		2DLA	I	1	90	2	30
	2	FH	I	1.5	15	2	5
		1DLA	II	1	24	2	10
		2DLA	III	2	12	1	12
	3	FH	II	2	20	3	5
		1DLA	III	1.5	16	2	7
		2DLA	I	1	45	1	45
	4	FH	I	1	3	1	3
		1DLA	II	1.5	20	2	7
		2DLA	III	2	38	2	12
	5	FH	II	2	2	1	2
		1DLA	I	1	10	1	10
		2DLA	III	1.5	18	1	18
	6	FH	III	1.5	5	2	2
		1DLA	II	1	27	1	27
		2DLA	I	2	24	1	24

References

1. Scholten, H.; Pourtaherian, A.; Mihajlovic, N.; Korsten, H.; Bouwman, R.A. Improving needle tip identification during ultrasound-guided procedures in anaesthetic practice. *Anaesthesia* **2017**, *72*, 889–904. [CrossRef] [PubMed]
2. Infiniti Plus™. In-Plane Ultrasound Needle Guides. Available online: https://www.civco.com/catalog/ultrasound-needle-guides/infiniti-plus-ultrasound-needle-guides/ (accessed on 16 July 2019).
3. Ungi, T.; Sargent, D.; Moult, E.; Lasso, A.; Pinter, C.; McGraw, R.C.; Fichtinger, G. Perk tutor: An open-source training platform for ultrasound-guided needle insertions. *IEEE Trans. Biomed. Eng.* **2012**, *59*, 3475–3481. [CrossRef] [PubMed]
4. Najafi, M.; Abolmaesumi, P.; Rohling, R. Single-camera closed-form real-time needle tracking for ultrasound-guided needle insertion. *Ultrasoun. Med. Biol.* **2015**, *41*, 2663–2676. [CrossRef] [PubMed]
5. Hoßbach, M.; Noll, M.; Wesarg, S. Simplified stereo-optical ultrasound plane calibration. *Int. Soc. Opt. Photonics* **2013**. [CrossRef]
6. Hemmerling, T.M.; Taddei, R.; Wehbe, M.; Cyr, S.; Zaouter, C.; Morse, J. Technical communication: First robotic ultrasound-guided nerve blocks in humans using the Magellan system. *Anesth. Analg.* **2013**, *116*, 491–494. [CrossRef] [PubMed]
7. Collins, G.B.; Fanou, E.-M.; Young, J.; Bhogal, P. A comparison of free-hand vs laser-guided long-axis ultrasound techniques in novice users. *Br. J. Radiol.* **2013**, *86*, 20130026. [CrossRef] [PubMed]
8. Tsui, B.C.H. Facilitating needle alignment in-plane to an ultrasound beam using a portable laser unit. *Reg. Anesth. Pain Med.* **2007**, *32*, 84–88. [CrossRef] [PubMed]
9. Liu, S.-Y.; Liao, M.-Y.; Wei, T.-S.; Yang, C.-P.; Chen, L.-W.; Lin, S.-J.; Ding, Y.-S.; Chang, Y.-J. Laser assisted ultrasound guided aspiration improves procedure time and reduces number of withdrawals. *Ultrasonics* **2008**, *48*, 647–651. [CrossRef] [PubMed]
10. Daehee, H.; Lan Seo, Y.; Choi, C.S.; Chul Kim, H.; Yoon, D.Y.; Bae, S.; Moon, J.; Kim, S.H.; Soo Kim, S.; Han, H. A steerable guiding device: The new method in ultrasound guidance. *Investig. Radiol.* **2002**, *37*, 626–631. [CrossRef]
11. Infiniti Plus™. Disposable Needle Guide Replacement Kits. Available online: https://www.civco.com/catalog/ultrasound-needle-guides/infiniti-plus-disposable-needle-guide-replacement-kits/ (accessed on 25 June 2019).
12. eTRAX. Needle Tip Tracking Guidance System—12GA. Available online: https://www.civco.com/catalog/fusion-tracking-technology/etrax-needle-tip-tracking-guidance-12ga/ (accessed on 25 June 2019).
13. Clear Guide Scenergy. Available online: https://clearguidemedical.com/ (accessed on 25 June 2019).

Journal of
Clinical Medicine

MDPI

Article

Comparison of Oncologic Outcomes in Laparoscopic versus Open Surgery for Non-Metastatic Colorectal Cancer: Personal Experience in a Single Institution

Chong-Chi Chiu [1,2,3], Wen-Li Lin [4], Hon-Yi Shi [5,6,7], Chien-Cheng Huang [8,9], Jyh-Jou Chen [10], Shih-Bin Su [11,12,13], Chih-Cheng Lai [14], Chien-Ming Chao [14], Chao-Jung Tsao [15], Shang-Hung Chen [16] and Jhi-Joung Wang [17,18,*]

1 Department of General Surgery, Chi Mei Medical Center, Liouying 73657, Taiwan; chiuchongchi@yahoo.com.tw
2 Department of General Surgery, Chi Mei Medical Center, Tainan 71004, Taiwan
3 Department of Electrical Engineering, Southern Taiwan University of Science and Technology, Tainan 71005, Taiwan
4 Department of Cancer Center, Chi Mei Medical Center, Liouying 73657, Taiwan; wenlilin2012@gmail.com
5 Department of Healthcare Administration and Medical Informatics, Kaohsiung Medical University, Kaohsiung 80708, Taiwan; hshi@kmu.edu.tw
6 Department of Business Management, National Sun Yat Sen University, Kaohsiung 80424, Taiwan
7 Department of Medical Research, Kaohsiung Medical University Hospital, Kaohsiung 80708, Taiwan
8 Department of Emergency Medicine, Chi-Mei Medical Center, Tainan 71004, Taiwan; chienchenghuang@yahoo.com.tw
9 Department of Senior Services, Southern Taiwan University of Science and Technology, Tainan 71005, Taiwan
10 Department of Gastroenterology and Hepatology, Chi Mei Medical Center, Liouying 73657, Taiwan; jjchen@mail.chimei.org.tw
11 Department of Occupational Medicine, Chi Mei Medical Center, Liouying 73657, Taiwan; shihbin.su@msa.hinet.net
12 Department of Occupational Medicine, Chi Mei Medical Center, Tainan 71004, Taiwan
13 Department of Leisure, Recreation and Tourism Management, Southern Taiwan University of Science and Technology, Tainan 71005, Taiwan
14 Department of Intensive Care Medicine, Chi Mei Medical Center, Liouying 73657, Taiwan; dtmed141@gmail.com (C.-C.L.); ccm870958@yahoo.com.tw (C.-M.C.)
15 Department of Oncology, Chi Mei Medical Center, Liouying 73657, Taiwan; m961193@mail.chimei.org.tw
16 National Institute of Cancer Research, National Health Research Institutes, Tainan 70403, Taiwan; bryanchen@nhri.edu.tw
17 Department of Medical Research, Chi Mei Medical Center, Tainan 71004, Taiwan
18 AI Biomed Center, Southern Taiwan University of Science and Technology, Tainan 71005, Taiwan
* Correspondence: 400002@mail.chimei.org.tw; Tel.: +886-6-2517844; Fax: +886-6-2832639

Received: 30 April 2019; Accepted: 17 June 2019; Published: 19 June 2019

Abstract: The oncologic merits of the laparoscopic technique for colorectal cancer surgery remain debatable. Eligible patients with non-metastatic colorectal cancer who were scheduled for an elective resection by one surgeon in a medical institution were randomized to either laparoscopic or open surgery. During this period, a total of 188 patients received laparoscopic surgery and the other 163 patients received the open approach. The primary endpoint was cancer-free five-year survival after operative treatment, and the secondary endpoint was the tumor recurrence incidence. Besides, surgical complications were also compared. There was no statistically significant difference between open and laparoscopic groups regarding the average number of lymph nodes dissected, ileus, anastomosis leakage, overall mortality rate, cancer recurrence rate, or cancer-free five-year survival. Even though performing a laparoscopic approach used a significantly longer operation time, this technique was more effective for colorectal cancer treatment in terms of shorter hospital stay and less blood loss. Meanwhile, fewer patients receiving the laparoscopic approach developed postoperative urinary tract infection, wound infection, or pneumonia, which reached statistical significance. For non-metastatic

colorectal cancer patients, laparoscopic surgery resulted in better short-term outcomes, whether in several surgical complications and intra-operative blood loss. Though there was no significant statistical difference in terms of cancer-free five-year survival and tumor recurrence, it is strongly recommended that patients undergo laparoscopic surgery if not contraindicated.

Keywords: laparoscopic; open surgery; non-metastatic colorectal cancer; surgical complication; oncologic outcome; single surgeon experience

1. Introduction

Since the first laparoscopic-assisted colon resection introduced in 1991 by Jacobs et al., it has gradually become popular [1]. Increasingly more colorectal surgeons admit that the laparoscopic technique leads to quicker functional recovery [2–5] and improved short-term results when compared with the open approach [6–12]. However, the laparoscopic technique has not previously been proven to gain significant benefits in colorectal surgeries [13–17]. Recently, oncologic outcomes of colorectal cancer resection, in terms of lymph node harvest number and excision safety margin lengths, achieved under laparoscopy could be comparable to those obtained using the conventional open technique. However, the curability of colorectal cancer under the laparoscopic technique remains controversial because of the uncertainty about the overall recurrence rate [18]. Besides, three principal, randomized clinical trials have proven that the laparoscopic technique can lead to the same oncological outcomes related to an open approach, but did not distinguish a survival benefit favoring laparoscopy [2,4,6].

It is believed that the role of the laparoscopic technique for advanced non-metastatic colorectal cancer management will be clarified through this study. The aim of this research was a comparison of surgical complications and five-year oncologic results of non-metastatic colorectal cancer patients receiving laparoscopic resection (LR) or open resection (OR) by one surgeon in a medical institution.

2. Materials and Methods

2.1. Ethics Statement

The institutional Ethics of Research Committee of Chi Mei Medical Center, Taiwan permitted this study. The protocol conformed to ethical standards according to the Declaration of Helsinki published in 1964. Moreover, written or verbal consent from patients was acquired for this study.

2.2. Study Population

From January 2008 to December 2013, a total of 375 consecutive colorectal cancer patients scheduled for resection by Dr. Chiu in a regional hospital with LR or OR were assessed (Figure 1). The treatment protocol was based on the National Comprehensive Cancer Network (NCCN) Guidelines®. The exclusion criteria included patients with cancer distant metastasis, synchronous tumors, adjacent organ invasion, intestinal obstruction, combined operations for other disease, history of trans-abdominal or trans-anal colorectal surgery, history of inflammatory bowel disease, polyposis, past episode of ileus related to severe intra-abdominal adhesions, morbid obesity, severe medical disease, pregnancy, emergent surgeries, patient unwilling to participate in the study, or conversion to open approach. Conversion to open approach was defined as an abdominal incision larger than necessary for specimen retrieval. Written informed consent was obtained from all patients in this study. This study divided patients into several groups according to the tumor locations. Patients of each tumor location group were randomly allocated to receive LR or OR using the random numbers belonging to that location group in the envelopes, blindly selected by the surgeon before the operation. In the LR group, all patients needed to pay for the extra fee of the harmonic scalpel and wound retractor.

Data were collected in a prospectively maintained database that was supplemented by a retrospective chart review.

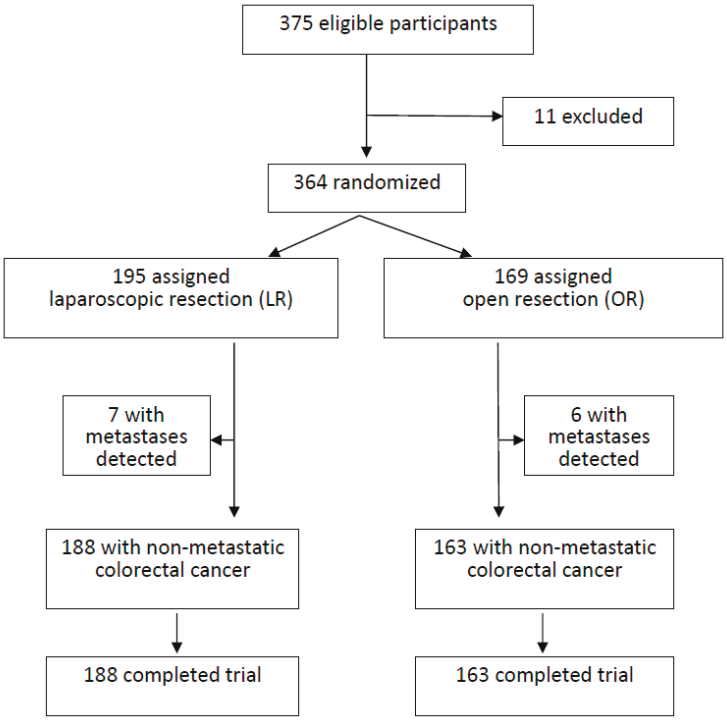


Figure 1. The flowchart of the study design.

2.3. Pre-Operative Staging Work-Up

The evaluation included physical examination, colonoscopy with biopsy, abdominal, and pelvic computed tomography (CT) scan. Pelvic magnetic resonance imaging was routinely performed for rectal cancer patients. Serum level of carcinoembryonic antigen (CEA) was sampled before the operation. The pre-operative clinical oncologic staging was classified by tumor node metastasis (TNM) system of the American Joint Committee on Cancer (AJCC)/International Union Against Cancer (UICC).

2.4. Surgical Techniques

All LR and OR procedures proceeded with a standardized medial-to-lateral approach and non-touch technique. During LR surgery, the surgeon and camera operator stood on the opposite side of the colorectal lesion, while the first assistant positioned to the same side of the lesion. Briefly, the right hemicolectomy including the range extended to the mid-transverse colon with lymphadenectomy about the ileocolic, right colic, and middle colic vessel origin was selected for proximal lesions (those sited proximal to the flexure of the spleen). The left hemicolectomy with lymphadenectomy at the level of the left colic and the left branch of the middle colic vessel origins was selected for lesions at the descending colon. The omentum was transected to allow entry into the omental bursa and mobilization of the liver flexure (right hemicolectomy) or splenic flexure (left hemicolectomy). As lesions of the sigmoid colon or rectosigmoid junction, the sigmoid colectomy with upper rectum resection and lymphadenectomy extended to the inferior mesenteric vessel origin were selected. At least 5 cm safety surgical clearance margin was mandatory for all patients. As for rectal cancer, the technique

was standardized as follows: (1) for upper third rectal lesions, a 5 cm mesorectal resection with end-to-end colorectal anastomosis was done; (2) for mid and low rectal lesions, total mesorectal excision with pouch supra-anal or anal anastomosis was performed; and (3) abdominoperineal excision was indicated once the levator muscle was involved by tumor. According to the principle of the non-touch technique, high ligation of the inferior mesenteric artery and mobilization of the splenic flexure were first systematically performed, whether the procedures were performed in LR or OR group. Dissected tissue was pulled out via a wound retractor at the extended umbilical wound for abdominal wall protection. For proximal lesions, anastomosis was routinely performed extra-corporeally. We routinely performed trans-anally intra-corporeal circular stapled anastomosis after the descending colon, sigmoid colon, or rectum lesion resection, because the residual distal intestine stump after resection was hardly managed extra-corporeally. In the OR group, the procedures were performed through a midline laparotomy with the same rules, and the wound was protected by gauze covering. The harmonic scalpel was generally used for soft tissue dissection in the LR group, but not in the OR group. Besides, in cases with difficult tumor localization by vision or palpation, an intra-operative colonoscopy would be routinely used to locate the actual tumor site instead of using other methods. However, no one in the OR group needed an intra-operative colonoscopy, but only three patients in the other group needed one.

2.5. Post-Operative Management

Post-operative treatment was the same for both groups. Patients were discharged when they had sufficient oral intake, well-controlled complications, or no complications. Complications designated as more severe than grade I according to the Clavien–Dindo classification system were categorized as ileus, urinary tract infection, wound infection, pneumonia, anastomotic leakage, and so on. Besides, most patients with stage III colorectal cancer would receive post-operative chemotherapy (oral or intravenous form), except six patients of the LR group and three of the OR group owing to general weakness or intolerance to chemotherapy.

2.6. Post-Operative Follow-Up

One specialized pathologist assessed all specimens. All patients were followed up with clinical examination, serum CEA assay, chest X-ray exam every three months, and liver ultrasound every six months for the first two years, and then annually. An abdominal CT exam was arranged annually. A colonoscopy was performed at one year after the operation, then every three years.

2.7. Statistical Analysis

The main endpoint of this study was cancer-free five-year survival. The secondary endpoint was the incidence of tumor recurrence. Predefined baseline variables are listed in Tables 1 and 2. Variables for the univariate analysis were gender, age, American Society of Anesthesiologists (ASA) class, tumor location, TNM stage, histopathology, pre-surgery serum CEA level, type of intervention, postoperative complications, and tumor recurrence. Categorical variables were compared using the χ^2 test. Continuous variables (e.g., number of lymph nodes removed, hospitalization period, intra-operative blood loss, and operation time) were compared using Student's *t*-test. Survival period was evaluated from the day of surgery to the last visit or death. For cancer-free survival, patients dying from other causes were censored at the time of death. Probability curves were constructed according to the Kaplan–Meier method and compared with the log-rank test (Table 3) (Figure 2). A *p*-value less than 0.05 was regarded as statistically significant. All calculations were performed using the SPSS software package version 20 (SPSS Inc., Chicago, IL, USA).

Table 1. Comparison of baseline characteristics between laparoscopic resection (LR) versus open resection (OR).

Items	LR (n = 188)	OR (n = 163)	p-Value
Gender			0.435
Male	102	87	
Female	86	76	
Age (mean ± SD)	68.6 ± 12.7	71.5 ± 12.1	0.23
ASA class			0.698
I	113	92	
II	75	71	
TNM stage (clinical/radiologic)			0.345
0	3	0	
I	70	55	
II	26	26	
III	89	82	
TNM stage (pathologic)			0.344
0	3	0	
I	68	53	
II	30	29	
III	87	81	
Histopathology			0.624
Well differentiated	92	79	
Moderate differentiated	71	65	
Poorly differentiated	25	19	
Tumor location			0.431
Cecum	29	20	
Ascending colon	41	36	
Transverse colon	12	15	
Descending colon	21	18	
Sigmoid colon	52	45	
Rectum	33	29	
Intervention			0.720
Right hemicolectomy	65	56	
Left hemicolectomy	24	21	
Transverse colectomy	14	10	
Sigmoid colectomy	50	44	
Protectomy	32	27	
Abdominal perineal resection	3	5	
Protective diversional stoma	13	15	
Pre-surgery			1.000
serum CEA level			
<5	24	13	
≥5	164	150	

ASA—American Society of Anesthesiologists; TNM—tumor node metastasis; CEA—carcinoembryonic antigen.

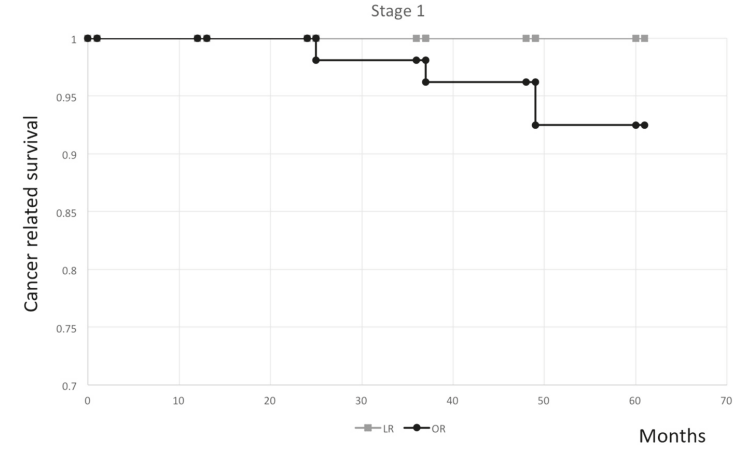
Table 2. Comparison of surgical outcomes between laparoscopic resection (LR) versus open resection (OR).

Items	LR (n = 188)	OR (n = 163)	p-Value
Tumor recurrence	17 (9.0%)	22 (13.5%)	0.186
Lymph nodes removed	16.0 ± 9.2	19.2 ± 13.7	0.07
Hospitalization (days)	13.2 ± 4.2	18.8 ± 9	<0.001 **
Blood loss (mL)	23.5 ± 14.6	162.2 ± 63.4	<0.001 **
Operation time (min)	191.4 ± 71.1	150.8 ± 46.3	<0.001 **
Postoperative complications			
Total	8	25	
Ileus	3	5	0.273
Urinary tract infection	1	5	<0.001 **
Wound infection	2	7	<0.001 **
Pneumonia	2	6	0.048 *
Anastomosis leakage	0	2	0.140

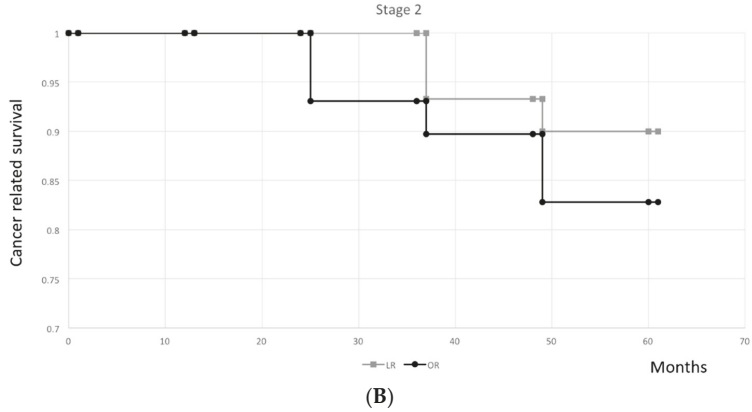
* $p \leq 0.05$ ** $p \leq 0.001$. LR Tumor recurrence: liver metastasis ×13, lung metastasis ×5, carcinomatosis ×4, anastomotic recurrence ×1, local recurrence ×3; OR Tumor recurrence: liver metastasis ×19, lung metastasis ×8, carcinomatosis ×6, anastomotic recurrence ×3, local recurrence ×2.

Table 3. Cancer-free survival rates between laparoscopic resection (LR) versus open resection (OR).

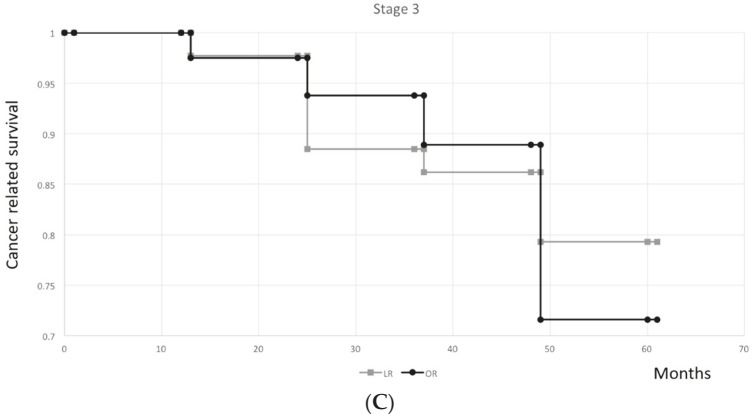
Stage	Group	n	Death		Survival					p-Value
					1st Year	2nd Year	3rd Year	4th Year	5th Year	
			n	%	% (n)	% (n)	% (n)	% (n)	% (n)	
0	LR	3	0	0	100 (3)	100 (3)	100 (3)	100 (3)	100 (3)	-
	OR	0	-	-	-	-	-	-	-	
I	LR	68	0	0	100 (68)	100 (68)	100 (68)	100 (68)	100 (68)	0.206
	OR	53	4	7.5	100 (53)	100 (53)	98.1 (52)	96.2 (51)	92.5 (49)	
II	LR	30	3	10.0	100 (30)	100 (30)	100 (30)	93.3 (28)	90.0 (27)	0.713
	OR	29	5	17.2	100 (29)	93.1 (27)	93.1 (27)	89.7 (26)	82.8 (24)	
III	LR	87	18	20.7	100 (87)	97.7 (85)	88.5 (77)	86.2 (75)	79.3 (69)	0.426
	OR	81	23	28.4	97.5 (79)	93.8 (76)	88.9 (72)	82.7 (67)	71.6 (58)	
Total	LR	188	21	11.2	100 (188)	98.9 (186)	94.7 (178)	92.6 (174)	88.8 (167)	0.328
	OR	163	32	19.6	98.8 (161)	95.7 (156)	92.6 (151)	88.3 (144)	80.3 (131)	



(**A**)

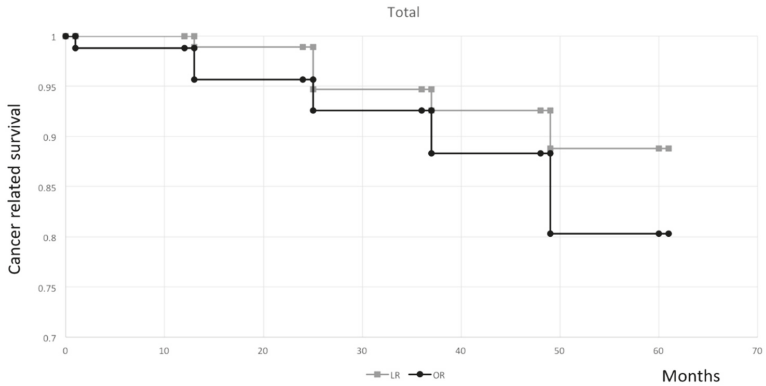


(**B**)



(**C**)

Figure 2. *Cont*.



(D)

Figure 2. (**A**) Kaplan–Meier curve of cancer-free five-year survival in stage I patients (p = 0.206); (**B**) Kaplan–Meier curve of cancer-free five-year survival in stage II patients (p = 0.713); (**C**) Kaplan–Meier curve of cancer-free five-year survival in stage III patients (p = 0.426); (**D**) Kaplan–Meier curve of cancer-free five-year survival in all stage patients (p = 0.328).

3. Results

3.1. Baseline Characteristics of Patients

The basic profile of this study is shown in Figure 1. Initially, 375 colorectal cancer patients under Dr. Chiu's service were sorted. Of these, 11 were excluded from the study. A total of 364 patients receiving curative resection were assessed in this study; 195 received LR and 169 received OR. Carcinomatosis was detected intra-operatively in seven patients of LR and six patients of OR, which were excluded. The remaining patients were compliant with the follow-up protocol. The median surveillance period was about 60 months.

In Table 1, both groups of patients were well matched in terms of demographic and clinicopathologic parameters. During this study period, 188 patients of the LR group were compared with the data obtained from the other 163 patients of the OR group. In the LR group, the mean age was 68.6 ± 12.7 years, and 102 (54.3%) patients were male. According to the final pathology report, three patients were classified in stage 0, 68 in stage I, 30 in stage II, and 87 in stage III. In the OR group, the mean age was 71.5 ± 12.1 years, and 87 (53.4%) patients were male; none were affected by tumors in stage 0, 53 in stage I, 29 in stage II, and 81 in stage III. There was a little disparity between clinical/radiologic staging and pathological staging in this study. Other characteristics of tumors and patients were summarized, and there was no statistical difference between these two groups.

3.2. Surgical Outcomes

In Table 2, the rate of tumor recurrence was 9.0% (17/188) in the LR group and 13.5% (22/163) in the OR group. Although the difference was not statistically significant, tumor recurrence seemed to be lower in the LR group (p = 0.186). The average number of lymph nodes removed in LR was 16.0 ± 9.2 and 19.2 ± 13.7 in OR (p = 0.07). Tumor margins were non-involved in patients of both groups. However, this study demonstrated that LR was more effective for the treatment of colorectal cancer in terms of hospital stay ($p < 0.001$) and blood loss ($p < 0.001$). Conversely, operation time was significantly longer in LR than in OR (191.4 ± 71.1 min vs. 150.8 ± 46.3 min, $p < 0.001$). Compared with the LR group, more patients in the OR group encountered postoperative urinary tract infection, wound infection, and pneumonia, which reached statistical significance. Only two patients in the OR group were found with mild anastomosis leakage from the drainage tube clinically. Abdominal CT confirmed the diagnosis and that the degree was mild. These two patients received conservative

treatment, including intravenous fluid supply and nil per os. However, further surgical intervention was not necessary.

3.3. Cancer-Free Survival Rates and Tumor Recurrence Incidence

In Table 3, twenty-one patients (11.2%) of the LR group and 32 patients (19.6%) of the OR group expired. There was a trend of higher overall mortality in the OR group, with 4 in stage I, 5 in stage II, and 23 in stage III, but it was not statistically different. In stage 0, there were only three patients in the LR group and none in the other. All three patients survived more than five years after surgery. In stage I, all four deaths of the OR group were non-cancer related. All patients survived for at least 30 months after surgery. In stage II, two patients in the OR group died within the second year after surgery, but they were non-cancer related. Others in both groups who died were all cancer-related. In stage III patients, all eighteen deaths in the LR group and twenty-three deaths in the OR group were cancer-related. In the OR group, two patients died fewer than six months after a second oncologic surgery for cancer recurrence, about three years after previous surgery.

There was a phenomenon of a higher cancer-free five-year survival in stage I ($p = 0.206$, Figure 2A), stage II ($p = 0.713$, Figure 2B), stage III ($p = 0.426$, Figure 2C), and all stages ($p = 0.328$, Figure 2D) in the LR group when compared with those in the OR group, although the difference was not statistically significant.

The median time for tumor recurrence was 57.0 months (range 25–68 months) in LR and 53.5 months (range 25–63 months) in OR. Importantly, no difference was observed in the cumulative incidence of recurrence between these two groups ($p = 0.186$) (Figure 3). Besides, there was no incidence of port-site recurrence in the LR group or wound recurrence in the OR group.

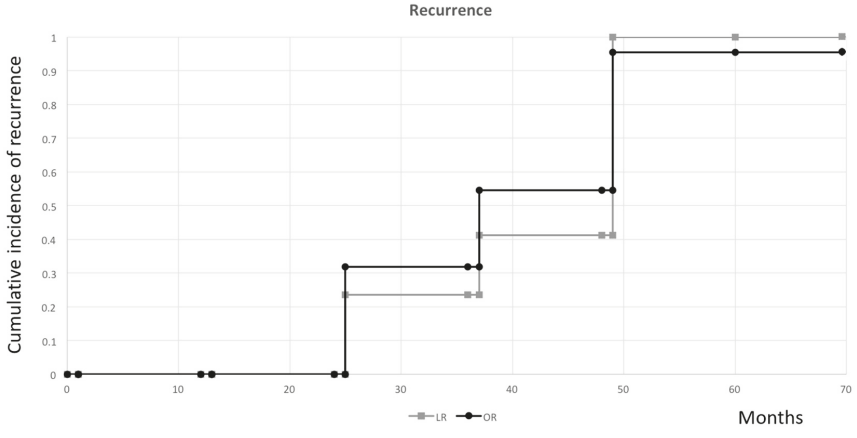


Figure 3. Cumulative incidence curve of tumor recurrence in all stage patients ($p = 0.186$).

4. Discussion

Previously, randomized controlled studies demonstrated that LR had favorable operative outcomes with less wound pain, earlier functional return of the gastrointestinal tract, a shorter hospital stay, and better cosmetics when compared with OR [19–22]. Moreover, a meta-analysis [23] and two large retrospective studies [24,25], which included a large number of patients, also showed a significant reduction in the mortality rate and lowered the morbidity after LR.

However, survival is the most crucial concern for assessing success for malignant disease treatment. This study included a 60-month follow-up and compared LR and OR for non-metastatic colorectal cancer. The results of cancer-related survival and incidence of tumor recurrence favored the LR group, despite that there was no statistically significant difference regarding the oncological results.

The Clinical Outcome of Surgical Therapy study, which was the largest randomized controlled trial conducted so far, also showed the same results as ours and even overall survival between the two groups after a median four-year follow-up [2]. However, in a single institution randomized study, Lacy et al. advocated that there was a cancer-related survival advantage after LR for stage III colon cancer patients [26]. Capussotti et al. also demonstrated that LR was related to significantly better disease-free and cancer-related survival stage III colon cancer patients [27]. Other studies have reported better survival for patients undergoing LR, even for those with stage II colorectal cancer [28].

One of the assumptions about better survival might be the number difference in dissected lymph nodes between the LR and OR groups. Laparoscopy provides better visualization of intra-abdominal conditions [29], including a more comprehensive, more precise, and brighter image to allow surgeons to perform a more radical and precise resection of the mesocolon and mesorectum, while facilitating an accurate and complete lymphadenectomy [30,31]. Complete lymphadenectomy for colorectal cancer is essential for the patient's oncological prognosis because of a reduced risk of residual nodal disease, as well as accurate nodal staging (achieving a better stratification of tumor staging) [19]. However, there was no statistical difference in the lymph node retrieval number between these two groups. Similarly, the retrieved and assessed lymph node number in many patients of both groups was higher than the threshold of 12 lymph nodes recommended by the American Joint Committee on Cancer (AJCC) in our study. Lymphadenectomy of colorectal cancer was a decisive factor for the prognostic and therapeutic staging of the patient. Different variables could affect the retrieval number of lymph nodes. Some, like the surgeon, the surgery, and the pathology exam, were without question modifiable; however, other both patient- and disease-related variables were non-modifiable and posed the question of whether the minimum number of examined lymph nodes must be individually assigned. However, since 2010, the AJCC classification subdivided patients treated for colorectal cancer into prognostic categories according to the number of metastatic lymph nodes [32]. The accuracy of the staging was influenced by the number of retrieved lymph nodes as the relationship between positive nodes divided by the total number of retrieved nodes. With regard to the prognosis prediction, this "lymph nodal ratio" is also effective in cases of reduced lymph nodal sampling [33–36]. Besides, the sentinel lymph nodes were thought to find valid application in this field in the future [37]. Of course, improvement of these modifiable factors is the only aspect that our team could strive for at this moment, as well as the opportunity to gain a better oncologic outcome of cancer remission or recurrence after reducing the risk of residual nodal disease.

Other proven benefits in oncologic results about LR include its effect on cellular immunity, intra-operative tumor manipulation, related stress response and subsequent cytokine release, surgical complication rate, and blood transfusion amount [26]. Conclusively, one of the most essentially beneficial theories of LR is regarded as the preservation of the patient's immunological response against cancer from the first postoperative days [38]. There has been significant evidence suggesting that surgical stress interferes with immunity, and this phenomenon is more apparent in OR than in LR [39]. The role of immunosuppression has been advocated because immunologic response mediators (e.g., C-reactive protein, interleukin 1–6, and tumor necrosis factor alpha) are decreased after LR in colorectal surgery compared with the OR approach. On the other hand, immunosuppression deteriorates both sepsis and cancer cell proliferation [40]. Lacy et al. have also pointed out that the post-LR stress response of colorectal cancer is less pronounced and finally leads to better preservation of cellular immune function, and attenuates inflammatory mediator interference [41,42]. Correlation of the stress response degree after the trauma of surgery with the host resistance to cancer has been proven in an animal model [26]. Immunity is a critical barrier against tumor progression and metastatic spread [39]. LR could, therefore, theoretically increase either overall or cancer-free survival. However, we should routinely examine these immunologic response mediators after surgery to improve the quality of our further study.

Tumor manipulation has been proven to contribute to cancer cell spread. There is some evidence that tumor mobilization is related to cancer cells' exfoliation into the peritoneal cavity and portal

vein bloodstream migration, which might be alleviated by non-touch surgical techniques or the avoidance of tumor manipulation. Preliminary reports have shown that cancer cell spread is not worsened [43], and dissemination of cancer cells is reduced by LR [39]. However, this phenomenon is difficult evaluated in this study based on the safety issue of blood sampling from portal vein bloodstream. Under the laparoscopic vision, limited access inside the abdominal cavity leads to minimal tumor handling and compliance of non-touch technique, both favoring the important oncology principle to avoid tumor cell spread during surgery. In this study, all patients received non-touch isolation techniques in both the LR and OR groups, which should cause no difference in prognosis. In the future, we could also perform intra-operative abdominal cavity normal saline irrigation after tumor resection to compare the difference of possible exfoliated cancer cells into the peritoneal cavity by two surgical techniques.

There is an evident statistical difference in fewer complication rates and the amount of blood loss in the LR group compared with the OR group. These factors theoretically contribute to better prognosis of tumor recurrence and cancer-free survival in LR patients. Despite that the differences regarding the oncological results did not reach the statistical significance of both groups, other better short-term outcomes, including smaller incisions, less postoperative pain, quicker functional recovery, shorter hospital stays, and earlier return to regular activity, suggested colorectal cancer patients should receive LR if not contraindicated. Meanwhile, although the operation length was longer in our LR group, the benefit of this minimally invasive technique on peri-operative care (quicker functional recovery and shorter hospital stays) further overcame this disadvantage. Besides, this benefit in shorter peri-operative care would significantly lower the total medical cost, especially in Western countries. Meanwhile, it is believed that we could set up some standard protocols and guidelines for routine use to decrease the operation time in the OR group in the future.

There are many debates about the effect of wound size. Several experienced colorectal surgeons pointed out that most colectomies could be performed with an abdominal wound of less than 7 cm, and thus opposed the wound benefits of LR [44]. However, the advantages of LR for colorectal cancer not only include a comparatively smaller wound size, but also relate to the properties of laparoscopy, especially the operation field magnification, more precise tumor resection, and its minimal invasiveness [5]. One meta-analysis including 3863 patients even showed that single-incision laparoscopic surgery (SILS) had comparable outcomes to LR in terms of operating time, conversion rate, reoperations, postoperative complications, and mortality, but only shorter mean hospital stay. There was no difference in the oncological results regarding average lymph node retrieval, adequate resection margins, survival rates, and local recurrence [45].

Application of LR for colorectal cancer encountered much criticism in the early 1990s as a result of several case reports about port site recurrence and suspicion of the adverse effect of oncologic outcome [25]. However, many surgeons advocated that LR did not aggravate cancer cell spillage intra-corporeally when surgeons strictly followed the oncologic principles [5]. However, the routine practice of the laparoscopic technique in colorectal cancer treatment is only performed in a few experienced centers in Taiwan. Localization of the tumors remains a major limiting factor of LR popularity among most surgeons, despite that there are several techniques, including conventional colonoscopy and colonoscopic tattooing, colonoscopic clip placement, radio-guided colorectal lesion localization, and the application of magnetic colonoscopic imaging [46]. Besides, some specialists are still quite hesitant about laparoscopy because of the lack of "direct" physical and visual contact of the lesion. Particularly during the LR process, the intestinal color is more difficult to assess, and direct palpation of blood vessels to the anastomosis is not possible [47]. The phenomenon of the slow popularity of this minimally invasive technique further reflects its complexity, especially at the initial stage of the learning curve; the lack of three-dimensional visualization, the absence of safe laparoscopic instruments, and the paucity of tactile feedback are still usually the causes of barriers to popularity and the causes of conversion during surgery [48].

Moreover, practicing a new or pioneer surgical technique on patients with a malignant disease is not permitted in the ethical aspect. However, inreasingly more improved new techniques of LR are being explored [47]. For the majority of cases, pioneers feel confident to employ LR. If further efforts are made to achieve standardization of these minimally invasive procedures and improvement of the related educational system, LR will undoubtedly become the standard and mainstay therapy for many bowel diseases, besides colorectal cancer. Furthermore, it is expected that other new techniques such as reduced port surgery and robotic surgery will be confirmed efficient and safe in the future [49].

Many experts pointed out that the learning curve for laparoscopic colorectal surgery is about greater than twenty cases [50]. In 2013, one meta-analysis by Comité de l'évolution des pratiques en oncologie (CEPO) recommended that LR be considered an option for the curative treatment of colon and rectal cancer in consideration of surgeon experience, tumor stage, potential contraindications, and patient expectations. Instead, CEPO also suggested only competent experts with sufficient annual surgeon volume should perform LR for rectal cancer patients [51]. However, safety control, quality monitor, and technique standardization applied to the surgical aspects of the study would provide a solution to the learning curve issues by the collaboration of interested experts to set up safe and reproducible experiment steps even in the setting of new technology [52]. As for the surgeon on our team, he had experience with LR of more than 100 cases before this study. Thus, the learning curve effect of our study series was not discussed.

Compared with previously published randomized studies in the literature, there were some weak points of this study that needed to be further addressed. Our hospital was an 800-bed regional hospital with a total of around 120 colorectal cancer operations per year. Admittedly, the number of patients included in this study was too small (only 351 patients) for comparison of the oncological outcomes; we should increase the sample size to make a reliable comparison between these two groups and to avoid the related bias in the future. Second, our surgeon excluded morbidly obese patients in this study because our surgeon preferred them to receive LR to lower the incidence of possible abdominal wound herniation in the future, which might cause a potential bias. However, only one morbidly obese patient was encountered during the study period, and he was excluded because he encountered the problem of severe intestinal adhesion and received conversion from LR to the OR approach. Third, recently developed and popular trans-anal surgery for patients with rectal lesions was not discussed in this study because we could not compare traditional OR and LR techniques. Fourth, some patients might encounter the problem of their retrieved and assessed lymph node number being lower than the threshold of twelve, which might cause inaccurate TNM staging and select inappropriate minor treatment in their protocol.

In Table 2, the result of our anastomosis leakage rate (0.57%) compared favorably to those of the published literature (0.9%–3.5%). Higher leak rates were typically reported for low pelvic anastomoses or anastomoses to the anal canal [53]. There were three reasons for this comparatively "better" result in our study. First, we largely selected surgical patients with good pre-operative nutrition status (Subjective Global Assessment of Nutritional Status class A or B). Second, we preferred to perform protective diversional stoma for patients with a higher risk of anastomosis leakage, especially those having pre-operative concurrent chemoradiotherapy. We believed nearly all small contained leaks of the anastomosis site would heal after fecal diversion for about six months. Third, the true incidence of anastomosis leakage was underestimated. We only performed post-operative CT for patients when turbid or stool-like discharge was noted from the drainage tube near the anastomosis site. Pickleman et al. advocated that some colorectal surgical patients ultimately found to have an anastomosis leakage developed a more insidious presentation, often with low-grade fever, prolonged ileus, or failure to thrive [54]. In these patients, confirmation of the diagnosis might be much more difficult, as the clinical course was often similar to other postoperative infectious complications. Radiologic imaging was usually required; even then, the definitive diagnosis might be elusive or at least uncertain [53]. Although there have been many studies that specify a rate of anastomotic leakage, it is seldom possible to know what constitutes a "leak". Bruce et al. performed a systematic review of studies measuring

the incidence of anastomotic leaks after gastrointestinal surgery; in the 97 studies reviewed, there was a total of 56 separate definitions of the anastomotic leak [55]. A leak may be defined by the need for reoperation, clinical findings, or radiologic criteria (CT scanning or contrast enema), making "accurate" comparisons between these studies difficult or impossible [53].

Meanwhile, colorectal cancers at different sites were included for analysis of oncologic outcomes and functional results in our study. This study design was debatable because the lymphatic drainage route, range of dissection during tumor resection, operation techniques, and even the biologic behavior were different in various colorectal locations [5]. However, all patients of the LR and OR groups were treated by a single surgeon, and this could avoid the related bias when patients were treated by multiple surgeons. Besides, if we could increase the patient number (sample size) in the future, we could analyze and discuss the tumor located at one specific site to decrease this bias.

Patients with liver metastasis (a common metastasis area of colorectal cancer) were excluded in our study because we thought these terminal patients (defined by current TNM staging system) should receive systemic treatment if no evident clinical lumen obstruction. However, in China, the attitudes of therapeutic approaches in these patients seem to vary among areas [56]. Specific guidelines regarding liver metastasis were revised in 2018 in order to improve the diagnosis and treatment strategy, including the overall clinical evaluation, personalized treatment goals, and comprehensive treatment protocol, in order to prevent the occurrence of liver metastases, and improve the resection rate of liver metastases and survival [57]. Although experts of different countries have investigated their treatment strategy for colorectal cancer thoroughly for many years, the TNM staging system is still commonly regarded as an essential tool to predict oncologic outcomes [58]. It has made an essential contribution to the clinical management of cancer patients over the past 50 years [59], but are we sure it delivers what is needed to provide adequate advice in the 21st century? Would patients face different oncologic outcomes despite that they have been labeled with the same TNM stage clinically?

Nowadays, the degree of cancer infiltration, the number of lymph nodes involved, and distant metastasis have generally been accepted as the most paramount items to predict outcomes [60]. Nevertheless, some patients in the same clinicopathological stage might exhibit unique variation in outcomes with different rates of cancer recurrence and mortality when merely evaluated with the current TNM staging system [60,61]. Specialists conducted intensive studies about the possible causes of this discrepancy. Maguire et al. pointed out that the classification of peritoneal involvement was different in TNM 5 and TNM 7. The Royal College of Pathologists in the United Kingdom still recommended the use of the TNM 5 staging system, while TNM 7 had been adopted in many other jurisdictions. In TNM 5, a tumor directly invading other organs was staged as pT4a, while a tumor involving the visceral peritoneum was staged as pT4b [60]. However, the reported incidence of peritoneal involvement ranged from 5% to 43% in studies of stage II colorectal cancer, which led to a wide statistical variation and an unreliable result [62–69]. Besides, Puppa et al. also advocated that identification and classification of morphologic features encountered in the pathologic examination of colorectal cancer specimens might be difficult and a source of subjective variability. They suggested that enhanced pathologic analysis, agreed-upon standard protocols, and standardization should improve the completeness and accuracy of pathology reports. In other words, the optimal staging system of colorectal cancer should encompass both anatomic and nonanatomic factors, the latter including molecular and treatment factors [61]. Some oncologists also advocated that cancer development and progression might depend partly on "changes" in several histological features, which might lead to this discrepancy clinically. These previously unrecognized features were closely related to the way cancerous cells interact with the surrounding stroma and obtain their potential for invasiveness [70]. These characteristics included tumor budding, poorly differentiated clusters, extramural vascular (vein) invasion, perineural invasion, tumor deposits, and mucin pools [58]. This discrepancy in the molecular signature of colorectal cancer has also revealed differences in phenotypic aggressiveness and therapeutic response rates [58]. Thus, we should remind colorectal cancer patients of the potential risk of having a disappointing result when choosing inappropriate minor post-operative treatment.

The discrepancy in staging colorectal cancer has critical effects on management, outcomes, and survival rates of the patients. Accurate predictions of the final pathological disease stage using high quality, accurate pre-operative clinical-radiological staging techniques enables multidisciplinary teams to plan prompt optimal management strategies for patients with colorectal neoplasms [71]. As cancer clinicians strive to improve survival by increasingly smaller steps, the accuracy of TNM staging becomes even more critical in the interpretation of reports of further clinical trials [59]. More importantly, it is essential to introduce effective preventive measures to this increasing global disease [72].

5. Conclusions

Within the limitations of this study, the results showed better short-term outcomes in terms of postoperative urinary tract infection, wound infection, pneumonia, and blood loss in LR versus OR for non-metastatic colorectal cancer. Although the differences regarding cancer-free five-year survival and tumor recurrence did not reach the statistical significance of both groups, it is strongly recommended that patients undergo laparoscopic surgery if not contraindicated.

Author Contributions: C.-C.C. performed the operations, cared the patients, designed the study, obtained IRB approval, collected clinical records, and wrote the first and final draft; W.-L.L. analyzed the statistical data; H.-Y.S. assisted with the statistical data analysis; C.-C.H. managed patients in the emergency room; J.-J.C. performed pre-operative colonoscopy exams; S.-B.S. performed stool occult blood screening; C.-C.L. and C.-M.C. cared some patients in the intensive care unit post-operatively; C.-J.T. and S.-H.C. performed chemotherapy or concurrent chemoradiotherapy for patients; J.-J.W. provided critical feedback, supervision, and opinion of this study.

Funding: This study was funded by Chi Mei Medical Center, Liouying, Research Grant CLFHR 10808.

Acknowledgments: This study was approved by the institutional review board of Chi Mei Medical Center, 10410-L03.

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

References

1. Guo, D.Y.; Eteuati, J.; Nguyen, M.H.; Lloyd, D.; Ragg, J.L. Laparoscopic assisted colectomy: Experience from a rural centre. *ANZ J. Surg.* **2007**, *77*, 283–286. [CrossRef] [PubMed]
2. Clinical Outcomes of Surgical Therapy Study Group. A comparison of laparoscopically assisted and open colectomy for colon cancer. *N. Engl. J. Med.* **2004**, *350*, 2050–2059. [CrossRef] [PubMed]
3. Leung, K.L.; Kwok, S.P.; Lam, S.C.; Lee, J.F.; Yiu, R.Y.; Ng, S.S.; Lai, P.B.; Lau, W.Y. Laparoscopic resection of rectosigmoid carcinoma: Prospective randomised trial. *Lancet* **2004**, *363*, 1187–1192. [CrossRef]
4. Guillou, P.J.; Quirke, P.; Thorpe, H.; Walker, J.; Jayne, D.G.; Smith, A.M.; Heath, R.M.; Brown, J.M.; MRC CLASICC Trial Group. Short-term endpoints of conventional versus laparoscopic-assisted surgery in patients with colorectal cancer (MRC CLASICC trial): Multicentre, randomised controlled trial. *Lancet* **2005**, *365*, 1718–1726. [CrossRef]
5. Liang, J.T.; Huang, K.C.; Lai, H.S.; Lee, P.H.; Jeng, Y.M. Oncologic results of laparoscopic versus conventional open surgery for stage II or III left-sided colon cancers: A randomized controlled trial. *Ann. Surg. Oncol.* **2007**, *14*, 109–117. [CrossRef] [PubMed]
6. Veldkamp, R.; Kuhry, E.; Hop, W.C.; Jeekel, J.; Kazemier, G.; Bonjer, H.J.; Haglind, E.; Påhlman, L.; Cuesta, M.A.; Msika, S.; et al. Laparoscopic surgery versus open surgery for colon cancer: Short-term outcomes of a randomised trial. *Lancet Oncol.* **2005**, *6*, 477–484. [PubMed]
7. Jayne, D.G.; Guillou, P.J.; Thorpe, H.; Quirke, P.; Copeland, J.; Smith, A.M.; Heath, R.M.; Brown, J.M.; UK MRC CLASICC Trial Group. Randomized trial of laparoscopic-assisted resection of colorectal carcinoma: 3-year results of the UK MRC CLASICC Trial Group. *J. Clin. Oncol.* **2007**, *25*, 3061–3068. [CrossRef]
8. Colon Cancer Laparoscopic or Open Resection Study Group; Buunen, M.; Veldkamp, R.; Hop, W.C.; Kuhry, E.; Jeekel, J.; Haglind, E.; Påhlman, L.; Cuesta, M.A.; Msika, S.; et al. Survival after laparoscopic surgery versus open surgery for colon cancer: Long-term outcome of a randomised clinical trial. *Lancet Oncol.* **2009**, *10*, 44–52.

9. Hemandas, A.K.; Abdelrahman, T.; Flashman, K.G.; Skull, A.J.; Senapati, A.; O'Leary, D.P.; Parvaiz, A. Laparoscopic colorectal surgery produces better outcomes for high risk cancer patients compared to open surgery. *Ann. Surg.* **2010**, *252*, 84–89. [CrossRef]
10. Bagshaw, P.F.; Allardyce, R.A.; Frampton, C.M.; Frizelle, F.A.; Hewett, P.J.; McMurrick, P.J.; Rieger, N.A.; Smith, J.S.; Solomon, M.J.; Stevenson, A.R.; et al. Long-term outcomes of the australasian randomized clinical trial comparing laparoscopic and conventional open surgical treatments for colon cancer: The Australasian Laparoscopic Colon Cancer Study trial. *Ann. Surg.* **2012**, *256*, 915–919. [CrossRef]
11. Cummings, L.C.; Delaney, C.P.; Cooper, G.S. Laparoscopic versus open colectomy for colon cancer in an older population: A cohort study. *World J. Surg. Oncol.* **2012**, *10*, 31. [CrossRef] [PubMed]
12. Hinoi, T.; Kawaguchi, Y.; Hattori, M.; Okajima, M.; Ohdan, H.; Yamamoto, S.; Hasegawa, H.; Horie, H.; Murata, K.; Yamaguchi, S.; et al. Laparoscopic versus open surgery for colorectal cancer in elderly patients: A multicenter matched case-control study. *Ann. Surg. Oncol.* **2015**, *22*, 2040–2050. [CrossRef] [PubMed]
13. Hutter, M.M.; Randall, S.; Khuri, S.F.; Henderson, W.G.; Abbott, W.M.; Warshaw, A.L. Laparoscopic versus open gastric bypass for morbid obesity: A multicenter, prospective, risk-adjusted analysis from the National Surgical Quality Improvement Program. *Ann. Surg.* **2006**, *243*, 657–662. [CrossRef] [PubMed]
14. Weeks, J.C.; Nelson, H.; Gelber, S.; Sargent, D.; Schroeder, G.; Clinical Outcomes of Surgical Therapy (COST) Study Group. Short-term quality-of-life outcomes following laparoscopic-assisted colectomy vs open colectomy for colon cancer: A randomized trial. *JAMA* **2002**, *287*, 321–328. [CrossRef] [PubMed]
15. Tong, D.K.; Law, W.L. Laparoscopic versus open right hemicolectomy for carcinoma of the colon. *JSLS* **2007**, *11*, 76–80. [PubMed]
16. Steele, S.R.; Brown, T.A.; Rush, R.M.; Martin, M.J. Laparoscopic vs open colectomy for colon cancer: Results from a large nationwide population-based analysis. *J. Gastrointest. Surg.* **2008**, *12*, 583–591. [CrossRef] [PubMed]
17. Kennedy, G.D.; Heise, C.; Rajamanickam, V.; Harms, B.; Foley, E.F. Laparoscopy decreases postoperative complication rates after abdominal colectomy: Results from the national surgical quality improvement program. *Ann. Surg.* **2009**, *249*, 596–601. [CrossRef] [PubMed]
18. Liang, Y.; Li, G.; Chen, P.; Yu, J. Laparoscopic versus open colorectal resection for cancer: A meta-analysis of results of randomized controlled trials on recurrence. *Eur. J. Surg. Oncol.* **2008**, *34*, 1217–1224. [CrossRef]
19. Cianchi, F.; Trallori, G.; Mallardi, B.; Macrì, G.; Biagini, M.R.; Lami, G.; Indennitate, G.; Bagnoli, S.; Bonanomi, A.; Messerini, L.; et al. Survival after laparoscopic and open surgery for colon cancer: A comparative, single-institution study. *BMC Surg.* **2015**, *15*, 33. [CrossRef]
20. Neudecker, J.; Klein, F.; Bittner, R.; Carus, T.; Stroux, A.; Schwenk, W.; LAPKON II Trialists. Short-term outcomes from a prospective randomized trial comparing laparoscopic and open surgery for colorectal cancer. *Br. J. Surg.* **2009**, *96*, 1458–1467. [CrossRef]
21. Biondi, A.; Grosso, G.; Mistretta, A.; Marventano, S.; Toscano, C.; Gruttadauria, S.; Basile, F. Laparoscopic-assisted versus open surgery for colorectal cancer: Short- and long-term outcomes comparison. *J. Laparoendosc. Adv. Surg. Tech. A* **2013**, *23*, 1–7. [CrossRef] [PubMed]
22. Małczak, P.; Mizera, M.; Torbicz, G.; Witowski, J.; Major, P.; Pisarska, M.; Wysocki, M.; Strzałka, M.; Budzyński, A.; Pędziwiatr, M. Is the laparoscopic approach for rectal cancer superior to open surgery? A systematic review and meta-analysis on short-term surgical outcomes. *Wideochir. Inne Tech. Maloinwazyjne* **2018**, *13*, 129–140. [CrossRef] [PubMed]
23. Tjandra, J.J.; Chan, M.K. Systematic review on the short-term outcome of laparoscopic resection for colon and rectosigmoid cancer. *Colorectal. Dis.* **2006**, *8*, 375–388. [CrossRef] [PubMed]
24. Senagore, A.J.; Stulberg, J.J.; Byrnes, J.; Delaney, C.P. A national comparison of laparoscopic vs. open colectomy using the National Surgical Quality Improvement Project data. *Dis. Colon Rectum* **2009**, *52*, 183–186. [CrossRef]
25. Law, W.L.; Poon, J.T.; Fan, J.K.; Lo, S.H. Comparison of outcome of open and laparoscopic resection for stage II and stage III rectal cancer. *Ann. Surg. Oncol.* **2009**, *16*, 1488–1493. [CrossRef] [PubMed]
26. Lacy, A.M.; Delgado, S.; Castells, A.; Prins, H.A.; Arroyo, V.; Ibarzabal, A.; Pique, J.M. The long-term results of a randomized clinical trial of laparoscopy-assisted versus open surgery for colon cancer. *Ann. Surg.* **2008**, *248*, 1–7. [CrossRef] [PubMed]

27. Capussotti, L.; Massucco, P.; Muratore, A.; Amisano, M.; Bima, C.; Zorzi, D. Laparoscopy as a prognostic factor in curative resection for node positive colorectal cancer: Results for a single-institution nonrandomized prospective trial. *Surg. Endosc.* **2004**, *18*, 1130–1135. [CrossRef]
28. Law, W.L.; Poon, J.T.; Fan, J.K.; Lo, O.S. Survival following laparoscopic versus open resection for colorectal cancer. *Int. J. Colorectal. Dis.* **2012**, *27*, 1077–1085. [CrossRef]
29. Li, J.; Guo, H.; Guan, X.D.; Cai, C.N.; Yang, L.K.; Li, Y.C.; Zhu, Y.H.; Li, P.P.; Liu, X.L.; Yang, D.J. The impact of laparoscopic converted to open colectomy on short-term and oncologic outcomes for colon cancer. *J. Gastrointest. Surg.* **2015**, *19*, 335–343. [CrossRef]
30. Patankar, S.K.; Larach, S.W.; Ferrara, A.; Williamson, P.R.; Gallagher, J.T.; DeJesus, S.; Narayanan, S. Prospective comparison of laparoscopic vs. open resections for colorectal adenocarcinoma over a ten-year period. *Dis. Colon Rectum* **2003**, *46*, 601–611. [CrossRef]
31. Lujan, H.J.; Plasencia, G.; Jacobs, M.; Viamonte, M., III; Hartmann, R.F. Long-term survival after laparoscopic colon resection for cancer: Complete five-year follow-up. *Dis. Colon Rectum* **2002**, *45*, 491–501. [CrossRef] [PubMed]
32. Edge, S.B.; Compton, C.C. The American Joint Committee on Cancer: The 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann. Surg. Oncol.* **2010**, *17*, 1471–1474. [CrossRef] [PubMed]
33. Qiu, H.B.; Zhang, L.Y.; Li, Y.F.; Zhou, Z.W.; Keshari, R.P.; Xu, R.H. Ratio of metastatic to resected lymph nodes enhances to predict survival in patients with stage III colorectal cancer. *Ann. Surg. Oncol.* **2011**, *18*, 1568–1574. [CrossRef] [PubMed]
34. Huh, J.W.; Kim, Y.J.; Kim, H.R. Ratio of metastatic to resected lymph nodes as a prognostic factor in node-positive colorectal cancer. *Ann. Surg. Oncol.* **2010**, *17*, 2640–2646. [CrossRef]
35. Wang, J.; Hassett, J.M.; Dayton, M.T.; Kulaylat, M.N. Lymph node ratio: Role in the staging of node-positive colon cancer. *Ann. Surg. Oncol.* **2008**, *15*, 1600–1608. [CrossRef] [PubMed]
36. Greenberg, R.; Itah, R.; Ghinea, R.; Sacham-Shmueli, E.; Inbar, R.; Avital, S. Metastatic lymph node ratio (LNR) as a prognostic variable in colorectal cancer patients undergoing laparoscopic resection. *Tech. Coloproctol.* **2011**, *15*, 273–279. [CrossRef] [PubMed]
37. Li Destri, G.; Di Carlo, I.; Scilletta, R.; Scilletta, B.; Puleo, S. Colorectal cancer and lymph nodes: The obsession with the number 12. *World J. Gastroenterol.* **2014**, *20*, 1951–1960.
38. Novitsky, Y.W.; Litwin, D.E.; Callery, M.P. The net immunologic advantage of laparoscopic surgery. *Surg. Endosc.* **2004**, *18*, 1411–1419. [CrossRef]
39. Lacy, A.M.; García-Valdecasas, J.C.; Delgado, S.; Castells, A.; Taurá, P.; Piqué, J.M.; Visa, J. Laparoscopy-assisted colectomy versus open colectomy for treatment of non-metastatic colon cancer: A randomised trial. *Lancet* **2002**, *359*, 2224–2229. [CrossRef]
40. Laurent, C.; Leblanc, F.; Wütrich, P.; Scheffler, M.; Rullier, E. Laparoscopic versus open surgery for rectal cancer: Long-term oncologic results. *Ann. Surg.* **2009**, *250*, 54–61. [CrossRef]
41. Whelan, R.L.; Franklin, M.; Holubar, S.D.; Donahue, J.; Fowler, R.; Munger, C.; Doorman, J.; Balli, J.E.; Glass, J.; Gonzalez, J.J.; et al. Postoperative cell mediated immune response is better preserved after laparoscopic vs open colorectal resection in humans. *Surg. Endosc.* **2003**, *17*, 972–978. [CrossRef] [PubMed]
42. Wichmann, M.W.; Hüttl, T.P.; Winter, H.; Spelsberg, F.; Angele, M.K.; Heiss, M.M.; Jauch, K.W. Immunological effects of laparoscopic vs open colorectal surgery: A prospective clinical study. *Arch. Surg.* **2005**, *140*, 692–697. [CrossRef] [PubMed]
43. Bessa, X.; Castells, A.; Lacy, A.M.; Elizalde, J.I.; Delgado, S.; Boix, L.; Piñol, V.; Pellisé, M.; García-Valdecasas, J.C.; Piqué, J.M. Laparoscopic-assisted vs. open colectomy for colorectal cancer: Influence on neoplastic cell mobilization. *J. Gastrointest. Surg.* **2001**, *5*, 66–73. [CrossRef]
44. Hsu, T.C. Feasibility of colectomy with mini-incision. *Am. J. Surg.* **2005**, *190*, 48–50. [CrossRef] [PubMed]
45. Hoyuela, C.; Juvany, M.; Carvajal, F. Single-incision laparoscopy versus standard laparoscopy for colorectal surgery: A systematic review and meta-analysis. *Am. J. Surg.* **2017**, *214*, 127–140. [CrossRef] [PubMed]
46. Acuna, S.A.; Elmi, M.; Shah, P.S.; Coburn, N.G.; Quereshy, F.A. Preoperative localization of colorectal cancer: A systematic review and meta-analysis. *Surg. Endosc.* **2017**, *31*, 2366–2379. [CrossRef] [PubMed]
47. Nishigori, N.; Koyama, F.; Nakagawa, T.; Nakamura, S.; Ueda, T.; Inoue, T.; Kawasaki, K.; Obara, S.; Nakamoto, T.; Fujii, H.; et al. Visualization of Lymph/Blood Flow in Laparoscopic Colorectal Cancer Surgery by ICG Fluorescence Imaging (Lap-IGFI). *Ann. Surg. Oncol.* **2016**, *23*, S266–S274. [CrossRef] [PubMed]

48. Chung, C.C.; Ng, D.C.; Tsang, W.W.; Tang, W.L.; Yau, K.K.; Cheung, H.Y.; Wong, J.C.; Li, M.K. Hand-assisted laparoscopic versus open right colectomy: A randomized controlled trial. *Ann. Surg.* **2007**, *246*, 728–733. [CrossRef] [PubMed]
49. Sato, T.; Watanabe, M. Present laparoscopic surgery for colorectal cancer in Japan. *World J. Clin. Oncol.* **2016**, *7*, 155–159. [CrossRef] [PubMed]
50. Tekkis, P.P.; Senagore, A.J.; Delaney, C.P.; Fazio, V.W. Evaluation of the learning curve in laparoscopic colorectal surgery: Comparison of right-sided and left-sided resections. *Ann. Surg.* **2005**, *242*, 83–91. [CrossRef] [PubMed]
51. Morneau, M.; Boulanger, J.; Charlebois, P.; Latulippe, J.F.; Lougnarath, R.; Thibault, C.; Gervais, N.; Comité de l'Évolution des Pratiques en Oncologie. Laparoscopic versus open surgery for the treatment of colorectal cancer: A literature review and recommendations from the Comité de l'évolution des pratiques en oncologie. *Can. J. Surg.* **2013**, *56*, 297–310. [CrossRef] [PubMed]
52. Fleshman, J.; Sargent, D.J.; Green, E.; Anvari, M.; Stryker, S.J.; Beart, R.W., Jr.; Hellinger, M.; Flanagan, R., Jr.; Peters, W.; Nelson, H.; et al. Laparoscopic colectomy for cancer is not inferior to open surgery based on 5-year data from the COST Study Group trial. *Ann. Surg.* **2007**, *246*, 655–662. [CrossRef] [PubMed]
53. Hyman, N.; Manchester, T.L.; Osler, T.; Burns, B.; Cataldo, P.A. Anastomotic leaks after intestinal anastomosis: it's later than you think. *Ann. Surg.* **2007**, *245*, 254–258. [CrossRef] [PubMed]
54. Pickleman, J.; Watson, W.; Cunningham, J.; Fisher, S.G.; Gamelli, R. The failed gastrointestinal anastomosis: An inevitable catastrophe? *J. Am. Coll. Surg.* **1999**, *188*, 473–482. [CrossRef]
55. Bruce, J.; Krukowski, Z.H.; Al-Khairy, G.; Russell, E.M.; Park, K.G. Systematic review of the definition and measurement of anastomotic leak after gastrointestinal surgery. *Br. J. Surg.* **2001**, *88*, 1157–1168. [CrossRef] [PubMed]
56. Zhang, Y.; Chen, Z.; Li, J. The current status of treatment for colorectal cancer in China: A systematic review. *Medicine* **2017**, *96*, e8242. [CrossRef] [PubMed]
57. Xu, J.; Ren, L. China Guideline for Diagnosis and Comprehensive Treatment of Colorectal Liver Metastases (Version 2018). *Zhonghua Wei Chang Wai Ke Za Zhi* **2018**, *21*, 601–626. [PubMed]
58. Athanasakis, E.; Xenaki, S.; Venianaki, M.; Chalkiadakis, G.; Chrysos, E. Newly recognized extratumoral features of colorectal cancer challenge the current tumor-node-metastasis staging system. *Ann. Gastroenterol.* **2018**, *31*, 525–534. [CrossRef]
59. Quirke, P.; Williams, G.T.; Ectors, N.; Ensari, A.; Piard, F.; Nagtegaal, I. The future of the TNM staging system in colorectal cancer: Time for a debate? *Lancet Oncol.* **2007**, *8*, 651–657. [CrossRef]
60. Maguire, A.; Sheahan, K. Controversies in the pathological assessment of colorectal cancer. *World J. Gastroenterol.* **2014**, *20*, 9850–9861. [CrossRef]
61. Puppa, G.; Sonzogni, A.; Colombari, R.; Pelosi, G. TNM staging system of colorectal carcinoma: A critical appraisal of challenging issues. *Arch. Pathol. Lab. Med.* **2010**, *134*, 837–852. [PubMed]
62. Lennon, A.M.; Mulcahy, H.E.; Hyland, J.M.; Lowry, C.; White, A.; Fennelly, D.; Murphy, J.J.; O'Donoghue, D.P.; Sheahan, K. Peritoneal involvement in stage II colon cancer. *Am. J. Clin. Pathol.* **2003**, *119*, 108–113. [CrossRef] [PubMed]
63. Kojima, M.; Nakajima, K.; Ishii, G.; Saito, N.; Ochiai, A. Peritoneal elastic laminal invasion of colorectal cancer: The diagnostic utility and clinicopathologic relationship. *Am. J. Surg. Pathol.* **2010**, *34*, 1351–1360. [CrossRef] [PubMed]
64. Burdy, G.; Panis, Y.; Alves, A.; Nemeth, J.; Lavergne-Slove, A.; Valleur, P. Identifying patients with T3-T4 node-negative colon cancer at high risk of recurrence. *Dis. Colon Rectum* **2001**, *44*, 1682–1688. [CrossRef] [PubMed]
65. Merkel, S.; Wein, A.; Günther, K.; Papadopoulos, T.; Hohenberger, W.; Hermanek, P. High-risk groups of patients with stage II colon carcinoma. *Cancer* **2001**, *92*, 1435–1443. [CrossRef]
66. Morris, M.; Platell, C.; de Boer, B.; McCaul, K.; Iacopetta, B. Population-based study of prognostic factors in stage II colonic cancer. *Br. J. Surg.* **2006**, *93*, 866–871. [CrossRef] [PubMed]
67. Newland, R.C.; Dent, O.F.; Chapuis, P.H.; Bokey, L. Survival after curative resection of lymph node negative colorectal carcinoma. A prospective study of 910 patients. *Cancer* **1995**, *76*, 564–571. [CrossRef]
68. Petersen, V.C.; Baxter, K.J.; Love, S.B.; Shepherd, N.A. Identification of objective pathological prognostic determinants and models of prognosis in Dukes' B colon cancer. *Gut* **2002**, *51*, 65–69. [CrossRef]

69. Grin, A.; Messenger, D.E.; Cook, M.; O'Connor, B.I.; Hafezi, S.; El-Zimaity, H.; Kirsch, R. Peritoneal elastic lamina invasion: Limitations in its use as a prognostic marker in stage II colorectal cancer. *Hum. Pathol.* **2013**, *44*, 2696–2705. [CrossRef]
70. Grizzi, F.; Celesti, G.; Basso, G.; Laghi, L. Tumor budding as a potential histopathological biomarker in colorectal cancer: Hype or hope? *World J. Gastroenterol.* **2012**, *18*, 6532–6536. [CrossRef]
71. Elfaedy, O.; Owens, P.; Aakif, M.; Mansour, E. Discrepancy in Colorectal Cancer Staging: A Single Center Experience. Available online: http://www.surgeryresearchjournal.com/full-text/wjssr-v1-id1054.php (accessed on 1 March 2019).
72. Gong, Y.; Peng, P.; Bao, P.; Zhong, W.; Shi, Y.; Gu, K.; Zheng, Y.; Wu, C.; Cai, S.; Xu, Y.; et al. The Implementation and First-Round Results of a Community-Based Colorectal Cancer Screening Program in Shanghai, China. *Oncologist* **2018**, *23*, 928–935. [CrossRef] [PubMed]

Journal of
Clinical Medicine

MDPI

Article

The Efficacy and Safety of Eravacycline in the Treatment of Complicated Intra-Abdominal Infections: A Systemic Review and Meta-Analysis of Randomized Controlled Trials

Shao-Huan Lan [1], Shen-Peng Chang [2], Chih-Cheng Lai [3], Li-Chin Lu [4] and Chien-Ming Chao [3,*]

1 School of Pharmaceutical Sciences and Medical Technology, Putian University, Putian 351100, Fujian, China; shawnlan0713@gmail.com
2 Department of Pharmacy, Chi Mei Medical Center, Liouying 73657, Taiwan; httremoon@ms.szmc.edu.tw
3 Department of Intensive Care Medicine, Chi Mei Medical Center, Liouying 73657, Taiwan; dtmed141@gmail.com
4 School of Management, Putian University, Putian 351100, China; jane90467@gmail.com
* Correspondence: ccm870958@yahoo.com.tw

Received: 27 May 2019; Accepted: 13 June 2019; Published: 17 June 2019

Abstract: This study aims to assess the clinical efficacy and safety of eravacycline for treating complicated intra-abdominal infection (cIAI) in adult patients. The PubMed, Web of Science, EBSCO, Cochrane databases, Ovid Medline, Embase, and ClinicalTrials.gov were searched up to May 2019. Only randomized controlled trials (RCTs) that evaluated eravacycline and other comparators for the treatment of cIAI were included. The primary outcome was the clinical cure rate at the test-of-cure visit based on modified intent-to-treat population, microbiological intent-to-treat population, clinically evaluable population, and microbiological evaluable population, and the secondary outcomes were clinical failure rate and the risk of adverse event. Three RCTs were included. Overall, eravacycline had a clinical cure rate (88.7%, 559/630) at test-of-cure in modified intent-to-treat population similar to comparators (90.1%, 492/546) in the treatment of cIAIs (risk ratio (RR), 0.99; 95% confidence interval (CI), 0.95–1.03; I^2 = 0%, Figure 3). In the microbiological intent-to-treat, clinically evaluable, and microbiological evaluable populations, no difference was found between eravacycline and comparators in terms of clinical cure rate at test-of-cure (microbiological intent-to-treat population, RR, 0.99; 95% CI, 0.95–1.04; I^2 = 0%, clinically evaluable population, RR, 1.00; 95% CI, 0.97–1.03; I^2 = 0%, microbiological evaluable population, RR, 0.98; 95% CI, 0.95–1.02; I^2 = 0%). In addition, eravacycline had clinical failure rate similar to comparators at test-of-cure in modified intent-to-treat population (RR, 1.01; 95% CI, 0.61–0.69; I^2 = 0%), microbiological intent-to-treat population (RR, 1.34; 95% CI, 0.77–2.31; I^2 = 16%), clinically evaluable population (RR, 1.03; 95% CI, 0.61–1.76; I^2 = 0%), and microbiological evaluable population (RR, 1.32; 95% CI, 0.75–2.32; I^2 = 10%). Although eravacycline was associated with higher risk of treatment-emergent adverse event than comparators (RR, 1.34; 95% CI, 1.13–1.58; I^2 = 0%), no significant differences were found between eravacycline and comparators for the risk of serious adverse event (RR, 1.04; 95% CI, 0.65–1.65; I^2 = 0%), discontinuation of study drug because of adverse event (RR, 0.68; 95% CI, 0.23–1.99; I^2 = 13%), and all-cause mortality (RR, 1.09; 95% CI, 0.41–2.9; I^2 = 28%). In conclusion, the clinical efficacy of eravacycline is as high as that of the comparator drugs in the treatment of cIAIs and this antibiotic is as well tolerated as the comparators.

Keywords: eravacycline; complicated intra-abdominal infection; efficacy; safety; mortality

1. Introduction

In contrast to uncomplicated abdominal infections, complicated intra-abdominal infections (cIAIs) can extend beyond the originally infected organ into peritoneal spaces, and can be associated with local or diffuse peritonitis [1,2]. *Enterobacteriaceae*, especially *Escherichia coli* and *Klebsiella pneumoniae*, are the most common pathogens causing cIAIs [3–5]. Emergence of multiple antibiotic resistances has become the major concern in this clinical entity and further limits the choice of optimal antibiotic treatment. *E. coli*, *Proteus* species, and *K. pneumoniae* are the most common pathogens; however, high resistance to broad-spectrum antibiotics, including extended-spectrum β-lactams and fluoroquinolones, among these pathogens, also emerges as a critical threat worldwide.

Eravacycline is a novel, synthetic fluorocycline antibacterial agent [6], and has excellent bactericidal activity against most antibiotic-resistant pathogens according to several in vitro studies [7–10]. Recently, the clinical efficacy of eravacycline in cIAI has been evaluated in several clinical studies [11–13]. However, an updated meta-analysis comparing the efficacy and safety of eravacycline and other comparators for the treatment of cIAI is lacking. Therefore, we conducted this meta-analysis to provide real-time evidence about the efficacy and safety of cIAI.

2. Methods

2.1. Study Search and Selection

All clinical studies were identified through a systematic review of the literature in PubMed, Web of Science, EBSCO, Cochrane databases, Ovid Medline, Embase, and ClinicalTrials.gov until May 2019 using the following search terms: "eravacycline", "Xerava™", "TP-434", and "abdom*" (Search strategy presented in Appendix A). Studies were considered eligible for inclusion if they directly compared the clinical efficacy and safety of eravacycline with other antimicrobial agents in the treatment of adult patients with cIAIs. Studies were excluded if they focused on in vitro activity, animal studies, or pharmacokinetic–pharmacodynamic assessment. Two authors (S.-P.C. and S.-H.L.) searched and examined publications independently. When they disagreed, the third author (C.-C.L.) resolved the issue. The following data including year of publication, study design, type of infections, patients' demographic features, antimicrobial regimens, clinical and microbiological outcomes, and adverse effects were extracted from every included study.

2.2. Outcome Measurement

The primary outcome of this meta-analysis was clinical response assessed at the test-of-cure visit, end-of-treatment, and follow-up visit based on modified intent-to-treat population, microbiological intent-to-treat population, clinically evaluable population, microbiological evaluable populations. The intent-to-treat population included all randomized patients, and the modified intent-to-treat population included all intent-to treat patients who received any amount of study drug. The microbiological intent-to-treat population included all modified intent-to-treat patients who met the minimal disease definition of cIAI and had a baseline pathogen identified. The clinically evaluable population included all modified intent-to-treat patients who met the minimal disease definition of cIAI and had a clinical response assessed at the test-of-cure visit. The microbiological evaluable population included all clinically evaluable patients who had a baseline pathogen identified and a microbiological response assessed. Clinical response was classified as cure, failure, or indeterminate based on clinical outcomes. Clinical cure was defined as resolution of all or most pretherapy signs or symptoms with no further requirement for antibiotics, radiological intervention, or surgery. The safety population included all patients who received any intravenous study therapy. Treatment-emergent adverse events were defined as adverse events that started during or after the first dose of study drug administration or increased in severity or relationship to the study drugs during the study. Serious adverse event is defined as an untoward medical occurrence or effect that at any dose results in

death, is life-threatening, requires hospitalization or extension of existing hospitalization, or results in persistent or significant disability.

2.3. Data Analysis

The quality of enrolled RCTs and the risk of bias were assessed using Cochrane Risk of Bias Assessment tool [14]. Statistical analyses were conducted using the software Review Manager, version 5.3. The degree of heterogeneity was evaluated with the Q statistic generated from the χ^2 test. The proportion of statistical heterogeneity was assessed using the I^2 measure. Heterogeneity was considered significant when the p was less than 0.1 or I^2 was greater than 50%. The random-effects model was used when data were significantly heterogeneous and the fixed-effect model was used when data were homogeneous. Pooled risk ratios (RRs) and 95% confidence intervals (CIs) were calculated for outcome analyses.

3. Results

3.1. Study Selection and Characteristics

The search program yielded 147 references. After excluding 90 duplications, the remaining 57 abstracts were screened. Among them, we retrieved 11 articles for full-text review. Finally, three studies [11–13] fulfilling the inclusion criteria were included in this meta-analysis (Figure 1). All enrolled studies had the same principal investigator. All studies [11–13] were randomized, multicenter, and multinational studies designed to compare the clinical efficacy and safety of eravacycline with other comparators for adult patients with cIAI (Table 1). The inclusion criterion of these three studies was that adult patients had to have clinical evidence of cIAI requiring urgent surgical or percutaneous intervention within 48 hours of diagnosis. Two studies [11,13] compared eravacycline with ertapenem, and one [12] compared with meropenem. The test-of-cure evaluation was conducted 25 to 31 calendar days in two studies [11,12] and 10 to 14 days in one study [13] after the first dose of the study drug was administered for the patients with cIAI. The follow-up visit was performed 38 to 50 calendar days in one study [11] and 28 to 42 days in one study [13] after the first dose of study drug was administered. All of the domains in each study were classified as having a low risk of bias (Table 2).

Figure 1. Study selection process flow.

Table 1. Characteristics of included studies.

Study, Published Year	Study Design	Study Site	Study Period	No. of Patients (ITT population)		Dose Regimen	
				Eravacycline	Comparator	Eravacycline	Comparator
Solomkin et al, 2014	Randomized, double-blind trial	19 sites in 6 countries	2011–2012	56 (1.5 mg/kg), 57 (1.0 mg/kg)	30	1.5 mg/kg or 1.0 mg/kg q24 h	Ertapenem 1 g q24 h
IGNITE1, 2017	Randomized, double-blind trial	66 sites in 11 countries	2013–2014	270	271	1.0 mg/kg q12 h	Ertapenem 1 g q24 h
IGNITE4, 2018	Randomized, double-blind trial	65 sites in 11 countries	2016–2017	250	250	1.0 mg/kg q12 h	Meropenem 1 g q8 h

ITT, intention to treat; q, every; h, hour; mg, milligram; g, gram.

Table 2. Risk of bias per study and domain.

Risk of Bias	Study		
	IGNITE1, 2017	IGNITE4, 2018	Solomkin et al, 2014
Random sequence generation	low	low	low
Allocation concealment	low	low	low
Blinding of participants and personnel	low	low	low
Blinding of outcome assessment	low	low	low
Incomplete outcome data	low	low	low
Selective reporting	low	low	low

3.2. Clinical Efficacy and Microbiologic Response

Overall, eravacycline had a clinical cure rate (88.7%, 559/630) at test-of-cure in modified intent-to-treat population similar to comparators (90.1%, 492/546) in the treatment of cIAIs (risk ratio (RR), 0.99; 95% CI, 0.95–1.03; $I^2 = 0\%$, Figure 2). In the microbiological intent-to-treat, clinically evaluable, and microbiological evaluable populations, no difference was found between eravacycline and comparators in terms of clinical cure rate at test-of-cure (Figure 2). In addition, no significant difference was observed between eravacycline and comparator in terms of clinical failure rate at test-of-cure in modified intent-to-treat population, microbiological intent-to-treat population, clinically evaluable population, and microbiological evaluable population (Figure 3).

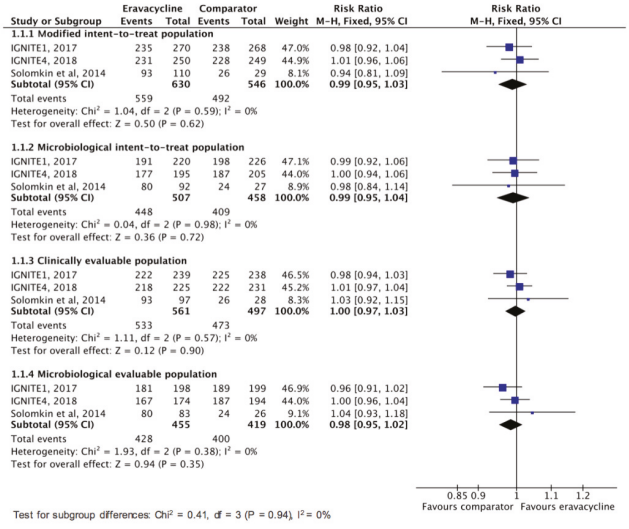


Figure 2. Overall clinical cure rates for eravacycline and comparators in the treatment of complicated intra-abdominal infections.

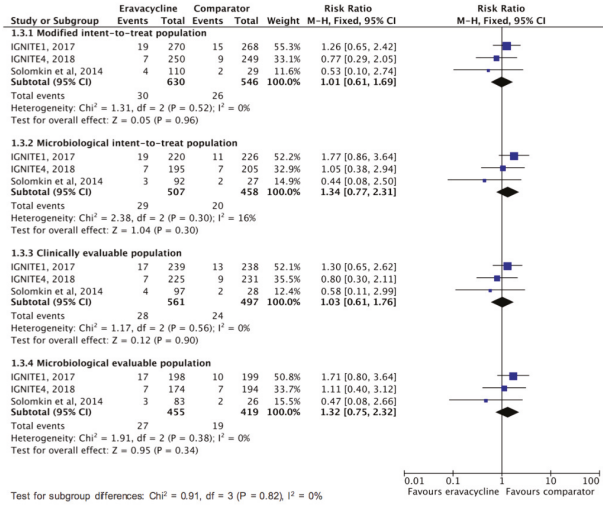


Figure 3. Overall clinical failure rates for eravacycline and comparators in the treatment of complicated intra-abdominal infections.

Only two studies [12,13] reported the outcome at end-of-treatment, and the pooled analysis showed no significant difference was observed between eravacycline and comparator in terms of clinical cure rate at end-of-treatment in modified intent-to-treat population (RR, 0.99; 95% CI, 0.95–1.03; I^2 = 1%), microbiological intent-to-treat population (RR, 0.98; 95% CI, 0.94–1.03; I^2 = 0%), clinically evaluable population (RR, 0.99; 95% CI, 0.96–1.01; I^2 = 0%), and microbiological evaluable population (RR, 0.98; 95% CI, 0.95–1.01; I^2 = 0%). In addition, these two studies[12,13] reported the outcome at follow-up, and the pooled analysis showed no significant difference was observed between eravacycline and comparator in terms of clinical cure rate at follow-up in modified intent-to-treat population (RR, 0.99; 95% CI, 0.93–1.04; I^2 = 0%), microbiological intent-to-treat population (RR, 0.98; 95% CI, 0.92–1.06; I^2 = 0%), clinically evaluable population (RR, 1.01; 95% CI, 0.97–1.06; I^2 = 0%), and microbiological evaluable population (RR, 1.02; 95% CI, 0.97–1.07; I^2 = 30%).

3.3. Adverse Events

In the pooled analysis of three studies reporting adverse events, we found that eravacycline was associated with a higher risk of treatment-emergent adverse events than comparators (Figure 4). However, no significant differences were found between eravacycline and comparators for the risk of serious adverse events, discontinuation of study drug because of adverse event, and all-cause mortality (Figure 4). The most common adverse event among the eravacycline group was nausea (6.5%, 41/629) and vomiting (3.8%, 24/629). Although the risks of nausea and vomiting in the eravacycline group were higher than those in the comparator group, these differences did not reach statistical significance (for nausea, RR, 4.79; 95% CI, 0.84–27.14;7 I^2 = 70%, for vomiting, RR, 1.46; 95% CI, 0.76–2.81;7 I^2 = 0%).

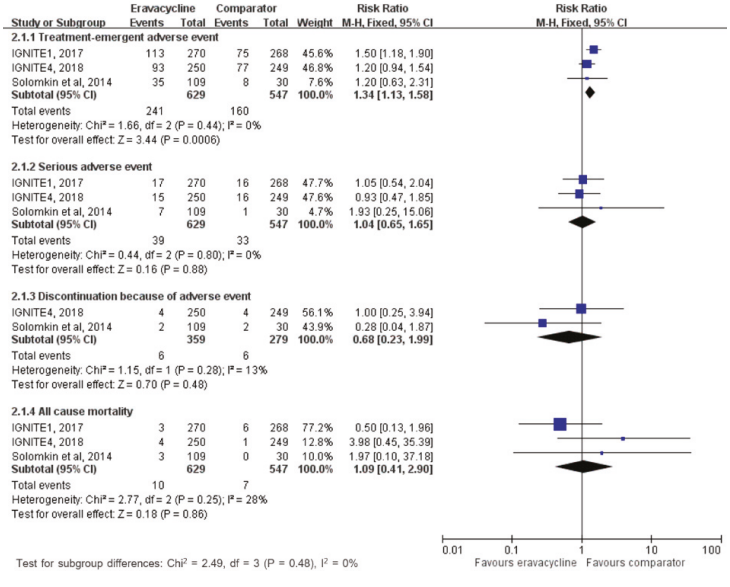


Figure 4. Adverse event risks with eravacycline and comparators in the treatment of complicated intra-abdominal infections.

4. Discussion

This first meta-analysis based on three RCTs [11–13] determined that the clinical efficacy of eravacycline is similar to that of other comparators in the treatment of adult patients with cIAIs. This significant finding is supported by the following analysis. First, the overall pooled clinical cure rate at test-of-cure of eravacycline in treating cIAIs was comparable to carbapenems in modified

intent-to-treat, microbiological intent-to-treat, clinically evaluable, and microbiological evaluable populations. Second, pooled clinical failure rate at test-of-cure of eravacycline was as low as comparators in modified intent-to-treat, microbiological intent-to-treat, clinically evaluable, and microbiological evaluable populations. Third, this similarity in terms of clinical efficacy between eravacycline and comparators did not change with the timing of the outcome measure at end-of-treatment and follow-up. In summary, all of these findings indicated that eravacycline can be an effective therapeutic option in the treatment of adult patients with cIAIs.

The effectiveness of ceftaroline in the treatment of cIAIs in adult patients can be supported by in vitro studies [7,9,10,15,16]. In the surveillance of 2213 Gram-negative and 2423 Gram-positive pathogens in 13 Canadian hospitals, the minimum inhibitory concentration$_{90}$ (MIC$_{90}$) ranged from 0.5 to 2μg/mL for 9 species of *Enterobacteriaceae* tested (n = 2067) and extended-spectrum β-lactamase producing *E. coli* (n = 141) and *K. pneumoniae* (n = 21) did not affect the potency of eravacycline in this study [10]. In another survey of more than 4000 Gram-negative pathogens in New York hospitals [7], eravacycline demonstrated great in vitro activity against *Enterobacteriaceae—E. coli*, *K. pneumoniae*, *Enterobacter aerogenes*, and *Enterobacter cloacae* with minimum inhibitory concentration$_{50}$ (MIC$_{50}$)/MIC$_{90}$ of 0.12/0.5 μg/mL, 0.25/1 μg/mL, 0.25/1 μg/mL, and 0.5/1 μg/mL, respectively. Moreover, the potent activity was retained against multidrug-resistant (MDR) isolates, including carbapenem nonsusceptible strains [7,9]. In addition to aerobic bacteria, anerobic bacteria play important roles in the cIAIs. Eravacycline showed good in vitro activity against *Bacteroides* spp., *Parabacteroides* spp., and *Clostridioides difficile* (formerly *Clostridium difficile*) and eravacycline remained potent against the strains with tetracycline-specific resistance determinants and MDR anaerobic pathogens [15,16]. Overall, the potent in vitro activity of eravacycline against commonly encountered pathogens causing cIAI largely explains the great in vivo clinical response in this meta-analysis.

In addition to clinical efficacy of eravacycline for the treatment of cIAIs, we should consider the risk of adverse event while prescribing eravacycline. Nausea and vomiting were the most common adverse events, and the overall incidence of these adverse events were higher than those of comparators. Moreover, the pooled risk of treatment-emergent adverse events was higher in the eravacycline group than in the control group. These findings are consistent with previous pooled analysis of IGNITE1 and IGNITE4, in which eravacycline recipients had higher incidence of nausea (6.5 vs. 0.6%) and vomiting (3.7 vs. 2.5%) [17]. In contrast, the incidence of serious adverse events, discontinuation of study drug because of adverse event, and all-cause mortality was similar between eravacycline and comparators. Therefore, the findings of this meta-analysis suggest that although eravacycline is associated with higher risk of mild adverse events than comparator, overall, eravacycline remains as safe as other comparators in the treatment of cIAI among adult patients.

This study has several limitations. First, only three RCTs were considered in this meta-analysis. Second, the usefulness of eravacycline in treating cIAIs was not assessed according to the disease severity. Third, we did not evaluate the correlation between in vitro activity and in vivo response of eravacycline against each specific pathogen, particularly antibiotic-resistant organisms, in this study.

5. Conclusions

In conclusion, eravacycline is as good as comparators in terms of efficacy and tolerance in the treatment of cIAI in adult patients.

Author Contributions: Conceptualization, S.-H.L., C.-C.L. and C.-M.C.; Data curation, S.-P.C., C.-C.L. and L.-C.L.; Formal analysis, S.-H.L., S.-P.C. and L.-C.L.; Writing—original draft, C.-C.L.; Writing—review & editing, C.-M.C.

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

Appendix A. Search Strategy

PubMed Search Strategy—Last Searched on 26 May 2019		Results
1	Eravacycline [Title/Abstract] OR TP-434 [Title/Abstract] OR Xerava [Title/Abstract]	74
2	abdom* [Title/Abstract]	330,674
3	1 AND 2	
4	Search (abdom* (Title/Abstract)) AND (((Eravacycline (Title/Abstract)) OR TP-434 (Title/Abstract)) OR Xerava (Title/Abstract))	22
Web of Science Search Strategy—Last Searched on 26 May 2019		**Results**
1	(Eravacycline) OR (Xerava) OR (TP-434)	71
2	(abdom*)	269,250
3	1 AND 2	
4	#1 AND #2	20
EBSCO Search Strategy—Last Searched on 26 May 2019		**Results**
1	AB Eravacycline OR AB Xerava OR AB TP-434	176
2	AB abdom*	495,125
3	1 AND 2	
4	S1 AND S2	40
Cochrane Library Search Strategy—Last Searched on 26 May 2019		**Results**
1	(Eravacycline):ti,ab,kw OR (TP-434):ti,ab,kw OR (Xerava):ti,ab,kw	12
2	(abdom*):ti,ab,kw	40,365
3	1 AND 2	
4	#1 AND #2	5
Ovid Medline Search Strategy—Last Searched on 26 May 2019		**Results**
1	(Eravacycline or Xerava or TP-434).ab.	82
2	abdom*.ab	373,974
3	1 AND 2	
4	1 and 2	24
Embase Search Strategy—Last Searched on 26 May 2019		**Results**
1	eravacycline:ti,ab,kw OR xerava:ti,ab,kw OR 'tp 434':ti,ab,kw	87
2	abdom*:ti,ab,kw	508,670
3	1 AND 2	
4	#1 AND #2	27
ClinicalTrials.gov Search Strategy—Last Searched on May 26, 2019		**Results**
1	Eravacycline (completed studies)	9

References

1. Solomkin, J.S.; Mazuski, J.E.; Baron, E.J.; Sawyer, R.G.; Nathens, A.B.; DiPiro, J.T.; Buchman, T.; Dellinger, E.P.; Jernigan, J.; Gorbach, S.; et al. Guidelines for the selection of anti-infective agents for complicated intra-abdominal infections. *Clin. Infect. Dis.* **2003**, *37*, 997–1005. [CrossRef] [PubMed]
2. Solomkin, J.S.; Mazuski, J.E.; Bradley, J.S.; Rodvold, K.A.; Goldstein, E.J.; Baron, E.J.; O'Neill, P.J.; Chow, A.W.; Dellinger, E.P.; Eachempati, S.R.; et al. Diagnosis and management of complicated intra-abdominal infection in adults and children: guidelines by the Surgical Infection Society and the Infectious Diseases Society of America. *Clin. Infect. Dis.* **2010**, *50*, 133–164. [CrossRef] [PubMed]

3. Brink, A.J.; Botha, R.F.; Poswa, X.; Senekal, M.; Badal, R.E.; Grolman, D.C.; Richards, G.A.; Feldman, C.; Boffard, K.D.; Veller, M.; et al. Antimicrobial susceptibility of Gram-negative pathogens isolated from patients with complicated intra-abdominal infections in South African hospitals (SMART Study 2004–2009): Impact of the new carbapenem breakpoints. *Surg. Infect. (Larchmt)* **2012**, *13*, 43–49. [CrossRef] [PubMed]
4. Lee, Y.L.; Chen, Y.S.; Toh, H.S.; Huang, C.C.; Liu, Y.M.; Ho, C.M.; Lu, P.L.; Ko, W.C.; Chen, Y.H.; Wang, J.H.; et al. Antimicrobial susceptibility of pathogens isolated from patients with complicated intra-abdominal infections at five medical centers in Taiwan that continuously participated in the Study for Monitoring Antimicrobial Resistance Trends (SMART) from 2006 to 2010. *Int. J. Antimicrob. Agents* **2012**, *40*, S29–S36. [PubMed]
5. Sheng, W.H.; Badal, R.E.; Hsueh, P.R. Distribution of extended-spectrum beta-lactamases, AmpC beta-lactamases, and carbapenemases among *Enterobacteriaceae* isolates causing intra-abdominal infections in the Asia-Pacific region: Results of the study for Monitoring Antimicrobial Resistance Trends (SMART). *Antimicrob. Agents Chemother.* **2013**, *57*, 2981–2988. [PubMed]
6. Zhanel, G.G.; Cheung, D.; Adam, H.; Zelenitsky, S.; Golden, A.; Schweizer, F.; Gorityala, B.; Lagacé-Wiens, P.R.; Walkty, A.; Gin, A.S.; et al. Review of Eravacycline, a Novel Fluorocycline Antibacterial Agent. *Drugs* **2016**, *76*, 567–588. [CrossRef] [PubMed]
7. Abdallah, M.; Olafisoye, O.; Cortes, C.; Urban, C.; Landman, D.; Quale, J. Activity of eravacycline against *Enterobacteriaceae* and *Acinetobacter baumannii*, including multidrug-resistant isolates, from New York City. *Antimicrob. Agents Chemother.* **2015**, *59*, 1802–1805. [CrossRef] [PubMed]
8. Livermore, D.M.; Mushtaq, S.; Warner, M.; Woodford, N. In vitro activity of eravacycline against Carbapenem-Resistant *Enterobacteriaceae* and *Acinetobacter baumannii*. *Antimicrob. Agents Chemother.* **2016**, *60*, 3840–3844. [CrossRef] [PubMed]
9. Seifert, H.; Stefanik, D.; Sutcliffe, J.A.; Higgins, P.G. In-vitro activity of the novel fluorocycline eravacycline against carbapenem non-susceptible *Acinetobacter baumannii*. *Int. J. Antimicrob. Agents* **2018**, *51*, 62–64. [CrossRef] [PubMed]
10. Zhanel, G.G.; Baxter, M.R.; Adam, H.J.; Sutcliffe, J.; Karlowsky, J.A. In vitro activity of eravacycline against 2213 Gram-negative and 2424 Gram-positive bacterial pathogens isolated in Canadian hospital laboratories: CANWARD surveillance study 2014–2015. *Diagn. Microbiol. Infect. Dis.* **2018**, *91*, 55–62. [CrossRef] [PubMed]
11. Solomkin, J.; Evans, D.; Slepavicius, A.; Lee, P.; Marsh, A.; Tsai, L.; Sutcliffe, J.A.; Horn, P.A. Assessing the efficacy and safety of eravacycline vs ertapenem in complicated intra-abdominal infections in the Investigating Gram-Negative Infections Treated with Eravacycline (IGNITE 1) trial: A randomized clinical trial. *JAMA Surg.* **2017**, *152*, 224–232. [CrossRef] [PubMed]
12. Solomkin, J.S.; Gardovskis, J.; Lawrence, K.; Montravers, P.; Sway, A.; Evans, D.; Tsai, L. IGNITE4: Results of a phase 3, randomized, multicenter, prospective trial of eravacycline vs. meropenem in the treatment of complicated intra-abdominal infections. *Clin. Infect. Dis.* **2018**. [CrossRef] [PubMed]
13. Solomkin, J.S.; Ramesh, M.K.; Cesnauskas, G.; Novikovs, N.; Stefanova, P.; Sutcliffe, J.A.; Walpole, S.M.; Horn, P.T. Phase 2, randomized, double-blind study of the efficacy and safety of two dose regimens of eravacycline versus ertapenem for adult community-acquired complicated intra-abdominal infections. *Antimicrob. Agents Chemother.* **2014**, *58*, 1847–1854. [CrossRef]
14. Higgins, J.P.; Altman, D.G.; Gotzsche, P.C.; Juni, P.; Moher, D.; Oxman, A.D.; Savovic, J.; Schulz, K.F.; Weeks, L.; Sterne, J.A. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* **2011**, *343*, d5928. [CrossRef]
15. Goldstein, E.J.C.; Citron, D.M.; Tyrrell, K.L. In vitro activity of eravacycline and comparator antimicrobials against 143 recent strains of *Bacteroides* and *Parabacteroides* species. *Anaerobe* **2018**, *52*, 122–124. [CrossRef]
16. Snydman, D.R.; McDermott, L.A.; Jacobus, N.V.; Kerstein, K.; Grossman, T.H.; Sutcliffe, J.A. Evaluation of the in vitro activity of eravacycline against a broad spectrum of recent clinical anaerobic isolates. *Antimicrob. Agents Chemother.* **2018**, *62*, e02206-17. [CrossRef]
17. Tetraphase Pharmaceuticals Inc. Xerava (Eravacycline): US Prescribing Information. 2018. Available online: http://www.fda.gov (accessed on 23 May 2019).

Journal of
Clinical Medicine

MDPI

Article

Efficacy and Safety of Ceftaroline for the Treatment of Community-Acquired Pneumonia: A Systemic Review and Meta-Analysis of Randomized Controlled Trials

Shao-Huan Lan [1], Shen-Peng Chang [2], Chih-Cheng Lai [3], Li-Chin Lu [4] and Chien-Ming Chao [3,*]

1 School of Pharmaceutical Sciences and Medical Technology, Putian University, Putian 351100, Fujian, China; shawnlan0713@gmail.com
2 Department of Pharmacy, Chi Mei Medical Center, Liouying 73657, Taiwan; httremoon@ms.szmc.edu.tw
3 Department of Intensive Care Medicine, Chi Mei Medical Center, Liouying 73657, Taiwan; dtmed141@gmail.com
4 School of Management, Putian University, Putian 351100, Fujian, China; jane90467@gmail.com
* Correspondence: ccm870958@yahoo.com.tw

Received: 15 May 2019; Accepted: 7 June 2019; Published: 9 June 2019

Abstract: This study aimed to compare the clinical efficacy and safety of ceftaroline with those of ceftriaxone for treating community-acquired pneumonia (CAP). The PubMed, Cochrane Library, Embase, and clinicalTrials.gov databases were searched until April 2019. This meta-analysis only included randomized controlled trials (RCTs) that evaluated ceftaroline and ceftriaxone for the treatment of CAP. The primary outcome was the clinical cure rate, and the secondary outcome was the risk of adverse events (AEs). Five RCTs were included. Overall, at the test of cure (TOC), the clinical cure rate of ceftaroline was superior to the rates of ceftriaxone for the treatment of CAP (modified intent-to-treat population (MITT) population, odds ratio (OR) 1.61, 95% confidence interval (CI) 1.31–1.99, I^2 = 0%; clinically evaluable (CE) population, OR 1.38, 95% CI 1.07–1.78, I^2 = 14%). Similarly, the clinical cure rate of ceftaroline was superior to that of ceftriaxone at the end of therapy (EOT) (MITT population, OR 1.57, 95% CI 1.16–2.11, I^2 = 0%; CE population, OR 1.64, 95% CI 1.15–2.33, I^2 = 0%). For adult patients, the clinical cure rate of ceftaroline remained superior to that of ceftriaxone at TOC (MITT population, OR 1.66, 95% CI 1.34–2.06, I^2 = 0%; CE population, OR 1.39, 95% CI 1.08–1.80, I^2 = 30%) and at EOT (MITT population, OR 1.64, 95% CI 1.20–2.24, I^2 = 0%; CE population, OR 1.65, 95% CI 1.15–2.36, I^2 = 0%). Ceftaroline and ceftriaxone did not differ significantly in the risk of serious AEs, treatment-emergent AEs, and discontinuation of the study drug owing to an AE. In conclusion, the clinical efficacy of ceftaroline is similar to that of ceftriaxone for the treatment of CAP. Furthermore, this antibiotic is as tolerable as ceftriaxone.

Keywords: ceftaroline; ceftriaxone; community-acquired pneumonia; safety

1. Introduction

Community-acquired pneumonia (CAP) is a common acute bacterial infection among adults and children and has become a significant global health problem [1–4]. Moreover, severe CAP is associated with high morbidity and mortality, particularly when prompt and appropriate treatment is not provided [5,6]. However, the emergence of antibiotic resistance in this era—with the increase in resistant bacteria not treatable with existing antibiotics—and the lack of development of novel antibiotics has complicated the use of antibiotics unlike before [3,7]. In addition to the most common CAP pathogen—*Streptococcus pneumoniae*, less than 8% of CAP can be caused by the so-called PES pathogens—*Pseudomonas aeruginosa*, extended-spectrum β-lactamase producing *Enterobacteriaceae*,

and methicillin-resistant *Staphylococcus aureus* (MRSA), especially in intensive care unit (ICU) [8,9]. Among PES, MRSA is the most frequently reported, and it requires the use of specific antimicrobial agents for the treatment of typical CAP [10]. Currently, the antibiotics recommended for treating CAP when MRSA infection is suspected are vancomycin, teicoplanin, and linezolid [11–13].

Ceftaroline is a new cephalosporin with broad-spectrum activity against many commonly encountered pathogens causing CAP, including *S. pneumoniae, S. aureus, Moraxella catarrhalis, Haemophilus influenzae, and Klebsiella pneumonia* [14–16]. Moreover, several investigations have demonstrated the substantial in vitro activity of ceftaroline against MRSA from various clinical specimens, including skin/soft tissue and respiratory tract [15,17–19]. Global surveillance revealed that compared to ceftriaxone, ceftaroline showed superior in vitro activity against common CAP pathogens [17]. Subsequently, several randomized controlled trials (RCTs) [20–24] have investigated the efficacy and safety of ceftaroline for the treatment of CAP. In the present study, we conducted a comprehensive meta-analysis to provide high-quality evidence on the efficacy and safety of ceftaroline compared to those of ceftriaxone for treating CAP.

2. Methods

2.1. Study Search and Selection

All clinical studies were identified through a systematic review of the literature in the PubMed, Embase, ClinicalTrials.gov, and Cochrane databases until April 2019 using the following search terms: "ceftaroline", "Teflaro", "Zinforo", "pneumonia", and "RCT". Only RCTs that compared the clinical efficacy and adverse effects of ceftaroline and ceftriaxone were included. Two reviewers (Lan and Chang) searched and examined publications independently to avoid bias. When they disagreed, a third author (Lai) resolved the issue. The following data were extracted from each study included in the meta-analysis: year of publication, study design, duration, antibiotic regimens of ceftaroline and ceftriaxone, outcomes, and adverse events (AEs).

2.2. Definitions and Outcomes

The primary outcome was the overall clinical cure with the resolution of clinical signs and symptoms of pneumonia or improvement to the extent that no further antimicrobial therapy was necessary at the end of therapy (EOT) and test of cure (TOC) in the modified intent-to-treat population (MITT) and the clinically evaluable (CE) population. The EOT visit took place within 48 h after the last dose of oral study drug or within 24 h after the last dose of the IV study drug. The TOC visit was at 8–15 days after the last dose of the IV or oral study drug (whichever was given last). Patients in the MITT population who met minimal disease criteria and had ≥1 bacterial pathogen commonly associated with CAP identified at baseline were included in the microbiological modified MITT (mMITT) population, and those who met criteria for both the CE and mMITT populations were included in the microbiologically evaluable (ME) population. The secondary outcome was the risk of AEs, including mild, moderate, and severe degree and discontinuation because of AEs, relapse rate, and mortality.

2.3. Data Analysis

This study used the Cochrane risk-of-bias tool to assess the quality of enrolled RCTs and their risk of bias [25]. The Review Manager software program, version 5.3, was used to conduct statistical analyses. The degree of heterogeneity was evaluated using the Q statistic generated from the χ^2 test. The I^2 measure assessed the proportion of statistical heterogeneity. Heterogeneity was considered significant when the P value was less than 0.10 or the I^2 value was more than 50%. The random-effects model was used when data were significantly heterogeneous, and the fixed-effect model was used when the data were homogeneous. Pooled odds ratios (ORs) and 95% confidence intervals (CI) were calculated for outcome analyses.

3. Results

The search results yielded a total of 133 studies from the online databases, and 76 studies were excluded because of duplication. Additionally, 64 studies were found to be irrelevant after the title and abstract were screened (article type and language), and 7 studies were found to be irrelevant after the full text was screened. Eventually, five RCTs [20–24] were enrolled for the meta-analysis (Figure 1).

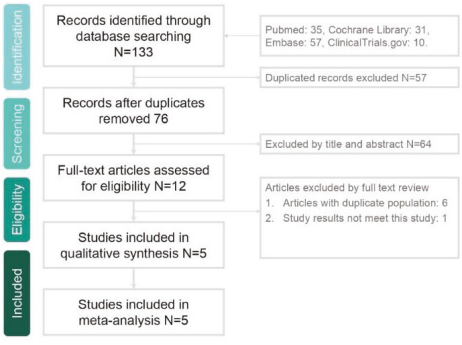


Figure 1. Flowchart of the study selection process.

3.1. Study Characteristics and Study Quality

All five RCTs [20–24] included were multinational and multicenter studies (Table 1). Three studies [20–22] focused on adult patients with CAP with Pneumonia Outcomes Research Term (PORT) [26] risk class III–IV, and two studies [23,24] enrolled pediatric patients only. Overall, the experimental group treated with ceftaroline and the control group treated with ceftriaxone comprised 1153 and 1050 patients, respectively. Almost all risks of basis in each study were low (Figure 2).

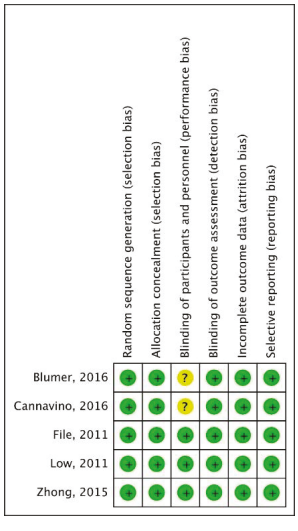


Figure 2. Risk of bias per study and domain.

Table 1. Characteristics of included studies.

Study, Published Year	Study Design	Study Period	Study Population	No of Patients		Dose Regimen	
				Ceftaroline	Comparator	Ceftaroline	Comparator
File et al., 2011 [20]	Multicenter, multinational, double-blinded, randomized trial	January 2008 to December 2008	Adult patients with PORT risk class III or IV CAP requiring hospitalization and IV therapy	304	309	600 mg q12 h	Ceftriaxone 1 g q24 h
Low et al., 2011 [21]	Multicenter, multinational, double-blinded, randomized trial	2007–2009	Patients (aged ≥18 years) with PORT risk class III or IV CAP requiring hospitalization and IV therapy	317	310	600 mg q12 h	Ceftriaxone 1 g q24 h
Zhong et al., 2015 [22]	Multicenter, multinational, double-blinded, randomized trial	2011–2013	Adult Asian patients with PORT risk class III–IV CAP	381	382	600 mg q12 h	Ceftriaxone 2 g q24 h
Cannavino et al., 2016 [23]	Multicenter, multinational, randomized	2012–2014	Ages of 2 months and <18 years with CAP requiring hospitalization and IV antibacterial therapy	121	39	Age < 6 m, 8 mg/kg q8 h; aged ≥ 6 m, 12 mg/kg q8 h for those weighing ≤ 33 kg or 400 mg q8 h for those weighing >33 kg	Ceftriaxone 75 mg/kg/d to a maximum 4g/d q12 h
Blumer et al., 2016 [24]	Multicenter, multinational randomized, observe-blinded	2012–2014	Pediatric patients between 2 months and 17 years of age with complicated CAP	30	10	15 mg/kg or 600 mg q8 h if weight > 40 kg if ≥6 m or 10 mg/kg q8 h if <6 m	Ceftriaxone, 75 mg/kg/d q12 h, and vancomycin 15 mg/kg q6 h

3.2. Clinical Efficacy

Notably, ceftaroline had a superior clinical cure rate at TOC compared with ceftriaxone for the treatment of CAP (MITT population, OR 1.61, 95% CI 1.31–1.99, I^2 = 0%; CE population, OR 1.38, 95% CI 1.07–1.78, I^2 = 14%; ME population, OR 1.98, 95% CI 1.20–3.25, I^2 = 0%; Figure 3). Similarly, at EOT, the clinical cure rate of ceftaroline was superior compared with that of ceftriaxone (MITT population, OR 1.57, 95% CI 1.16–2.11, I^2 = 0%; CE population, OR 1.64, 95% CI 1.15–2.33, I^2 = 0%).

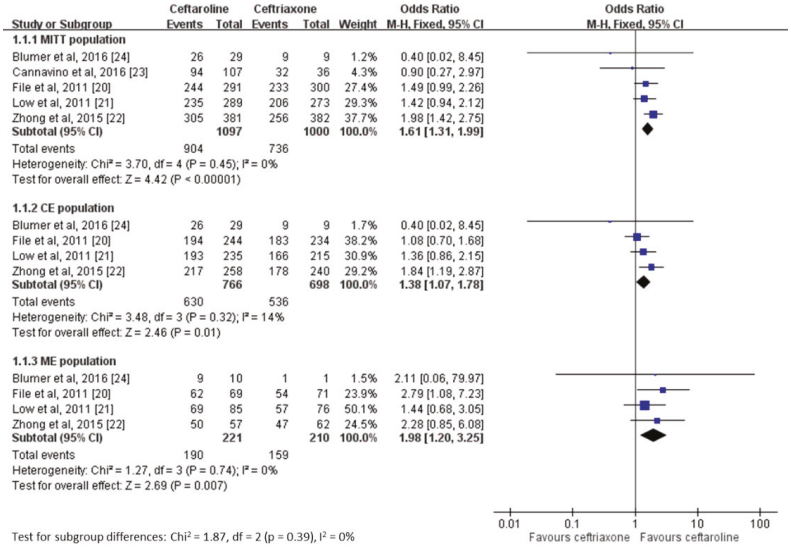


Figure 3. Overall clinical cure rates of ceftaroline and ceftriaxone for the treatment of community-acquired pneumonia. MITT, modified intent-to-treat population; CE, clinically evaluable; ME, microbiologically evaluable.

In the subgroup analysis of three studies [20–22] including adult patients, the clinical cure rate of ceftaroline remained superior to that of ceftriaxone at TOC (MITT population, OR 1.66, 95% CI 1.34–2.06, I^2 = 0%; CE population, OR 1.39, 95% CI 1.08–1.80, I^2 =30%) and at EOT (MITT population, OR 1.64, 95% CI 1.20–2.24, I^2 = 0%; CE population, OR 1.65, 95% CI 1.15–2.36, I^2 =0%). On the other hand, the pooled analysis of two studies showed that the clinical cure rates at TOC and EOT were similar between pediatric patients treated with ceftaroline or ceftriaxone (at TOC, OR 0.79, 95% CI 0.26–2.97, I^2 = 0%; at EOT, OR 1.02, 95% CI 0.38–2.75, I^2 = 0%)[23,24].

Figure 4 shows further analysis of the clinical cure rate (ceftaroline vs. ceftriaxone) at the TOC visit in various patient subgroups. Ceftaroline showed a superior clinical cure rate than ceftriaxone for patients with PORT risk III (OR 1.83, 95% CI 1.26–2.67, I^2 = 14%) but not for patients with PORT risk IV (OR 1.39, 95% CI 0.91–2.12, I^2 = 0%). The efficacy of ceftaroline was superior compared to that of ceftriaxone in patients who did not receive prior antibiotics (OR 1.90, 95% CI 1.22–2.95, I^2 = 37%) but not in those who received prior antibiotics (OR 1.18, 95% CI 0.75–1.87, I^2 = 0%). No differences were observed in the clinical cure rate between elderly patients (age ≥65 years) treated with ceftaroline or ceftriaxone (OR 1.72, 95% CI 0.95–3.11, I^2 = 58%) and between patients with bacteremia treated with ceftaroline or ceftriaxone (OR 1.62, 95% CI 0.46–5.72, I^2 = 0%).

We also assessed the clinical cure rate based on pathogens among the mMITT population, and we found that ceftaroline was superior to ceftriaxone in the overall population (OR 1.94, 95% CI 1.25–3.01, I^2 = 0%, Figure 5). Ceftaroline was superior to ceftriaxone in patients with gram-positive coccal (GPC) infection (OR 2.65, 95% CI 1.40–5.01, I^2 = 0%) but not in those with gram-negative bacterial

(GNB) infection (OR 1.26, 95% CI 0.65–2.42, I^2 = 0%). No significant difference was noted between the ceftaroline and ceftriaxone groups for each of the following pathogens: *S. pneumoniae*, *S. aureus*, *H. influenzae*, *H. parainfluenzae*, *Escherichia coli*, and *K. pneumoniae*.

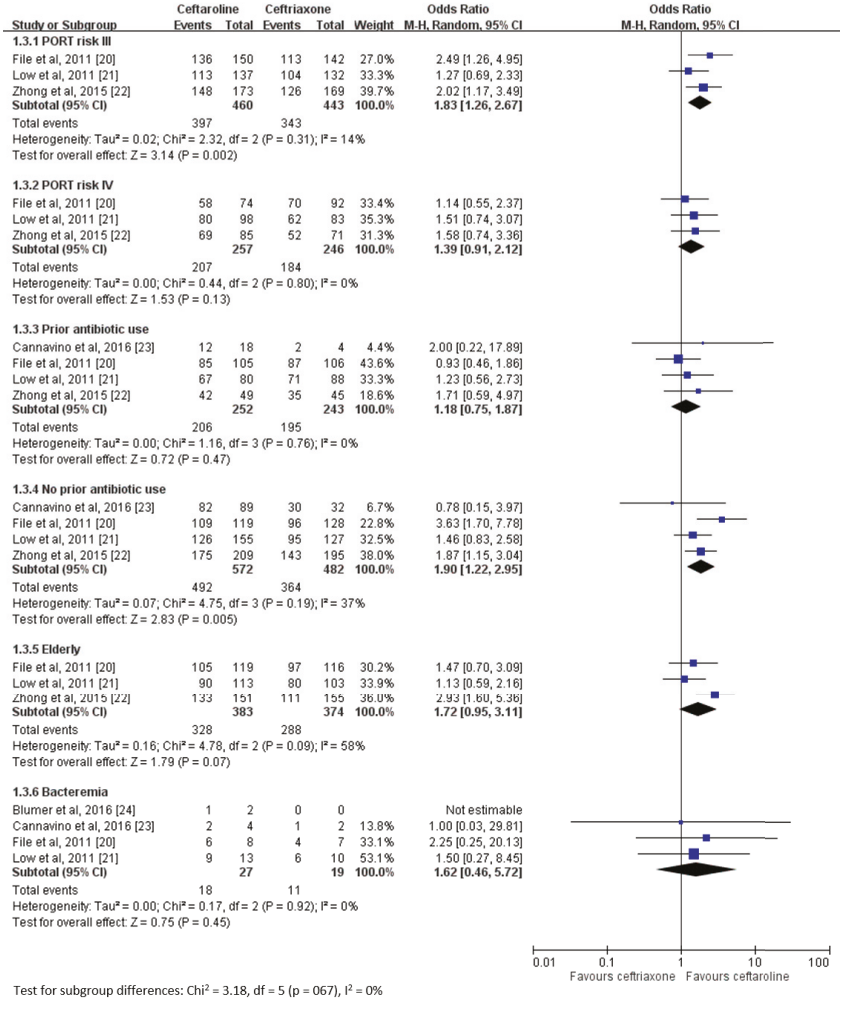


Figure 4. Overall clinical cure rates of ceftaroline and ceftriaxone for the treatment of community-acquired pneumonia based on patient group.

3.3. Adverse Events

No significant differences were observed in the risk of overall treatment-emergent adverse events (TEAEs) between the ceftaroline and ceftriaxone groups (OR 0.99, 95% CI 0.75–1.30, I^2 = 43%), and the similarity was not changed by the degree of severity (Figure 6). The risks of serious AEs and discontinuation of the study drug were similar between the ceftaroline and ceftriaxone groups (Figure 5). In addition, no relapse was noted among all enrolled patients. Finally, the mortality rate was similar between the ceftaroline and ceftriaxone groups (OR 1.13, 95% CI 0.57–2.23, I^2 = 0), and none of the cases of mortality were related to the study drug.

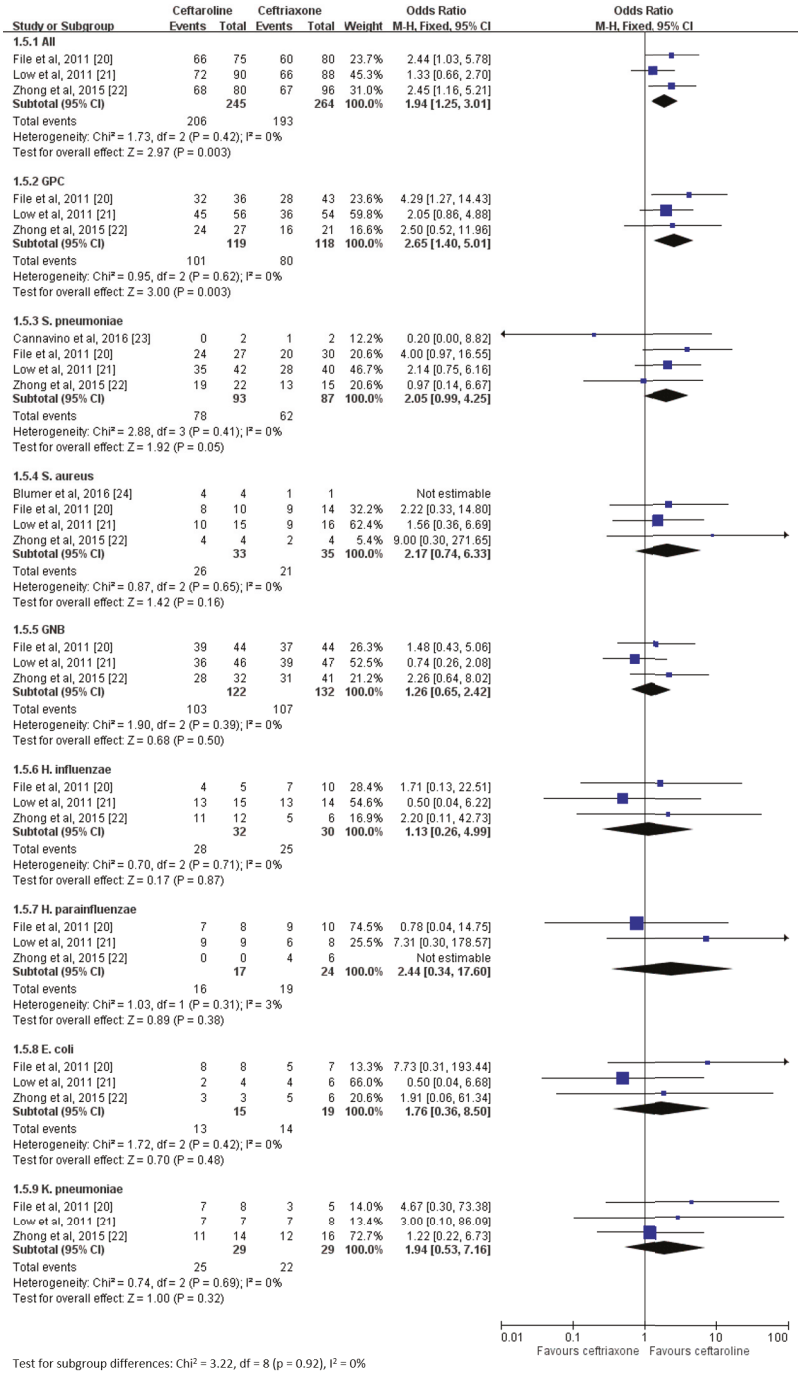


Figure 5. Overall clinical cure rates of ceftaroline and ceftriaxone for the treatment of community-acquired pneumonia based on pathogens.

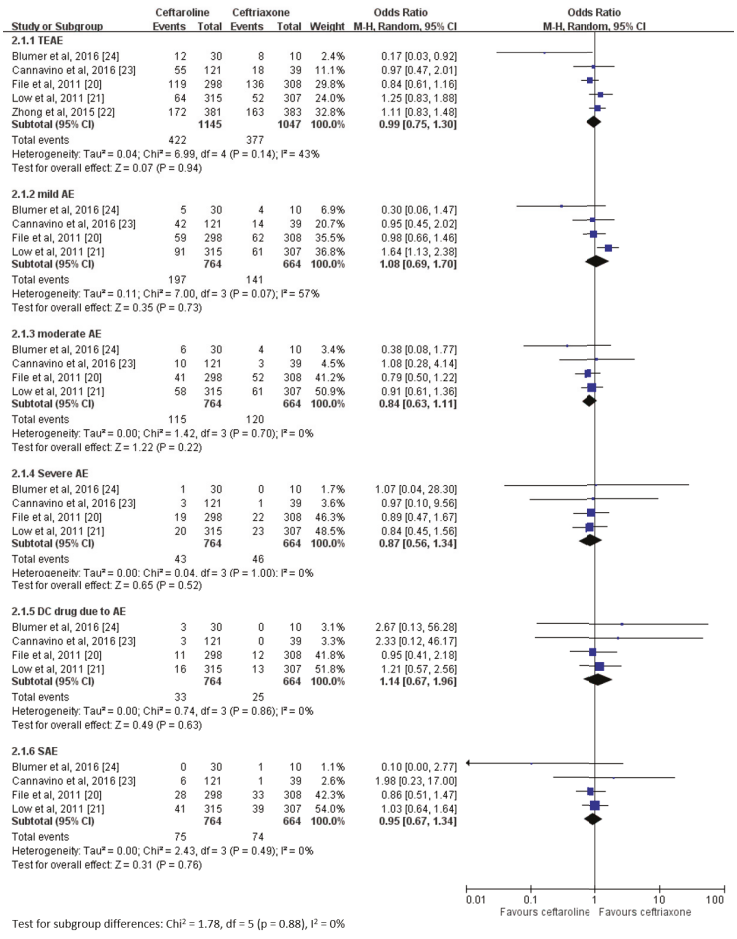


Figure 6. Risk of adverse events between ceftaroline and ceftriaxone for the treatment of community-acquired pneumonia.

4. Discussion

This meta-analysis of five RCTs determined that the clinical efficacy of ceftaroline was superior to that of ceftriaxone for the treatment of patients with CAP. First, the overall clinical cure rate of ceftaroline was superior to that of ceftriaxone for treating CAP in the pooled populations of the five RCTs, including pediatric and adult patients [20–24]. The superiority of ceftaroline compared to ceftriaxone remained significant at different times of outcome measurement, including EOT and TOC, and in different populations, including MITT, CE, and ME populations. Second, we found that ceftaroline had a higher clinical cure rate than ceftriaxone among adult patients in the subgroup analysis of three studies [20–22] including adult patients, but the pooled analysis of two studies [23,24] including pediatric patients showed similar clinical cure rates for ceftaroline and ceftriaxone. However, the two studies [23,24] that focused on pediatric patients had a limited number of patients. Therefore, more pediatric studies are warranted to clarify this issue. Third, the subgroup analysis of CAP in various populations demonstrated that ceftaroline was at least similar to ceftriaxone in patients with PORT risk IV, those who received previous antibiotics, those who were aged ≥65 years, and those with bacteremia but superior to ceftriaxone in patients with PORT risk III and those who did not

receive previous antibiotics. In summary, the overall clinical efficacy of ceftaroline is similar to that of ceftriaxone for the treatment of CAP. For other populations, ceftaroline is at least similar to ceftriaxone in terms of the clinical cure rate. However, the case numbers of several subgroup analyses, such as bacteremia, PORT risk IV or different pathogens were limited, which may limit the significance of differences between ceftaroline and ceftriaxone. Therefore, a further large-scale study is warranted to prove our findings.

In the mMITT population, the present meta-analysis determined that the clinical cure rate of ceftaroline was superior to that of ceftriaxone for CAP caused by GPC, but no significant difference was found for CAP caused by GNB, *S. pneumoniae*, *S. aureus*, *H. influenzae*, *H. parainfluenzae*, *E. coli*, and *K. pneumoniae* between the ceftaroline and ceftriaxone groups. The effectiveness of ceftaroline for the treatment of CAP is supported by in vitro studies. In a surveillance study at a US medical center, ceftaroline was noted to be more potent against *S. pneumoniae* ($MIC_{50} \leq 0.015$ vs. ≤ 0.06 μg/mL; $MIC_{90} = 0.12$ vs. 1 μg/mL) and even remained active against strains nonsusceptible to ceftriaxone ($MIC_{90} = 0.25$ μg/mL) [18]. Similar findings were demonstrated in the analysis of bacterial isolates in pediatric patients [15]. Upon global surveillance, ceftaroline was noted to be more potent than ceftriaxone against MSSA and *S. pneumoniae,* and ceftaroline had similar efficacy to ceftriaxone against *H. influenzae* [17]. Overall, the in vitro activity of ceftaroline that is greater or at least equal to ceftriaxone against most commonly encountered pathogens causing CAP could largely explain the high in vivo clinical response in this meta-analysis. However, we can only see the trend of better efficacy of ceftaroline than the comparator in each subgroup; these differences do not reach statistical significance. This may be due to the limited case number of each pathogen, so further large-scale study is warranted.

Although this study demonstrated the clinical efficacy of ceftaroline in the treatment of CAP, antibiotics may have a limited effect on the outcome of CAP, particularly these severely affected cases. This could be due to the fact that pneumonia is caused by a variety of pathogens, including respiratory viruses, Mycoplasma pneumoniae, and bacteria. In addition, incidence of primary bacterial pneumonia may be very low and be far less than that of nonbacterial pneumonia in developed countries as well as in developing countries [27]. Incidences of each pathogen pneumonia may differ in children and adults (older persons) across the populations, but severe pneumonia of viral or nonpathogen origin can induce secondary bacterial infection caused by lung injuries from primary insults; hence, it is reasonable that any pneumonia patients could be treated with antibiotics. However, antibiotics have a limited effect on the natural course of infection-related extrapulmonary manifestations. Further, outcomes of severe pneumonia may be affected by underlying comorbidities or the immune status of the host, not only by antibiotic treatment. Moreover, the pattern of antimicrobial resistance may vary in different sites; therefore, the guidelines for antibiotic treatment for CAP may differ and could be changed in each country over time. In summary, although the appropriate use of antibiotics is essential for the successful treatment of pneumonia, many factors, including disease severity, underlying comorbidity, immune status, pathogens, and the timing of antibiotic use are also significantly associated with the outcome of pneumonia.

The risk of AEs is another major concern when treating CAP with this antimicrobial agent. The most common AEs are headache, diarrhea, and insomnia [28]. In this analysis, the pooled risks of TEAEs of all degrees and even serious AEs were similar between the ceftaroline and ceftriaxone groups. Additionally, ceftaroline is associated with the risk of discontinuation of the study drug that is similar to that of ceftriaxone; this risk is because of the development of AEs. Although the overall mortality of the ceftaroline group was only 1.81%—which was comparable to that of ceftriaxone group—none of the cases of mortality were associated with the study drug. Therefore, all these findings revealed that ceftaroline is as safe as ceftriaxone for the treatment of CAPs.

A major strength of this meta-analysis is that only RCTs were included, thereby reducing the risk of bias and providing strong evidence. However, this meta-analysis also has several limitations. First, the number of MRSA-associated pneumonia cases was limited in this study. Therefore, the anti-MRSA effect of ceftaroline, which is not owned by ceftriaxone, cannot be elucidated in this meta-analysis.

Second, this meta-analysis had a limited number of studies and patients in subgroup analyses, such as different pathogens among different age groups. Therefore, some differences between the ceftaroline and ceftriaxone groups did not reach statistical significance.

5. Conclusions

In conclusion, based on the findings of this meta-analysis of five RCTs, the clinical efficacy of ceftaroline is similar to that of ceftriaxone for the treatment of CAP. Additionally, ceftaroline was as tolerable as ceftriaxone. However, clinicians should cautiously use ceftaroline in the selected population at high risk of MRSA to avoid the unnecessary coverage of MRSA by ceftaroline. Overall, ceftaroline can be recommended as an appropriate antibiotic therapy for CAP.

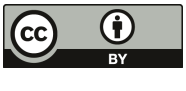
Author Contributions: Conceptualization, S.-H.L., S.-P.C., C.-C.L., C.-M.C.; methodology: S.-H.L., S.-P.C., L.-C.L.; writing—original draft preparation, C.-C.L., C.-M.C.; writing—review and editing, C.-M.C.

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

References

1. Le Roux, D.M.; Zar, H.J. Community-acquired pneumonia in children—A changing spectrum of disease. *Pediatr. Radiol.* **2017**, *47*, 1392–1398. [CrossRef] [PubMed]
2. Musher, D.M.; Thorner, A.R. Community-acquired pneumonia. *N. Engl. J. Med.* **2014**, *37*, 1619–1628. [CrossRef] [PubMed]
3. Peyrani, P.; Mandell, L.; Torres, A.; Tillotson, G.S. The burden of community-acquired bacterial pneumonia in the era of antibiotic resistance. *Expert Rev. Respir. Med.* **2019**, *13*, 139–152. [CrossRef] [PubMed]
4. Cilloniz, C.; Dominedo, C.; Garcia-Vidal, C.; Torres, A. Community-acquired pneumonia as an emergency condition. *Curr. Opin. Crit. Care* **2018**, *24*, 531–539. [CrossRef] [PubMed]
5. Leoni, D.; Rello, J. Severe community-acquired pneumonia: Optimal management. *Curr. Opin. Infect. Dis.* **2017**, *30*, 240–247. [CrossRef] [PubMed]
6. Pereira, J.M.; Goncalves-Pereira, J.; Ribeiro, O.; Baptista, J.P.; Froes, F.; Paiva, J.A. Impact of antibiotic therapy in severe community-acquired pneumonia: Data from the Infauci study. *J. Crit. Care* **2018**, *43*, 183–189. [CrossRef] [PubMed]
7. Wunderink, R.G.; Yin, Y. Antibiotic Resistance in Community-Acquired Pneumonia Pathogens. *Semin. Respir. Crit. Care Med.* **2016**, *37*, 829–838.
8. Cilloniz, C.; Dominedo, C.; Nicolini, A.; Torres, A. PES Pathogens in Severe Community-Acquired Pneumonia. *Microorganisms* **2019**, *7*, 49. [CrossRef]
9. Torres, A.; Chalmers, J.D.; Dela Cruz, C.S.; Dominedo, C.; Kollef, M.; Martin-Loeches, I.; Niederman, M.; Wunderink, R.G. Challenges in severe community-acquired pneumonia: A point-of-view review. *Intensiv. Care Med.* **2019**, *45*, 159–171. [CrossRef]
10. Thomas, R.; Ferguson, J.; Coombs, G.; Gibson, P.G. Community-acquired methicillin-resistant Staphylococcus aureus pneumonia: A clinical audit. *Respirology* **2011**, *16*, 926–931. [CrossRef]
11. Chou, C.C.; Shen, C.F.; Chen, S.J.; Chen, H.M.; Wang, Y.C.; Chang, W.S.; Chang, Y.T.; Chen, W.Y.; Huang, C.Y.; Kuo, C.C.; et al. Recommendations and guidelines for the treatment of pneumonia in Taiwan. *J. Microbiol. Immunol. Infect.* **2019**, *52*, 172–199. [CrossRef] [PubMed]
12. Liu, C.; Bayer, A.; Cosgrove, S.E.; Daum, R.S.; Fridkin, S.K.; Gorwitz, R.J.; Kaplan, S.L.; Karchmer, A.W.; Levine, D.P.; Murray, B.E.; et al. Clinical practice guidelines by the infectious diseases society of America for the treatment of methicillin-resistant *Staphylococcus aureus* infections in adults and children. *Clin. Infect. Dis.* **2011**, *52*, e18–e55. [CrossRef] [PubMed]
13. Lim, W.S.; Baudouin, S.V.; George, R.C.; Hill, A.T.; Jamieson, C.; Le Jeune, I.; Macfarlane, J.T.; Read, R.C.; Roberts, H.J.; Levy, M.L.; et al. BTS guidelines for the management of community acquired pneumonia in adults: Update 2009. *Thorax* **2009**, *64* (Suppl. 3), iii1–55. [CrossRef]
14. Pfaller, M.A.; Mendes, R.E.; Duncan, L.R.; Flamm, R.K.; Sader, H.S. In Vitro Activities of Ceftaroline and Comparators against Streptococcus pneumoniae Isolates from U.S. Hospitals: Results from Seven Years of the AWARE Surveillance Program (2010 to 2016). *Antimicrob. Agents Chemother.* **2018**, *62*, e01555-17. [CrossRef] [PubMed]

15. Pfaller, M.A.; Mendes, R.E.; Castanheira, M.; Flamm, R.K.; Jones, R.N.; Sader, H.S. Ceftaroline Activity Tested Against Bacterial Isolates Causing Community-acquired Respiratory Tract Infections and Skin and Skin Structure Infections in Pediatric Patients From United States Hospitals: 2012–2014. *Pediatr. Infect. Dis. J.* **2017**, *36*, 486–491. [CrossRef]
16. Karlowsky, J.A.; Biedenbach, D.J.; Bouchillon, S.K.; Hackel, M.; Iaconis, J.P.; Sahm, D.F. In vitro activity of Ceftaroline against bacterial pathogens isolated from patients with skin and soft tissue and respiratory tract infections in African and Middle Eastern countries: AWARE global surveillance program 2012–2014. *Diagn. Microbiol. Infect. Dis.* **2016**, *86*, 194–199. [CrossRef]
17. Biedenbach, D.J.; Iaconis, J.P.; Sahm, D.F. Comparative in vitro activities of ceftaroline and ceftriaxone against bacterial pathogens associated with respiratory tract infections: Results from the AWARE surveillance study. *J. Antimicrob. Chemother.* **2016**, *71*, 3459–3464. [CrossRef]
18. Sader, H.S.; Farrell, D.J.; Mendes, R.E.; Flamm, R.K.; Castanheira, M.; Jones, R.N. Antimicrobial activity of ceftaroline tested against bacterial isolates causing respiratory tract and skin and skin structure infections in US medical centers in 2013. *Diagn. Microbiol. Infect. Dis.* **2015**, *82*, 78–84. [CrossRef]
19. Poon, H.; Chang, M.H.; Fung, H.B. Ceftaroline fosamil: A cephalosporin with activity against methicillin-resistant *Staphylococcus aureus*. *Clin. Ther.* **2012**, *34*, 743–765. [CrossRef]
20. File, T.M., Jr.; Low, D.E.; Eckburg, P.B.; Talbot, G.H.; Friedland, H.D.; Lee, J.; Llorens, L.; Critchley, I.A.; Thye, D.A.; Pullman, J.; et al. FOCUS 1: A randomized, double-blinded, multicentre, Phase III trial of the efficacy and safety of ceftaroline fosamil versus ceftriaxone in community-acquired pneumonia. *J. Antimicrob. Chemother.* **2011**, *66* (Suppl. 3), iii19–32. [CrossRef]
21. Low, D.E.; File, T.M., Jr.; Eckburg, P.B.; Talbot, G.H.; David Friedland, H.; Lee, J.; Llorens, L.; Critchley, I.A.; Thye, D.A.; Corral, J.; et al. FOCUS 2: A randomized, double-blinded, multicentre, Phase III trial of the efficacy and safety of ceftaroline fosamil versus ceftriaxone in community-acquired pneumonia. *J. Antimicrob. Chemother.* **2011**, *66* (Suppl. 3), iii33–44. [CrossRef] [PubMed]
22. Zhong, N.S.; Sun, T.; Zhuo, C.; D'Souza, G.; Lee, S.H.; Lan, N.H.; Chiang, C.-H.; Wilson, D.; Sun, F.; Iaconis, J.; et al. Ceftaroline fosamil versus ceftriaxone for the treatment of Asian patients with community-acquired pneumonia: A randomised, controlled, double-blind, phase 3, non-inferiority with nested superiority trial. *Lancet Infect. Dis.* **2015**, *15*, 161–171. [CrossRef]
23. Cannavino, C.R.; Nemeth, A.; Korczowski, B.; Bradley, J.S.; O'Neal, T.; Jandourek, A.; Friedland, H.D.; Kaplan, S.L. A Randomized, Prospective Study of Pediatric Patients with Community-acquired Pneumonia Treated with Ceftaroline Versus Ceftriaxone. *Pediatr. Infect. Dis. J.* **2016**, *35*, 752–759. [CrossRef] [PubMed]
24. Blumer, J.L.; Ghonghadze, T.; Cannavino, C.; O'Neal, T.; Jandourek, A.; Friedland, H.D.; Bradley, J. A Multicenter, Randomized, Observer-blinded, Active-controlled Study Evaluating the Safety and Effectiveness of Ceftaroline Compared with Ceftriaxone Plus Vancomycin in Pediatric Patients With Complicated Community-acquired Bacterial Pneumonia. *Pediatr. Infect. Dis. J.* **2016**, *35*, 760–766. [CrossRef] [PubMed]
25. Higgins, J.P.; Altman, D.G.; Gotzsche, P.C.; Juni, P.; Moher, D.; Oxman, A.D.; Savović, J.; Schulz, K.F.; Weeks, L.; Sterne, J.A.C.; et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* **2011**, *343*, d5928. [CrossRef] [PubMed]
26. Fine, M.J.; Auble, T.E.; Yealy, D.M.; Hanusa, B.H.; Weissfeld, L.A.; Singer, D.E.; Coley, C.M.; Marrie, T.J.; Kapoor, W.N. A prediction rule to identify low-risk patients with community-acquired pneumonia. *N. Engl. J. Med.* **1997**, *336*, 243–250. [CrossRef] [PubMed]
27. Rhedin, S.; Lindstrand, A.; Hjelmgren, A.; Ryd-Rinder, M.; Öhrmalm, L.; Tolfvenstam, T.; Örtqvist, Å.; Rotzén-Östlund, M.; Zweygberg-Wirgart, B.; Henriques-Normark, B.; et al. Respiratory viruses associated with community-acquired pneumonia in children: Matched case-control study. *Thorax* **2015**, *70*, 847–853. [CrossRef]
28. Sotgiu, G.; Aliberti, S.; Gramegna, A.; Mantero, M.; Di Pasquale, M.; Trogu, F.; Saderi, L.; Blasi, F. Efficacy and effectiveness of Ceftaroline Fosamil in patients with pneumonia: A systematic review and meta-analysis. *Respir. Res.* **2018**, *19*, 205. [CrossRef]



Journal of
Clinical Medicine

MDPI

Article

Use of Secukinumab in a Cohort of Erythrodermic Psoriatic Patients: A Pilot Study

Giovanni Damiani [1,2,3,4,*,†], Alessia Pacifico [5,†], Filomena Russo [6], Paolo Daniele Maria Pigatto [2,3], Nicola Luigi Bragazzi [7], Claudio Bonifati [5], Aldo Morrone [5], Abdulla Watad [8,9,‡] and Mohammad Adawi [10,‡]

1 Department of Dermatology, Case Western Reserve University, Cleveland, OH 44124, USA
2 Clinical Dermatology, IRCCS Istituto Ortopedico Galeazzi, 20161 Milan, Italy; paolo.pigatto@unimi.it
3 Department of Biomedical, Surgical and Dental Sciences, University of Milan, 20122 Milan, Italy
4 Young Dermatologists Italian Network (YDIN), Centro Studi GISED, 24121 Bergamo, Italy
5 San Gallicano Dermatological Institute, IRCCS, 00144 Rome, Italy; alessia.pacifico@gmail.com (A.P.); claudio.bonifati@ifo.gov.it (C.B.); aldomorrone54@gmail.com (A.M.)
6 Dermatology Section, Department of Clinical Medicine and Immunological Science, University of Siena, 53100 Siena, Italy; file.russo@libero.it
7 School of Public Health, Department of Health Sciences (DISSAL), University of Genoa, 16132 Genoa, Italy; robertobragazzi@gmail.com
8 NIHR Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds LS7 4SA, UK; watad.abdulla@gmail.com
9 Department of Medicine 'B', Sheba Medical Center, Tel-Hashomer and Sackler Faculty of Medicine, Tel Aviv University, 5265601 Tel Aviv, Israel
10 Padeh and Ziv Hospitals, Azrieli Faculty of Medicine, Bar-Ilan University, 5290002 Ramat Gan, Israel; adawimo1802@gmail.com
* Correspondence: dr.giovanni.damiani@gmail.com; Tel.: +39-035-223-753
† These authors contributed equally to this work.
‡ These authors contributed equally to this work.

Received: 17 May 2019; Accepted: 29 May 2019; Published: 31 May 2019

Abstract: Erythrodermic psoriasis (EP) is a dermatological emergency and its treatment with secukinumab is still controversial. Furthermore, no data exist regarding the prognostic value of drug abuse in such a condition. We performed a multi-center, international, retrospective study, enrolling a sample of EP patients (body surface area > 90%) who were treated with secukinumab (300 mg) during the study period from December 2015 to December 2018. Demographics and clinical data were collected. Drug abuses were screened and, specifically, smoking status (packages/year), cannabis use (application/week) and alcoholism—tested with the Alcohol Use Disorders Identification Test (AUDIT)—were assessed. All patients were followed for up to 52 weeks. We enrolled 13 EP patients, nine males, and four females, with a median age of 40 (28–52) years. Patients naïve to biologic therapy were 3/13. Regarding drug use, seven patients had a medium-high risk of alcohol addiction, three used cannabis weekly, and seven were smokers with a pack/year index of 295 (190–365). The response rate to secukinumab was 10/13 patients with a median time to clearance of three weeks (1.5–3). No recurrences were registered in the 52-week follow-up and a Psoriasis Area Severity Index (PASI) score of 90 was achieved. The entire cohort of non-responders (n = 3) consumed at least two drugs of abuse (alcohol, smoking or cannabis). Non-responders were switched to ustekinumab and obtained a PASI 100 in 24 weeks. However, given our observed number of patients using various drugs in combination with secukinumab in EP, further studies are needed to ascertain drug abuse prevalence in a larger EP cohort. Secukinumab remains a valid, effective and safe therapeutic option for EP.

Keywords: erythrodermic psoriasis; secukinumab; addiction; smoking; alcohol; cannabis

1. Introduction

Erythroderma is an uncommon and severe dermatological manifestation of a variety of diseases. The most common form of erythroderma is erythrodermic psoriasis (EP), which accounts for 1–2.25% of all psoriatic patients, with a male predominance as demonstrated by a male to female ratio of 3:1 [1]. EP clinically manifests with diffuse erythema (body surface area (BSA) > 75%) involving also skin folds with or without exfoliate dermatitis.

Several triggers have been described to elicit EP in predisposed subjects such as environmental factors (sunburn, alcoholism, and infections), drugs (lithium, anti-malarial drugs), and the rebound phenomenon following discontinuation of anti-psoriatic treatments (oral steroids or methotrexate) [1]. However, the pathogenesis of EP remains elusive, which can limit a physician's capability to deliver safe and effective therapy. In 2010, the National Psoriasis Foundation (NPF) published a guideline describing the current evidence regarding EP treatment, stating that cyclosporine and infliximab should be the first line treatment in acute and unstable patients, whilst methotrexate and acitretin are recommended in more stable patients [2].

Despite this clear advice, prominent limitations included that few high-quality studies assessing EP treatment were present in the literature [2]. In a clinical setting, EP treatment faces two other prominent challenges, namely the difficulty in differential diagnosis and in implementing a biological approach that rules out non-inflammatory conditions. Although histological confirmation is mandatory in suspected EP cases, it is sometimes challenging due to the potential lack of histological parameters resembling classical psoriasis, such as parakeratosis or acanthosis [1].

The exclusion of neoplastic causes (Sézary syndrome) is mandatory if biologics are the selected approach. In fact, in the last 5 years, neoplasia has been a relative contraindication [3]. The NPF guidelines did not include IL-17 inhibitors [2], such as secukinumab, and recently two case series studies described the use of secukinumab in EP patients [4,5]. Current evidence seems to support the use of secukinumab in EP patients, even though there is a dearth of data concerning potential predictors of responsiveness in these patients.

Remarkably, among psoriatic patients, alcohol use/abuse and smoking are described and linked to both psoriasis development and exacerbations but are not studied in EP [6–10]. Conversely, the prevalence of cannabis users among psoriatic patients and the effect of cannabis use on psoriasis are still missing. Furthermore, in vitro or murine studies explored keratinocyte changes only in response to a single cannabis compound [11]. Thus, due to the increasing prevalence of cannabis users in the general population and also its promoting role in medicine [11], reports focusing on the effect in psoriasis are needed.

The current study aimed to evaluate (i) first the efficacy and safety of secukinumab in psoriatic erythroderma and (ii) second to describe the prevalence of drug abuses, namely alcohol, tobacco, and cannabis smoking, in EP patients.

2. Experimental Section

This multi-center, international, retrospective, pilot study enrolled a sample of EP patients (BSA > 90%) treated with a loading dose of 300 mg subcutaneous secukinumab at weeks 0, 1, 2, 3 and 4, followed by 300 mg every 4 weeks, in the period from December 2015 to December 2018.

All erythrodermic patients were biopsied and malignancies were ruled out by complete blood count, blood smear, transaminases, lactate dehydrogenase (LDH), gamma-glutamyl transferase (GGT), anion gap, Sézary cell search, and total body computed tomography. Smoking history (pack/years), cannabis use (smoking episodes/week) and alcohol use (Alcohol Use Disorders Identification Test (AUDIT)) status were assessed.

AUDIT is a 10-question screening tool (0–40 points) developed by the World Health Organization (WHO) in order to evaluate alcohol consumption, drinking behavior, and alcohol-related complications. According to AUDIT, patients are stratified as follows: 0–7 points indicate a low risk, 8–15 points a medium risk, 16–19 points a high risk, and 20–40 points a probable addiction.

All EP patients underwent a 52-week follow-up to evaluate recurrent erythrodermic episodes.

Demographics and clinical charts were recorded, including: age; gender; previous Psoriasis Area Severity Index (PASI) score before erythroderma, if any; previous anti-psoriatic therapy; biologic therapy exposure; secukinumab response; side effects; drug use history; PASI and Dermatologic life quality index (DLQI) at weeks 8, 12, 16, 24, and 52. We stratified erythroderma clearance (BSA < 75%) based on PASI 75, PASI 90, PASI 100.

3. Results

3.1. Study Population

In the current study, 13 EP patients (female/male ratio equal to 9/4), with a median age of 40 (28–52), and body mass index of 24 (22–27) kg/m^2 were included. Family history of psoriasis was positive in 9/13 patients.

3.2. Drug History

In Table 1 we assessed drug history. Only 3/13 patients were naïve to biologic therapy. Among patients treated with biologics, eight had switched more than two biologics. Furthermore, 8/13 had a previous episode of erythroderma and six patients had more than two episodes. Drug history indicated that some of the EP patients had previously received therapeutic agents that could potentially trigger psoriasis, namely four underwent beta blockers, three received angiotensin II blockers (ARBs), two patients received angiotensin-converting enzyme (ACE) inhibitors, and one patient was previously treated with thiazide diuretics.

Table 1. Pharmacological history in our cohort.

Variables	EP (n = 13)
Last anti-psoriatic therapy (*N* (%))	
Methotrexate	1 (7.7)
Phototherapy	1 (7.7)
Adalimumab	4 (30.8)
Etanercept	2 (15.4)
Ustekinumab	2 (15.4)
Apremilast	1 (7.7)
Combination therapy (MTX + Etanercept)	3 (23.1)
Biologics naïve (*N* (%))	3 (23.1)
Biologics switching (*N* (%))	10 (76.9)
1	2 (20.0)
2	5 (50.0)
3	1 (10.0)
>3	2 (20.0)
Other drugs capable to aggravate psoriasis (*N* (%))	
Beta-blockers	4 (30.8)
ACE inhibitors	2 (15.4)
ARBs	3 (23.1)
Thiazides diuretics	1 (7.7)

ACE: Angiotensin-converting enzyme, ARBs: Angiotensin II receptor blockers, EP: erythrodermic psoriasis, MTX: Methotrexate.

3.3. Drug Abuses and Comorbidities

Drug abuse screening revealed that seven patients had a medium-high risk of alcohol abuse, three patients used cannabis on a weekly basis, and seven patients were smokers with a pack/year index of 295 (190–365). The comorbidities represented in our cohort included: dyslipidemia (five patients), hypertension (three patients), osteoporosis (two patients), atrial fibrillation (one patient) and pulmonary tuberculosis (one patient), respectively (Table 2).

Table 2. Prevalence of drug abuses in our cohort.

Addictions	EP (n = 13)
Smokers (N (%))	7 (53.8)
Pack/year (median (IQR))	295 (190–365)
AUDIT test (median ± SD)	9 (4–14)
Zone I (0–7 points) (N (%))	6 (46.2)
Zone II (8–15 points) (N (%))	6 (46.2)
Zone III (16–19 points) (N (%))	1 (7.7)
Zone IV (20–40 points) (N (%))	0 (0.0)
Cannabis use (N (%))	3 (23.1)

AUDIT: Alcohol Use Disorders Identification Test, EP: erythrodermic psoriasis, IQR: Interquartile range, SD: standard deviation.

3.4. Clinical Response to Secukinumab

Clinical and therapeutic data are summarized in Table 3. The median value of the last recorded PASI was 10 (7–15). Responders to secukinumab were 10/13 (Figure 1a,b) and the median clearing time was three (1.5–3) weeks.

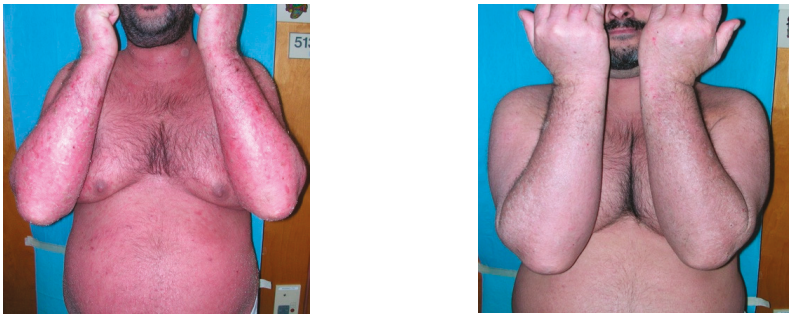
Table 3. Clinical and therapeutic records in our cohort.

Variables	EP (n = 13)
Last control PASI (median (IQR))	10 (7–15)
Secukinumab responders (N (%))	10 (76.9)
Secukinumab non-responders (N (%))	3 (23.1)
Previous erythroderma episodes (N (%))	8 (61.5)
1	2 (25.0)
2	3 (37.5)
3	1 (12.5)
>3	2 (25.0)
Erythroderma clearing time (median (IQR), weeks)	3 (1–5.3)
PASI (median (IQR))	
Week 8	15 (13–17)
PASI 75 (N (%))	4 (30.8)
PASI 90 (N (%))	0 (0.0)
PASI 100 (N (%))	0 (0.0)
Week 12	4.5 (0–10)
PASI 75 (N (%))	5 (38.5)
PASI 90 (N (%))	3 (23.1)
PASI 100 (N (%))	4 (30.8)

Table 3. *Cont.*

Variables	EP (*n* = 13)
Week 16	2 (0–5)
PASI 75 (*N* (%))	1 (7.7)
PASI 90 (*N* (%))	5 (38.5)
PASI 100 (*N* (%))	4 (30.8)
Week 24	2 (0–2.75)
PASI 75 (*N* (%))	1 (7.7)
PASI 90 (*N* (%))	5 (38.5)
PASI 100 (*N* (%))	4 (30.8)
DLQI (median (IQR))	
Week 8	17 (13–22)
Week 12	12 (9–17)
Week 16	11 (7–16)
Week 24	8 (6–12)
Week 52	8 (5–12)
Side effects (*N* (%))	5 (38.5)
Recurrent oral candidiasis	1 (20.0)
Urticaria	1 (20.0)
Injection-site pain	3 (60.0)

DLQI: Dermatologic Life Quality Index, EP: erythrodermic psoriasis, IQR: Interquartile range, MTX: Methotrexate, PASI: Psoriasis Area Severity Index.



(**a**) (**b**)

Figure 1. A 34-year-old patient with erythrodermic psoriasis that underwent secukinumab therapy. (**a**) Erythrodermic patient before treatment, (**b**) Patient after three weeks of secukinumab treatment.

Side effects were reported in 5/13 patients and remarkably were the only cause of treatment interruption, in contrast to other previously reported cases series [4,5]. All patients were on continuous secukinumab treatment and no recurrences were registered in the 52 weeks of follow up. After recovering from erythroderma at week eight, four patients achieved PASI 75, while none achieved PASI 90 or PASI 100. At week 52, five patients achieved PASI 90 and five achieved PASI 100. Interestingly, looking at the PASI trends of this cohort (Figure 2), all three non-responders used two out of three of the aforementioned drugs (alcohol, cannabis, and smoking) and no recorded comorbidities. Non-responders were switched to ustekinumab (90 mg) and obtained a PASI 100 in 24 weeks.

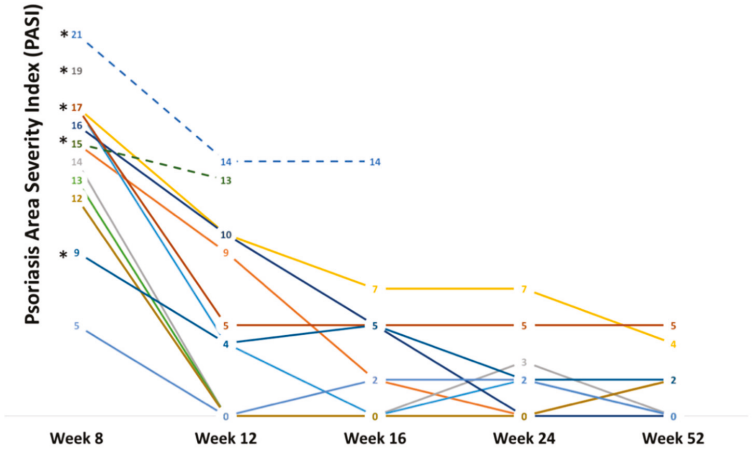


Figure 2. PASI trends in erythrodermic patients from week eight to week 52. * Patients that displayed more than one type of drug use (tobacco, cannabis, alcohol).

4. Discussion

Our study further supports that secukinumab is an effective therapy in EP and suggests that the use of recreational and accepted drugs (alcohol, cannabis, and tobacco) is prevalent in EP patients.

Furthermore, in the literature, EP patients treated with secukinumab had 16 [4] or 24 [5] months of follow up, lacking an assessment of long-term DLQI. Thus we assessed DLQI at 8, 12, 16, 24, and 52 weeks and found that secukinumab contributed to the improvement, not only in skin disease, but also in the long-term quality of life of EP patients. In our cohort, EP patients that responded to secukinumab did not exhibit recurrences and maintained long-term responsiveness to the drug. This study further supports the results described in a retrospective 52-week, observational, multicenter study, evaluated by Galluzzo et al., which suggested long-term efficacy of secukinumab in plaque psoriasis [12].

Focusing on EP patients, we assessed for the first time in detail the timing related to clearance of erythroderma, and after that, how secukinumab managed to clear the residual plaque psoriasis during the 52-week follow-up period. These two parameters together, are of pivotal importance in the clinical setting to guide therapeutic decisions made by dermatologists. In addition, the 52-week follow-up data highlighted that the EP responders to secukinumab achieved at least PASI 90 after clearing EP.

Among the three patients who did not respond to secukinumab therapy, one patient developed generalized urticaria at week three, the second patient experienced recurrent oral candidiasis and stopped the drug at week 12, and the third patient lost response at week 16. Remarkably, the second non-responder also smoked cannabis. Non-responders have not previously been treated with ustekinumab, and in accord with the recent real-life data on secukinumab non-responders, they were switched to ustekinumab and achieved a complete remission [13]. Ustekinumab is an IL-12/IL-23 blocker that targets the p40 subunit shared by these two cytokines. Furthermore, IL-12 plays a pivotal role in T helper cell type 1 (Th-1) polarization, as does IL-23 in Th-17 polarization [14]. We interpret the non-responsiveness of our patients as potentially due to the development of anti-secukinumab antibodies or up-regulation of Th-1-related pro-inflammatory cytokines, as previously demonstrated by Zaba et al. [15].

Evaluation of clinical characteristics in secukinumab non-responders indicated that all three had a familial history of EP and had used more than one drug, including smoking, alcohol, and cannabis. However, none of them were treated with any drug known to trigger psoriasis.

In the literature, the prevalence of drug abuse in the rare subset of EP is not reported, conversely, in plaque psoriasis patients several authors addressed the problem of drug abuse prevalence (alcohol and tobacco smoking) and its impact on anti-psoriatic therapies [7,8,16,17].

Alcohol intake and, consequently, also the abuse, may favor psoriasis-related systemic inflammation by promoting lipopolysaccharide (LPS) translocation from intestine to blood flow, increasing the pro-inflammatory activation of several immune cells, including lymphocytes (producing TNF-α and IFN-γ) and monocytes/macrophages (producing TNF-α), and directly by triggering keratinocytes pro-inflammatory activation via keratinocyte growth factor receptor (KGFR) [9]. These observations are further supported by Brenaut and colleagues, who conducted a meta-analysis on the epidemiological link between psoriasis and alcohol intake, and found that alcohol is a risk factor in developing psoriasis [17]. Furthermore, Qureshi et al. described a correlation between heavy beer intake and psoriasis severity during exacerbation [8]. This concept is translatable to EP patients, where erythroderma is an acute and very severe exacerbation of pre-existent psoriasis. Thus, alcohol abuse seems to increase TNF-α levels and may theoretically explain a possible lack or loss of response to IL-17 blockers, as with secukinumab in our EP patients.

Tobacco smoking and its link with psoriasis was assessed by Armstrong and colleagues in a large meta-analysis, involving 28 studies. They found an odds-ratio (OR) of 1.78 (95% confidence interval = 1.52–2.06) and a higher PASI in psoriatic smokers compared to non-smokers [16]. Remarkably, psoriasis severity gradually increases with the number of cigarettes smoked per day [17], but may benefit from a stop in smoking [18,19]. The nicotine contained in cigarettes activates nicotinic acetylcholine receptors (nAChRs) on the surface of dendritic cells, macrophages, endotheliocytes and keratinocytes, leading to an increased Th-1/Th-17 polarization of naïve T cells and to an increased production of pro-inflammatory cytokines, such as TNF-α, IL-12, IL-17, IL-23, IL-1β and IFN-γ [20]. These are all capable of decreasing the therapeutic effects of both TNF-α [7] and IL-17 blockers.

Conversely, fragmentary data exist regarding the effects of cannabis on the immune system and skin [21,22], but no data have been published about cannabis smoking in psoriatic patients or in murine models of psoriasis. However, some purified extracts derived from cannabis may inhibit in vitro keratinocyte proliferation [21] and Th-17 cell-related cytokine production in a dose-dependent manner [22]. Cannabinoids mainly interact with two receptors, cannabinoid-1 receptor (CB1R) and (CB2R), and both inhibit adenylate cyclase and activate mitogen-activated protein kinase (MAPK) [11]. This theoretically contrasts the anti-psoriatic function of apremilast, with regard to the intracellular cyclic adenosine monophosphate (AMPc) increase due to phosphodiesterase-4 inhibition. CB1R is prevalently present in keratinocytes, whilst CB2R is prevalent in immune cells, such as T cells and monocytes/macrophages [11]. Upon stimulation in the presence of purified cannabis extracts, namely cannabidiol (CBD) and tetrahydrocannabinol (THC), Th-17 cells massively decrease both transcription and release of IL-17A [22], which may theoretically act synergistically with IL-17 blockers. This aspect may be also confirmed by reports that list candidiasis as a side effect of both IL-17 blockers and chronic cannabis use [23]. Consequently, patients under IL-17 blockers that use cannabis may be exposed to a higher risk of candidiasis. Remarkably, Russo and colleagues pointed out that, in order to evaluate the global effect of cannabis, it is necessary to take into consideration the synergism existing among different cannabis compounds that altogether determine the final so-called entourage effect, capable of enhancing or even obscuring the properties of single compound [24]. Furthermore, no studies evaluated how smoking cannabis can modify these compounds and their biological effect. Thus, this is the first report to evaluate this relevant use of such drugs in a cohort of patients affected by EP, a chronic systemic inflammatory disease.

Moreover, both smoking and alcohol consumption were found to increase IL-17 and TNF-α production [9,12,16], corroborating our hypothesis that drug use may promote systemic inflammation, contributing to less favorable results from anti-psoriatic therapies.

The main limitation of the present study remains the small sample of enrolled patients, which was due to EP rarity and due to the fact secukinumab is still off-label in treating EP. Therefore, we cannot conclude that drug use in the non-responding patient group was causal. Other plausible reasons for the failed response in the small number of patients with addiction problems in the present cohort might well be insufficient compliance, even though all of our patients regularly attended dermatological appointments and reported to have auto-injected secukinumab.

5. Conclusions

Although not conclusive, our preliminary results in EP patients treated with secukinumab enlighten two presently unmet needs: (i) the need of therapy-specific biomarkers/prognostic factors and (ii) the prevalence of drug use in EP.

In conclusion, secukinumab may be a safe and effective treatment in EP, however, larger studies are needed to validate our results.

Author Contributions: Conceptualization, G.D. and A.W.; methodology, G.D., F.R. and N.L.B.; software, G.D. and N.L.B.; validation, G.D. and A.W.; formal analysis, G.D. and N.L.B.; investigation, G.D.; resources, M.A.; data curation, G.D. and N.L.B.; writing—original draft preparation, G.D. and C.B.; writing—review and editing, G.D., A.P., F.R., P.D.M.P., N.L.B., C.B., A.M., A.W. and M.A.; visualization, A.W. and M.A.; supervision, G.D., A.M. and M.A.; project administration, G.D., P.D.M.P., C.B. and A.M.; funding acquisition, M.A.

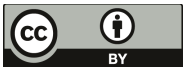
Acknowledgments: G.D. is supported by the P50 AR 070590 01A1 National Institute of Arthritis and Musculoskeletal and Skin Diseases.

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

References

1. Singh, R.K.; Lee, K.M.; Ucmak, D.; Brodsky, M.; Atanelov, Z.; Farahnik, B.; Abrouk, M.; Nakamura, M.; Zhu, T.H.; Liao, W. Erythrodermic psoriasis: Pathophysiology and current treatment perspectives. *Psoriasis (Auckl)* **2016**, *6*, 93–104.
2. Rosenbach, M.; Hsu, S.; Korman, N.J.; Lebwohl, M.G.; Young, M.; Bebo, B.F., Jr.; Van Voorhees, A.S.; National Psoriasis Foundation Medical Board. Treatment of erythrodermic psoriasis: From the medical board of the National Psoriasis Foundation. *J. Am. Acad. Dermatol.* **2010**, *62*, 655–662. [CrossRef] [PubMed]
3. Nast, A.; Gisondi, P.; Ormerod, A.D.; Saiag, P.; Smith, C.; Spuls, P.I.; Arenberger, P.; Bachelez, H.; Barker, J.; Dauden, E.; et al. European S3-Guidelines on the systemic treatment of psoriasis vulgaris–Update 2015–Short version–EDF in cooperation with EADV and IPC. *J. Eur. Acad. Dermatol. Venereol.* **2015**, *29*, 2277–2294. [CrossRef] [PubMed]
4. Mateu-Puchades, A.; Santos-Alarcón, S.; Martorell-Calatayud, A.; Pujol-Marco, C.; Sánchez-Carazo, J.L. Erythrodermic psoriasis and secukinumab: Our clinical experience. *Dermatol. Ther.* **2018**, *31*, e12607. [CrossRef] [PubMed]
5. Weng, H.J.; Wang, T.S.; Tsai, T.F. Clinical experience of secukinumab in the treatment of erythrodermic psoriasis: A case series. *Br. J. Dermatol.* **2018**, *178*, 1439–1440. [CrossRef] [PubMed]
6. Li, W.; Han, J.; Qureshi, A.A. Smoking and risk of incident psoriatic arthritis in US women. *Ann. Rheum. Dis.* **2012**, *71*, 804–808. [CrossRef]
7. Hojgaard, P.; Glintborg, B.; Hetland, M.L.; Hansen, T.H.; Lage-Hansen, P.R.; Petersen, M.H.; Holland-Fischer, M.; Nilsson, C.; Loft, A.G.; Andersen, B.N.; et al. Association between tobacco smoking and response to tumour necrosis factor alpha inhibitor treatment in psoriatic arthritis: Results from the DANBIO registry. *Ann. Rheum. Dis.* **2015**, *74*, 2130–2136. [CrossRef]
8. Qureshi, A.A.; Dominguez, P.L.; Choi, H.; Han, J.; Curhan, G. Alcohol intake and risk of incident psoriasis in US women: A prospective study. *Arch. Dermatol.* **2010**, *146*, 1364–1369. [CrossRef] [PubMed]
9. Farkas, A.; Kemeny, L. Alcohol, liver, systemic inflammation and skin: A focus on patients with psoriasis. *Skin Pharmacol. Physiol.* **2013**, *26*, 119–126. [CrossRef]
10. Poikolainen, K.; Reunala, T.; Karvonen, J.; Lauharanta, J.; Kärkkäinen, P. Alcohol intake: A risk factor for psoriasis in young and middle aged men? *BMJ* **1990**, *300*, 780–783. [CrossRef] [PubMed]

11. Katz-Talmor, D.; Katz, I.; Porat-Katz, B.S.; Shoenfeld, Y. Cannabinoids for the treatment of rheumatic diseases—Where do we stand? *Nat. Rev. Rheumatol.* **2018**, *14*, 488–498. [CrossRef]
12. Galluzzo, M.; Talamonti, M.; De Simone, C.; D'Adamio, S.; Moretta, G.; Tambone, S.; Caldarola, G.; Fargnoli, M.C.; Peris, K.; Bianchi, L. Secukinumab in moderate-to-severe plaque psoriasis: A multi-center, retrospective, real-life study up to 52 weeks observation. *Expert Opin. Biol. Ther.* **2018**, *18*, 727–735. [CrossRef]
13. Damiani, G.; Conic, R.R.Z.; de Vita, V.; Costanzo, A.; Regazzini, R.; Pigatto, P.D.M.; Bragazzi, N.L.; Pacifico, A.; Malagoli, P. When IL-17 inhibitors fail: Real-life evidence to switch from secukinumab to adalimumab or ustekinumab. *Dermatol. Ther.* **2018**, e12793. [CrossRef]
14. van Vugt, L.J.; van den Reek, J.M.P.A.; Hannink, G.; Coenen, M.J.H.; de Jong, E.M.G.J. Association of HLA-C*06:02 Status with Differential Response to Ustekinumab in Patients With Psoriasis: A Systematic Review and Meta-analysis. *JAMA Dermatol.* **2019**. [CrossRef]
15. Zaba, L.C.; Suárez-Fariñas, M.; Fuentes-Duculan, J.; Nograles, K.E.; Guttman-Yassky, E.; Cardinale, I.; Lowes, M.A.; Krueger, J.G. Effective treatment of psoriasis with etanercept is linked to suppression of IL-17 signaling, not immediate response TNF genes. *J. Allergy Clin. Immunol.* **2009**, *124*, 1022–1030.e395. [CrossRef] [PubMed]
16. Armstrong, E.J.; Harskamp, C.T.; Dhillon, J.S.; Armstrong, A.W. Psoriasis and smoking: A systematic review and meta-analysis. *Br. J. Dermatol.* **2014**, *170*, 304–314. [CrossRef] [PubMed]
17. Brenaut, E.; Horreau, C.; Pouplard, C.; Barnetche, T.; Paul, C.; Richard, M.A.; Joly, P.; Le Maître, M.; Aractingi, S.; Aubin, F.; et al. Alcohol consumption and psoriasis: A systematic literature review. *J. Eur. Acad. Dermatol. Venereol.* **2013**, *27*, 30–35. [CrossRef]
18. Fortes, C.; Mastroeni, S.; Leffondré, K.; Sampogna, F.; Melchi, F.; Mazzotti, E.; Pasquini, P.; Abeni, D. Relationship between smoking and the clinical severity of psoriasis. *Arch. Dermatol.* **2005**, *141*, 1580–1584. [CrossRef] [PubMed]
19. Richer, V.; Roubille, C.; Fleming, P.; Starnino, T.; McCourt, C.; McFarlane, A.; Siu, S.; Kraft, J.; Lynde, C.; Pope, J.E.; et al. Psoriasis and smoking: A systematic literature review and meta-analysis with qualitative analysis of effect of smoking on psoriasis severity. *J. Cutan. Med. Surg.* **2016**, *20*, 221–227. [CrossRef]
20. Armstrong, A.W.; Armstrong, E.J.; Fuller, E.N.; Sockolov, M.E.; Voyles, S.V. Smoking and pathogenesis of psoriasis: A review of oxidative, inflammatory and genetic mechanisms. *Br. J. Dermatol.* **2011**, *165*, 1162–1168. [CrossRef]
21. Wilkinson, J.D.; Williamson, E.M. Cannabinoids inhibit human keratinocyte proliferation through a non-CB1/CB2 mechanism and have a potential therapeutic value in the treatment of psoriasis. *J. Dermatol. Sci.* **2007**, *45*, 87–92. [CrossRef] [PubMed]
22. Kozela, E.; Juknat, A.; Kaushansky, N.; Rimmerman, N.; Ben-Nun, A.; Vogel, Z. Cannabinoids decrease the th17 inflammatory autoimmune phenotype. *J. Neuroimmune Pharmacol.* **2013**, *8*, 1265–1276. [CrossRef] [PubMed]
23. Dhadwal, G.; Kirchhof, M.G. The Risks and Benefits of Cannabis in the Dermatology Clinic. *J. Cutan. Med. Surg.* **2018**, *22*, 194–199. [CrossRef]
24. Russo, E.B. Taming THC: Potential cannabis synergy and phytocannabinoid-terpenoid entourage effects. *Br. J. Pharmacol.* **2011**, *163*, 1344–1364. [CrossRef] [PubMed]



Journal of
Clinical Medicine

MDPI

Article

Epidemiology and Burden of Diabetic Foot Ulcer and Peripheral Arterial Disease in Korea

Dong-il Chun [1,†], Sangyoung Kim [2,†], Jahyung Kim [1], Hyeon-Jong Yang [3], Jae Heon Kim [4], Jae-ho Cho [5], Young Yi [6], Woo Jong Kim [7] and Sung Hun Won [1,*]

1 Department of Orthopaedic Surgery, Soonchunhyang University Hospital Seoul, 59, Daesagwan-ro, Yongsan-gu, Seoul 04401, Korea; orthochun@gmail.com (D.-i.C.); hpsyndrome@naver.com (J.K.)
2 SCH Biomedical Informatics Research Unit, Soonchunhyang University Seoul Hospital, Seoul 04401, Korea; kkimsy@naver.com
3 Department of Pediatrics, Soonchunhyang University Hospital Seoul, 59, Daesagwan-ro, Yongsan-gu, Seoul 04401, Korea; pedyang@schmc.ac.kr
4 Department of Urology, Soonchunhyang University Hospital Seoul, 59, Daesagwan-ro, Yongsan-gu, Seoul 04401, Korea; piacekjh@hanmail.net
5 Department of Orthopaedic Surgery, Chuncheon Sacred Heart Hospital, Hallym University, 77, Sakju-ro, Chuncheon-si 24253, Korea; hohotoy@nate.com
6 Department of Orthopaedic Surgery, Seoul Foot and Ankle Center, Inje University, 85, 2-ga, Jeo-dong, Jung-gu, Seoul 04551, Korea; 20vvin@naver.com
7 Department of Orthopaedic Surgery, Soonchunhyang University Hospital Cheonan, 31, Soonchunhyang 6-gil, Dongnam-gu, Cheonan 31151, Korea; kwj9383@hanmail.net
* Correspondence: orthowon@gmail.com; Tel.: +82-2-709-9250; Fax: +82-2-710-3191
† Dong-il Chun and Sangyoung Kim contributed equally to this work and should be considered co-first authors.

Received: 9 May 2019; Accepted: 23 May 2019; Published: 25 May 2019

Abstract: Information about the epidemiology of diabetic foot ulcer (DFU) with peripheral arterial disease (PAD) is likely to be crucial for predicting future disease progression and establishing a health care budget. We investigated the incidence and prevalence of DFU and PAD in Korea. In addition, we examined costs of treatments for DFU and PAD. This study was conducted using data from Health Insurance Review and Assessment Service from 1 January 2011 to 31 December 2016. The incidence of DFU with PAD was 0.58% in 2012 and 0.49% in 2016. The prevalence of DFU with PAD was 1.7% in 2011 to 1.8% in 2016. The annual amputation rate of DFU with PAD was 0.95% in 2012 and 1.10% in 2016. Major amputation was decreased, while minor amputation was increased. The direct cost of each group was increased, especially the limb saving group. which was increased from 296 million dollars in 2011 to 441 million dollars in 2016. The overall incidence of DFU with PAD was about 0.5% of total population in Korea, from 2012 to 2016. Furthermore, costs for treatments of diabetic foot ulcer are increasing, especially those for the limb saving group.

Keywords: diabetic foot ulcer; peripheral arterial disease; incidence; prevalence; cost; National Health Insurance Service data

1. Introduction

The prevalence of diabetes mellitus is expected to increase and the number of diabetic patients worldwide is on the rise. The global prevalence of diabetic foot varies from 3% in Oceania to 13% in North America, with a global average of 6.4% [1]. The annual incidence of diabetic foot ulcer (DFU) or necrosis in diabetic patients is known to be about 2% to 5% and the lifetime risk ranges from 15% to 20% [2–4]. Peripheral arterial disease (PAD), like cardiovascular disease, is a major arterial disease caused by atherosclerosis [5]. Diabetes is one of the high risk factors of PAD [5], and Olinic et al. [6]

reported that the prevalence of PAD in Europe is increasing, parallel with increasing age and other risk factors for cardiovascular disease. PAD is associated with a 20-fold higher prevalence in patients with diabetes. It is known to be a risk factor for the highest severity of single factors in diabetic patients [7–9]. In addition, the probability of amputation within one year after the first ulcer or gangrene is 34.1% and the mortality rate has been reported to be 5.5% [8]. Information about the epidemiology of peripheral arterial disease associated with DFU is likely to be crucial for predicting future disease progression and establishing a health care budget.

About 20% to 33% of costs related to diabetes mellitus are used for treatments of diabetic foot [3,10]. The incidence of diabetes represented by chronic diseases is increasing. The cost of medical care for diabetic foot is increasing. Korea has recently entered an aging society. The increase in the number of diabetic patients has become an important issue in the decision of the health and welfare budget in Korea. In addition, the increase in complications due to diabetes is a burden, not only for patients, but also for the nation. Furthermore, such information is important for public health policy makers to advocate for implementation of prevention and treatment recommendations. However, there are no recent studies on the incidence, prevalence, or costs of treatments of DFU and PAD in Korea.

Thus, the primary objective of this study was to investigate the incidence and prevalence of DFU and PAD in Korea. The secondary objective was to analyze the costs of treatments for DFU and PAD using National Health Insurance Service data provided by the Health Insurance Review and Assessment Service (HIRA).

2. Materials and Methods

This study was approved by the Institutional Review Board of Soonchunhyang University Hospital Seoul (Institutional Review Board number: SCHUH 2018-01-007). The use of codes directly signifying DFU began on 1 January 2011, when the sixth edition of the Korean statistical classification of disease and related health problems-6 system (KCD-6) was applied. Until the year 2010, the disease code indicating the gangrene and ulcer was used separately from the diabetes code. If the disease code of the foot wound was not actively recorded, even if there was a DFU, DFU patients were inevitably missing. Therefore, we judged that it was not accurate to investigate data before 2010. Finally, data after 2011 were examined in this study. This study was conducted using data from HIRA from 1 January 2011 to 31 December 2016.

The annual incidence and prevalence of diabetes foot ulcer and PAD among the total population of Korea (estimated population) reported by the National Statistical Office were calculated. We considered the wash-out period as one year to determine the annual incidence of DFU and PAD. Therefore, the annual incidence of newly diagnosed DFU and PAD patients was calculated from 2012.

The amputation rates in diabetic foot ulcer and PAD patients were also calculated according to amputation level (minor vs. major (above ankle)). Diabetic foot ulcer and PAD codes and behavior codes (such as amputation, debridement, etc.) included in this study are summarized in Table 1.

Table 1. Diabetic foot ulcer (DFU) and peripheral arterial disease (PAD) codes and behavior codes included in this study.

Diabetic Foot
E105: E1050, E1051, E1058
E107: E1070, E1071, E1072, E1078
E115: E1150, E1151, E1158
E117: E1170, E1171, E1172, E1178
E125: E1250, E1251, E1258
E127: E1270, E1271, E1272, E1278
E135: E1350, E1351, E1358
E137: E1370, E1371, E1372, E1378

Table 1. *Cont.*

E145: E1450, E1451, E1458 E147: E1470, E1471, E1472, E1478
Peripheral Arterial Disease
I7022, I7023, I7024, I7025, I7029
Behavior
SC021, SC022, SC023, SC024, SC025, SC026, SC027 M0111, M0115, M0121, M0122, M0123, M0125, M0135, M0137 N0571, N0572, N0573, N0574, N0575, N0579

The direct cost for each amputation was calculated. We also analyzed the direct costs of DFU and PAD care in three groups. Group I was a limb-saving group. Group II was for those who had one amputation. Group III was for patients who had repeated amputation. The cost was based on the direct cost of patient contributions plus insurance claims. The direct cost was adjusted by taking into account the medical price index presented by the Korean Statistical Information Service (KOSIS, Daejeon, Korea). The data of this study were analyzed using SAS Enterprise Guide, ver. 6.1 M1 (SAS Institute Inc., Cary, NC, USA).

3. Results

Regarding the overall annual incidence of DFU from 2012 to 2016, 0.43% of total populations were diagnosed with DFU in 2012 whereas 0.34% were diagnosed in 2016, showing a remarkable incidence plateau with a mild decrease over five years. The annual incidence of PAD was 0.19% in 2012 and 0.20% in 2016, showing an incidence plateau with a mild increase over five years. The annual incidence of DFU with PAD was 0.58% in 2012 and 0.49% in 2016 (Figure 1). The overall prevalence of DFU in the study period was 1.4% in 2011 and 1.3% in 2016. The prevalence of PAD was 0.4% in 2011 and 0.5% in 2016. The prevalence of DFU with PAD showed a mild increase from 1.7% in 2011 to 1.8% in 2016 (Figure 1).

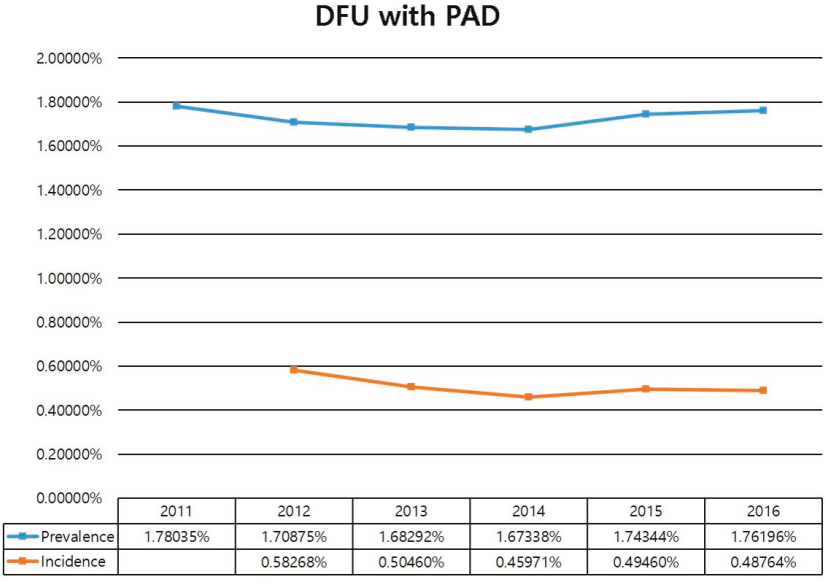


Figure 1. Annual the incidence and prevalence of diabetic foot ulcer with PAD.

The annual amputation rate of DFU with PAD was increased from 0.95% in 2012 to 1.10% in 2016. Of these, the major amputation rate was decreased from 0.28% in 2012 to 0.27% in 2016, while the minor amputation rate was increased from 0.66% in 2012 to 0.82% in 2016 (Figure 2).

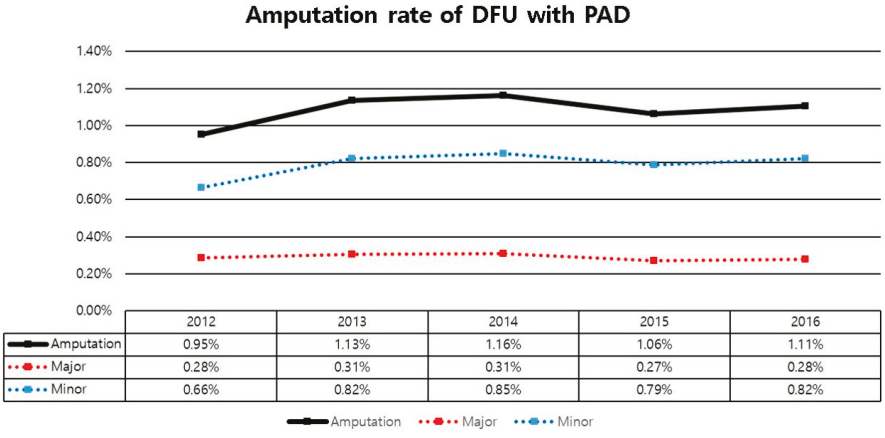


Figure 2. Annual amputation rate of diabetic foot ulcer with PAD.

The direct cost of amputation was increased from 17 million dollars in 2011 to 25 million dollars in 2016. Especially, the sum of direct costs of minor amputation increased from 11 million dollars in 2011 to 17 million dollars in 2016 (Figure 3). The average cost of amputation per person was also increased from 6100 dollars in 2011 to 7300 dollars in 2016 (Figure 4). The direct cost of each group was increased from 2011 to 2016. Direct costs for group 1 increased from 296 million dollars in 2011 to 441 million dollars in 2016. These costs for group 2 increased from 7.1 million dollars in 2011 to 9.3 million dollars in 2016, while those for group 3 increased from 10 million dollars in 2011 to 15 million dollars in 2016 (Figure 5).

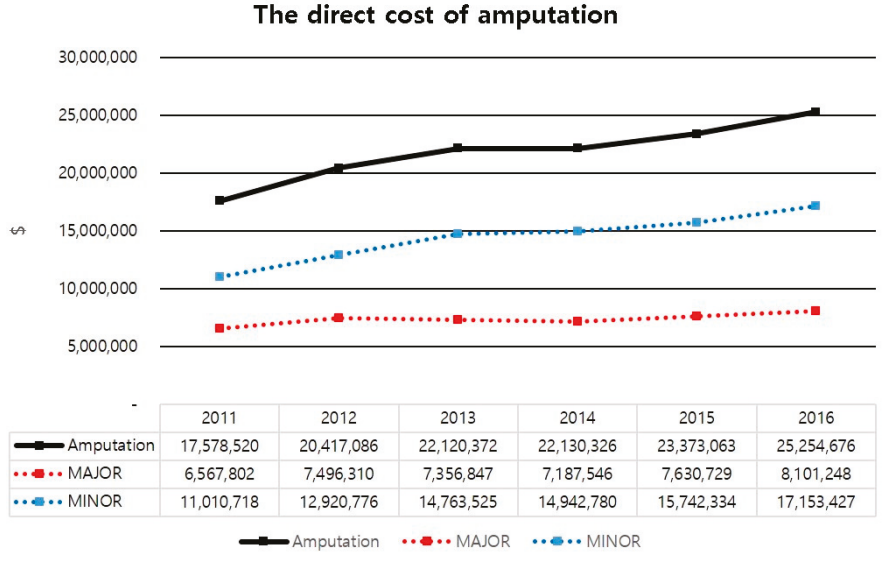


Figure 3. The direct cost of amputation of diabetic foot ulcer with PAD.

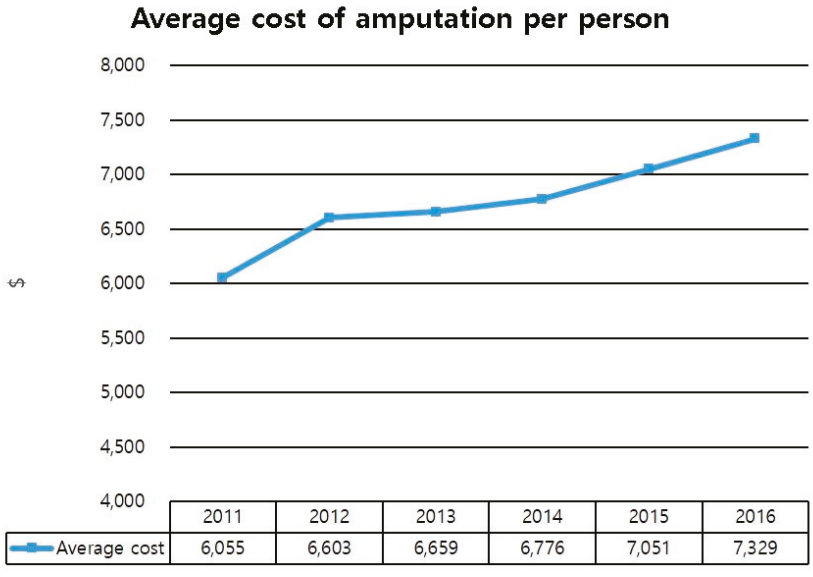


Figure 4. The average cost of amputation, per person, of diabetic foot ulcer with PAD.

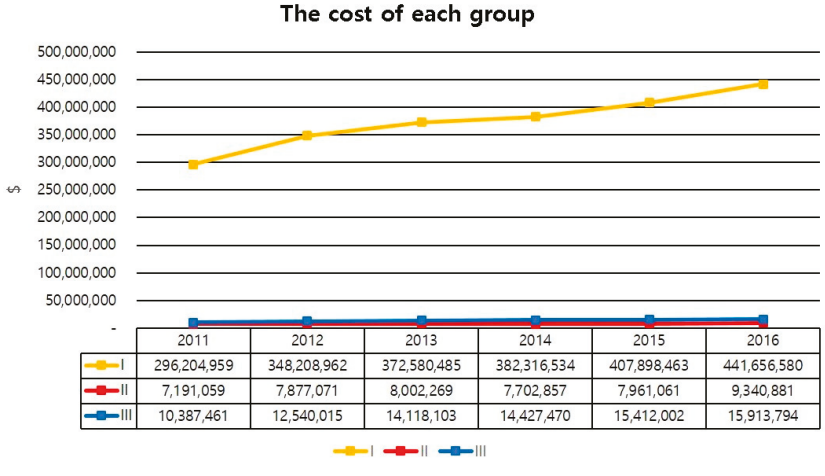


Figure 5. The direct cost of each group for diabetic foot ulcer with PAD.

4. Discussion

Overall incidence and prevalence of DFU with PAD in Korea from 2012 to 2016 were about 0.5% and 1.7% of the total population, respectively. The amputation rate was increased, especially the minor amputation rate, which increased from 0.66% in 2011 to 0.82% in 2016. Furthermore, direct costs for diabetes treatment were increased, especially the expense for the limb saving group.

We investigated the annual incidence and prevalence of DFU and PAD among the total population in Korea. The incidence and prevalence of DFU are also important for determining the number of diabetic patients. We requested HIRA to provide the total data for diabetic patients. However, the organization explained to us that these data were too large to release. Therefore, we could not obtain information for the total number of diabetic patients during the study period. Thus, we calculated the incidence and prevalence of DFU patients in the total population. However,

the Korean diabetic association reported that the prevalence of diabetes increased from 12.4% in 2012 to 14.4% in 2016 through a diabetic fact sheet in 2018. Thus, we could investigate the prevalence of diabetic foot ulcer among the diabetic patients indirectly, showing 10% in 2012 and 9% in 2016. Recently, Zhang et al. [1] reported that the global prevalence of diabetic foot ulceration is 6.3% and the prevalence is 13.0% in North America and 5.5% in Asia. However, in their systematic review and meta-analysis study, the definitions for diabetic foot and diabetic foot ulceration were ambiguous. Furthermore, two epidemiologic studies using only Korean data focused on the epidemiology of diabetic peripheral neuropathy [11,12]. However, our study investigated not only DFU, but also PAD. Thus, we believe that our study has a more accurate prevalence of DFU in Korea.

Concerning amputation, our results showed that annual amputation rate of DFU with PAD increased from 0.95% in 2012 to 1.10% in 2016. Of these, the major amputation rate decreased from 0.28% in 2012 to 0.27% in 2016, while the minor amputation rate was increased from 0.66% in 2012 to 0.82% in 2016. The overall amputation rate was increased. This might be due to the increased minor amputation rate, rather than the decrease of the major amputation rate. Goodney et al. reported that lower extremity amputation decreased by 45% from 1996 to 2011 (above the knee amputation decreased by 48% and below the knee amputation decreased by 39%) [13]. Although, in our study, we did not show a clear causal relationship about the reason for this situation, two reasons might be important. The first one is the increased awareness of the risk of diabetic foot in diabetic patients. Previous studies have reported that education about foot care to diabetic patients is important because it is associated with a significant reduction in lower extremity amputation. In addition, monthly foot checks are associated with the reduction of major lower limb amputations in diabetic incident hemodialysis patients [14,15]. Future studies are needed to examine awareness of risk of DFU in diabetic patients. The second reason is the improvement of vascular conditions due to increase of revascularization. This is also important for the reduction of major amputation. Peripheral vascular disease is known to be the most significant risk factor for diabetic foot amputation [8]. A previous study also reported that it is evident that the increasing use of vascular and preventive care, especially among patients with diabetes, is temporally associated with lower rates of major amputation [16]. Future studies focusing on understanding this relationship are needed.

The increase in the cost of medical care, due to the increased number of diabetic patients, has been reported all over the world. It is a useful indicator for planning and enforcing health care policies and budgets [17–19]. To the best of our knowledge, our research is the first to investigate the cost of medical care for diabetic foot ulcer in Korea. Our study showed that, although the cost of amputation was increased a lot, the expense for the limb saving group increased exponentially. Such a cost increase might be due to the development of medications and dressing materials. This could increase the burden on the patient. Understanding the cost of DFU should support future decisions on investment in diabetic foot care.

Some limitations of the study need to be addressed. First, DFU and PAD codes are diverse, unclear, and sometimes missing. There was no defined disease code for DFU until 2010. Thus, data on DFU from 2007 to 2010 could not be used. This is considered the limit for using National Health Insurance Service data provided by HIRA. It is necessary to agree on codes of diabetic foot and PAD more clearly and uniformly in the future. Second, diabetic neuropathy was not included in this study, because the disease code and operational definition for diabetic neuropathy were not established in the Big Data. Furthermore, extracted data using the provided diabetic neuropathy code revealed that too many patients were included, so the extracted data could not be trusted. Therefore, in this study, diabetic neuropathy was excluded in order to improve the quality of the study, but it is thought that studies to include neuropathy in the diabetic foot should be done through the operational definition defined later. Third, when calculating the incidence of chronic diseases, such as diabetes and PAD, a wash-out period of at least 2 years should be used. However, we used wash-out period of one year to calculate incidence of DFU and PAD. This was an inevitable choice due to too much missing data. If data for

a longer period of time can be used, the wash-out period of 2 years can be used. Such study is needed in the future.

5. Conclusions

In conclusion, over 5 years, we found that the overall incidence and prevalence of DFU with PAD in Korea were about 0.5% and 1.7% of the total population, respectively. The amputation rate was increased, especially the minor amputation rate. Furthermore, the direct costs for DFU treatment were increased, especially the expense for the limb saving group. Our results suggest that we should pay attention to effective implementation of the budget when we make future health policies for diabetic foot. Further studies warrant the importance of productive and cost-effective methods for saving limbs in the healthcare system.

Author Contributions: Conceptualization, J.-h.C. and S.H.W.; Data curation, J.K.; Formal analysis, D.-i.C. and S.K.; Investigation, H.-J.Y., Y.Y., and W.J.K.; Methodology, J.H.K.

Funding: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. M20180308941). This work was supported by the Soonchunhyang University Research Fund.

Acknowledgments: We thank Mi-soon Lim, Dong won Shin for contributing to collect the full text articles.

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

References

1. Zhang, P.; Lu, J.; Jing, Y.; Tang, S.; Zhu, D.; Bi, Y. Global epidemiology of diabetic foot ulceration: A systematic review and meta-analysis (dagger). *Ann. Med.* **2017**, *49*, 106–116. [CrossRef] [PubMed]
2. Schaper, N.C.; Apelqvist, J.; Bakker, K. The international consensus and practical guidelines on the management and prevention of the diabetic foot. *Curr. Diabetes Rep.* **2003**, *3*, 475–479. [CrossRef] [PubMed]
3. Mayfield, J.A.; Reiber, G.E.; Sanders, L.J.; Janisse, D.; Pogach, L.M.; American Diabetes, A. Preventive foot care in diabetes. *Diabetes Care* **2004**, *27* (Suppl. 1), S63–S64.
4. Jeffcoate, W.J.; Harding, K.G. Diabetic foot ulcers. *Lancet* **2003**, *361*, 1545–1551. [CrossRef]
5. Katsilambros, N.L.; Tsapogas, P.C.; Arvanitis, M.P.; Tritos, N.A.; Alexiou, Z.P.; Rigas, K.L. Risk factors for lower extremity arterial disease in non-insulin-dependent diabetic persons. *Diabet. Med.* **1996**, *13*, 243–246. [CrossRef]
6. Olinic, D.M.; Spinu, M.; Olinic, M.; Homorodean, C.; Tataru, D.A.; Liew, A.; Schernthaner, G.H.; Stanek, A.; Fowkes, G.; Catalano, M. Epidemiology of peripheral artery disease in Europe: VAS Educational Paper. *Int. Angiol.* **2018**, *37*, 327–334. [CrossRef] [PubMed]
7. Hirsch, A.T.; Criqui, M.H.; Treat-Jacobson, D.; Regensteiner, J.G.; Creager, M.A.; Olin, J.W.; Krook, S.H.; Hunninghake, D.B.; Comerota, A.J.; Walsh, M.E.; et al. Peripheral arterial disease detection, awareness, and treatment in primary care. *JAMA* **2001**, *286*, 1317–1324. [CrossRef] [PubMed]
8. Won, S.H.; Chung, C.Y.; Park, M.S.; Lee, T.; Sung, K.H.; Lee, S.Y.; Kim, T.G.; Lee, K.M. Risk factors associated with amputation-free survival in patient with diabetic foot ulcers. *Yonsei Med. J.* **2014**, *55*, 1373–1378. [CrossRef] [PubMed]
9. Behrendt, C.A.; Sigvant, B.; Szeberin, Z.; Beiles, B.; Eldrup, N.; Thomson, I.A.; Venermo, M.; Altreuther, M.; Menyhei, G.; Nordanstig, J.; et al. International Variations in Amputation Practice: A VASCUNET Report. *Eur. J. Vasc. Endovasc. Surg.* **2018**, *56*, 391–399. [CrossRef] [PubMed]
10. Driver, V.R.; Fabbi, M.; Lavery, L.A.; Gibbons, G. The costs of diabetic foot: The economic case for the limb salvage team. *J. Vasc. Surg.* **2010**, *52*, 17S–22S. [CrossRef] [PubMed]
11. Kim, S.S.; Won, J.C.; Kwon, H.S.; Kim, C.H.; Lee, J.H.; Park, T.S.; Ko, K.S.; Cha, B.Y. Prevalence and clinical implications of painful diabetic peripheral neuropathy in type 2 diabetes: Results from a nationwide hospital-based study of diabetic neuropathy in Korea. *Diabetes Res. Clin. Pract.* **2014**, *103*, 522–529. [CrossRef] [PubMed]
12. Won, J.C.; Kwon, H.S.; Kim, C.H.; Lee, J.H.; Park, T.S.; Ko, K.S.; Cha, B.Y. Prevalence and clinical characteristics of diabetic peripheral neuropathy in hospital patients with Type 2 diabetes in Korea. *Diabet. Med.* **2012**, *29*, e290–e296. [CrossRef] [PubMed]

13. Goodney, P.P.; Tarulli, M.; Faerber, A.E.; Schanzer, A.; Zwolak, R.M. Fifteen-year trends in lower limb amputation, revascularization, and preventive measures among medicare patients. *JAMA Surg.* **2015**, *150*, 84–86. [CrossRef] [PubMed]
14. Dargis, V.; Pantelejeva, O.; Jonushaite, A.; Vileikyte, L.; Boulton, A.J. Benefits of a multidisciplinary approach in the management of recurrent diabetic foot ulceration in Lithuania: A prospective study. *Diabetes Care* **1999**, *22*, 1428–1431. [CrossRef] [PubMed]
15. Marn Pernat, A.; Persic, V.; Usvyat, L.; Saunders, L.; Rogus, J.; Maddux, F.W.; Lacson, E., Jr.; Kotanko, P. Implementation of routine foot check in patients with diabetes on hemodialysis: Associations with outcomes. *BMJ Open Diabetes Res. Care* **2016**, *4*, e000158. [CrossRef] [PubMed]
16. Goodney, P.P.; Holman, K.; Henke, P.K.; Travis, L.L.; Dimick, J.B.; Stukel, T.A.; Fisher, E.S.; Birkmeyer, J.D. Regional intensity of vascular care and lower extremity amputation rates. *J. Vasc. Surg.* **2013**, *57*, 1471–1480. [CrossRef] [PubMed]
17. Toscano, C.; Sugita, T.; Rosa, M.; Pedrosa, H.; Rosa, R.; Bahia, L. Annual Direct Medical Costs of Diabetic Foot Disease in Brazil: A Cost of Illness Study. *Int. J. Environ. Res. Public Health* **2018**, *15*, 89. [CrossRef] [PubMed]
18. Hicks, C.W.; Selvarajah, S.; Mathioudakis, N.; Sherman, R.E.; Hines, K.F.; Black III, J.H.; Abularrage, C.J. Burden of Infected Diabetic Foot Ulcers on Hospital Admissions and Costs. *Ann. Vasc. Surg.* **2016**, *33*, 149–158. [CrossRef] [PubMed]
19. Kerr, M.; Rayman, G.; Jeffcoate, W.J. Cost of diabetic foot disease to the National Health Service in England. *Diabet. Med.* **2014**, *31*, 1498–1504. [CrossRef] [PubMed]

Journal of
Clinical Medicine

MDPI

Article

Association between Serum Urate and Risk of Hypertension in Menopausal Women with XDH Gene

Jong-Han Lee [1], Tae Hwa Go [2], San-Hui Lee [3], Juwon Kim [1], Ji Hye Huh [4], Jang Young Kim [4], Dae Ryong Kang [2], Seongmun Jeong [5], Sang-Baek Koh [6] and Jung Ran Choi [7,*]

1 Department of Laboratory Medicine, Yonsei University Wonju College of Medicine, Wonju 26426, Korea; cello425@yonsei.ac.kr (J.-H.L.); juwon76@yonsei.ac.kr (J.K.)
2 Center of Biomedical Data Science, Institute of Genomic Cohort, Yonsei University Wonju College of Medicine, Wonju 26426, Korea; thgo1123@gmail.com (T.H.G.); dr.kang@yonsei.ac.kr (D.R.K.)
3 Department of Obstetrics and Gynecology, Yonsei University Wonju College of Medicine, Wonju 26426, Korea; sanaram@yonsei.ac.kr
4 Department of Internal Medicine, Yonsei University Wonju College of Medicine, Wonju 26426, Korea; png1212@yonsei.ac.kr (J.H.H.); kimjang713@gmail.com (J.Y.K.)
5 Genome Editing Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon 34141, Korea; likemun@gmail.com
6 Department of Preventive Medicine and Institute of Occupational and Environmental Medicine, Wonju College of Medicine, Yonsei University, Wonju 26426, Korea; kohhj@yonsei.ac.kr
7 Institute of Genomic Cohort, Yonsei University Wonju College of Medicine, Wonju 26426, Korea
* Correspondence: christinae@yonsei.ac.kr

Received: 10 April 2019; Accepted: 20 May 2019; Published: 23 May 2019

Abstract: Elevated serum urate (sUA) concentrations have been associated with an increased risk of hypertension. We aimed to examine the association of sUA concentration on the risk of hypertension in pre- and post-menopausal women and investigated the association between the polymorphism of the xanthine dehydrogenase gene and the risk of hypertension. Among 7294 women, 1415 premenopausal and 5879 postmenopausal women were recruited. Anthropometric parameters as risk factors of hypertension were identify by logistic regression models. In addition, we investigated an association between xanthine dehydrogenase gene and sUA and their combined associations on the risk of hypertension. Body mass index (BMI) and waist circumference (WC) were significantly increased in accordance to the increase of sUA levels ($p < 0.001$). Multivariate logistic regression analysis showed postmenopausal women with a high sUA and high BMI were 3.18 times more likely to have hypertension than in those with normal and lower sUA (Odds ratio: 3.18, 95% confidence interval: 2.54–3.96). Postmenopausal women with a high WC were 1.62 times more likely to have hypertension than in those with normal and lower sUA. Subjects with the AG genotype of rs206860 was found to be at lower risk of hypertension (odd ratio: 0.287, 95% confidence interval: 0.091–0.905, $p = 0.033$). This cross-sectional study indicated a high sUA is associated with a higher risk of hypertension in postmenopausal women. Further well-designed prospective studies in other populations are warranted to validate our results.

Keywords: serum urate; menopause; hypertension; xanthine dehydrogenase; cross-sectional cohort study

1. Introduction

An elevated serum urate (sUA) concentration is a common phenomenon in subjects with hypertension, insulin resistance, or obesity [1], and previous epidemiological studies have demonstrated that high sUA concentrations are associated with an increased risk of hypertension [1–5]. Although longitudinal studies showed sUA might play a role in the development of hypertension, the association

between sUA and blood pressure may be affected by various factors [5–7]. However, it is not clear whether urate elevation is the cause or a consequence of hypertension [1,6,8].

Urate is the catabolic end-product of endogenous and dietary purine metabolism in human, and is mainly produced by xanthine oxidase, which is involved in the production of reactive-oxygen species (ROS) [5,9–11]. In addition, excessive sUA accumulation can cause various diseases [11,12]. Several studies using animal models and cell cultures have identified mechanisms whereby high sUA concentrations might lead to hypertension by reducing endothelial nitric oxide release and activating the renin–angiotensin system leading to smooth muscle cell proliferation [13–16]. Furthermore, associations between sUA and metabolic disorders are gender dependent. sUA concentrations tend to be lower in women than in men, partially due to the uricosuric effect of estrogens [10]. In addition, sUA concentrations seem to be increased in both physiologic and post-surgical menopause independently of other confounders [9,17], presumably due to the uricosuric effect of estrogens and hormone replacement therapy induced sUA reduction [18]. Estrogens have some indirect uricosuric effects that can contribute to modulate the urate before and after menopause. The sUA concentration is unlikely to reflect urate production, as an increased sUA production is compensated by increased excretion to maintain sUA within the normal range [19]. To the best of our knowledge, urate production has a limited influence on sUA concentration and both, urate concentration and the production, might influence blood pressure by a different underlying mechanism.

Mammalian xanthine oxidoreductase (XOR) has the characteristics of being found in two interconvertible forms: constitutively expressed in vivo NAD+-dependent xanthine dehydrogenase (XDH, EC 1.1.1.204) and post-transcriptionally modified xanthine oxidase (XO, EC 1.1.3.22) [20]. The XOR enzyme exists in XDH form, but when released into the circulation it is converted into XO. Although XDH preferentially reduces NAD+, both forms of the enzyme can also reduce molecular oxygen to form the ROS superoxide and hydrogen peroxide. Wu et al. reported the *XDH* gene might be associated with constitutional susceptibility to hypertension [21]. XDH alters xanthine oxidase by reversible sulfhydryl oxidation or by irreversible proteolytic modification, and the production of urate results from the metabolism of purines by XDH [21,22].

We examined the association between sUA and hypertension risk in post-menopausal women, and investigated the association between the polymorphism of *XDH* gene and the risk of hypertension.

2. Methods

2.1. Study Population

This study was conducted with participants from a population-based cohort within the Korean Genome and Epidemiology Study on Atherosclerosis Risk of Rural Areas in the Korean General Population (KoGES-ARIRANG) to assess the genetic and environmental etiology of common metabolic common metabolic and cardiovascular diseases in South Koreans [23,24]. The KoGES-ARIRANG cohort study contained all adults aged 40–70 years that resided in rural areas of Wonju and Pyeongchang, Gangwon-do, Republic of Korea.

In this cross-sectional study, the baseline survey was performed from November 2005 to January 2008 and contained 28,338 adults aged 40 to 70 years. Among 17,517 women, 8666 women with a sUA level were included in this study. After excluding 736 with experience of hormone therapy, 625 with no history of menopause, and 11 missing for blood pressure, a total of 1415 premenopausal and 5879 postmenopausal women comprised in this cross-sectional study (Figure 1). Subjects who were treated antihypertensive drugs were 1991 participants. Those who experienced anti-hypertensive drugs were considered having hypertension. Additionally, we excluded the participants without genotype, 3666 participants with the genetic variations in *XDH* were eligible for this cross-sectional study. The study protocol was approved by the Institutional Review Board of Wonju Severance Christian Hospital.

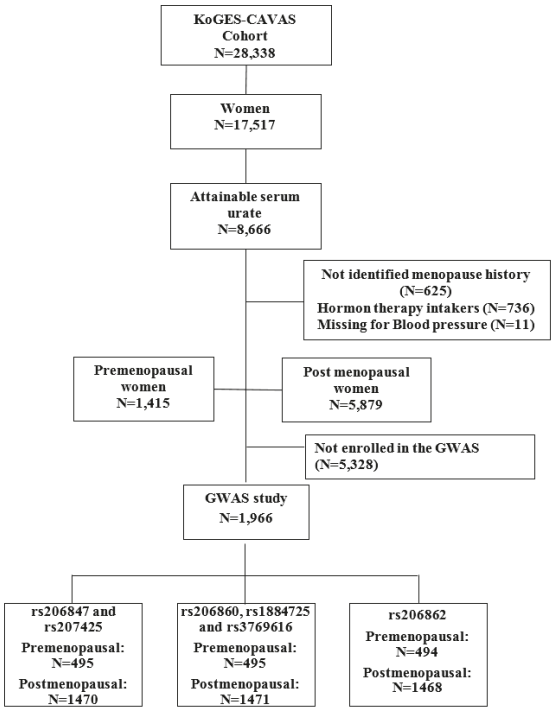


Figure 1. Flow chart of study populations in the KoGES-CAVAS cohort. Abbreviations: KoGES-CAVAS, Korean Genome and Epidemiology Study on Atherosclerosis Risk of Rural Areas in the Korean General Population; GWAS, genome wide association study.

2.2. Data Collection

At baseline examination, study participants completed a standardized medical history and lifestyle questionnaire and underwent a comprehensive health examination according to standard procedures.

Body weight and height were measured, while participants were wearing light indoor clothing without shoes. Waist circumference (WC) was measured in a horizontal plane, midway between the inferior margin of the ribs and the superior border of the iliac crest using a tape measure (SECA-200, SECA, Hamburg, Germany).

Systolic (SBP) and diastolic blood pressures (DBP) were measured twice in right arms within 5 min using a standard mercury sphygmomanometer (Baumanometer, Copiague, NY, USA). The means of the two blood pressure readings were used for the data analyses. Hypertension was defined as a SBP of ≥140 mmHg or a DBP ≥90 mmHg and/or current treatment with antihypertensive medications at the baseline survey. All the participants were examined after fasting.

A venous blood sample was drawn from study participants after fasting for >12 h or overnight. Serum aliquots were stored at −80 °C until thawed for analysis. The sUA was measured by an enzymatic coloric method that can detect the absorbance differences using uricase and peroxidase as reaction enzymes. Fasting glucose was measured using the hexokinase method. The serum concentrations of high-density lipoprotein (HDL) cholesterol, total cholesterol (TC), and triglycerides (TG) were determined using the enzymatic calorimetric method.

Alcohol and smoking habits were estimated using self-questionnaires. Individuals who had smoked ≥100 cigarettes in their lifetime were defined as current smokers, and those who had not smoked for ≥3 months were defined as ex-smokers. An interview was performed to confirm the use of medications for hypertension, and the status of regular physical exercise.

At baseline examination, study participants completed a standardized medical history and lifestyle questionnaire and underwent a comprehensive health examination according to standard procedures. At first, we estimated menopausal status using self-questionnaire. Women who had pre- and postmenopausal status were defined as "Have you always had your periods at regular 28-day intervals? Or, have you had not your periods for three months recently?" Also, an interview was performed to confirm the use of medications for hormone therapy.

2.3. Genome Wide Association Study Genotyping

Samples were analyzed using an Affymetrix Genome-Wide Human single nucleotide polymorphism (SNP) array 6.0, which contains 906,600 genome-wide SNPs and 946,000 copy number variations. Briefly, the genomic DNA was digested with two restriction enzymes (NSP I and Sty I) and processed according to the Affymetrix protocol. Digested segments were ligated to enzyme specific adaptors incorporating a universal PCR priming sequence. PCR amplification was performed using universal primers in a reaction optimized for the amplification of fragments between 200–1100 base pairs. A fragmentation step was then used to reduce the PCR products to segments of approximately 25–50 bp, which were then end-labeled using biotinylated nucleotides. The labeled products were then hybridized to a chip, washed, and detected. Images were analyzed using GeneChip Operating System software (Affymetrix, Santa Clara, CA, USA). Internal quality control measures were to ensure data fidelity, that is, a QC call rate (Dynamic Model algorithm) always was over 86% and correct identification of subject gender based on heterozygosity on the X chromosome. Genotype calling was performed using the Birdseed v2 algorithm [25].

2.4. Analysis of the XDH Genomic Polymorphism

For this study, *XDH* fragments were independently amplified by polymerase chain reaction (PCR). PCR products were purified and then sequenced using a BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) and an ABI 3730 × 1 automated sequencer (Applied Biosystems). SNPs identified in the *XDH* gene by whole gene sequencing were genotyped. Genomic DNA was extracted from 5 mL of peripheral venous blood using a commercially available isolation kit (QuickGene SP Kit DNA whole blood, Fujifilm, Tokyo, Japan). Genotyping was performed using the TaqMan fluorogenic 5′ nuclease assay (Applied Biosystems) [25].

2.5. Statistical Analysis

We analyzed the study population divided in quartiles of sUA. Categorical variables were analyzed using the chi-square test and continuous variables were analyzed by ANOVA and post hoc using Scheffe's test in pre- and post-menopausal women. Interactions between drinking status, body mass index (BMI), WC, TG, and sUA on hypertension were investigated. In order to identify an association between sUA and hypertension we analyzed multivariate logistic regression which was used to evaluate the independence of associations between sUA and risk of hypertension according to menopausal status after adjusting for fasting glucose, BMI, and WC. Additionally, there might be collinearity between BMI and WC. Consequently, we demonstrated a part of odds ratio for sUA and hypertension without considering all confounding factors including age. We used SBP as confounding factors because of one of components defined participants with hypertension and without hypertension. In additional analysis, we adjusted for age to investigate an association of sUA and clinical variables, because postmenopausal women were older than premenopausal women.

Results were expressed as ORs ratios and 95% confidence intervals (CI). All analyses were performed using SAS version 9.3 (SAS Institute, Cary, NC, USA), and SPSS version 23.0 (IBM Corp., Armonk, NY, USA). P value less than 0.05 was considered as statistically significant.

3. Results

3.1. Baseline Characteristics

The percentages of BMI, SBP, fasting glucose, and TG were significantly higher, and HDL cholesterol were significantly lower in the highest quartile of sUA (UA ≥ 5.0 mg/dL) (Table 1). BMI and WC were significantly higher in the participants with the highest quartile of sUA than those with the lowest quartile of sUA. (25.3 ± 3.4 vs. 23.6 ± 3.2; 85.2 ± 8.9 vs. 81.0 ± 9.0; respectively, $p < 0.01$) (Table 1). In the beginning of study design, we analyzed an association of sUA and hypertension with substantial confounding factors that related with sUA such as BMI and alcohol consumption. However, we did not confirm an association between sUA and hypertension after adjusted for all potential confounding factors.

Table 1. Sociodemographic characteristics of the study population.

Variable	Serum Urate (sUA)				*p*-Value	Post-Hoc *
	Quartile 1 (<3.7 mg/dL)	Quartile 2 (3.7 ≤ 4.3 mg/dL)	Quartile 3 (4.3 ≤ 5.0 mg/dL)	Quartile 4 (≥5.0 mg/dL)		
Women						
Premenopausal	411 (24.3)	385 (23.4)	381 (20.0)	238 (11.6)	<0.01	
Postmenopausal	1282 (75.7)	1258 (76.6)	1524 (80.0)	1815 (88.4)		
Age	60.5 ± 10.6	60.2 ± 10.5	60.4 ± 10.3	62.6 ± 9.6	<0.01	a, b, c < d
Smoking status						
Non, ex-smoker	1039 (96.9)	978 (96.4)	1140 (96.9)	1116 (35.1)	0.2	
Current smoker	22 (3.1)	36 (3.6)	37 (3.1)	58 (4.9)		
Alcohol						
Yes	458 (27.1)	447 (27.3)	573 (30.1)	602 (29.3)	0.11	
No	1235 (72.9)	1193 (72.7)	1330 (69.9)	1451 (70.7)		
Weight (kg)	54.6 ± 8.5	56.1 ± 8.2	57.4 ± 8.7	58.8 ± 9.5	<0.01	a < b < c < d
Height (cm)	152.0 ± 6.2	152.6 ± 6.2	152.5 ± 5.9	152.1 ± 6.0	0.02	
BMI (kg/m^2)	23.6 ± 3.2	24.1 ± 3.0	24.7 ± 3.2	25.3 ± 3.4	<0.01	a < b < c < d
WC (cm)	81.0 ± 9.0	81.9 ± 8.8	83.4 ± 9.0	85.2 ± 8.9	<0.01	a < b < c < d
HC (cm)	91.7 ± 6.3	92.5 ± 6.2	93.3 ± 6.8	94.0 ± 6.9	<0.01	a < b < c < d
Creatinine (mg/dL)	0.8 ± 0.1	0.8 ± 0.1	0.9 ± 0.1	0.9 ± 0.3	<0.01	a < b < c < d
HDL-C (mg/dL)	47.9 ± 11.2	46.9 ± 10.9	45.8 ± 11.2	43.7 ± 10.4	<0.01	a, b > c > d
Fasting glucose (mg/dL)	97.9 ± 24.5	95.9 ± 19.3	98.0 ± 19.2	100.4 ± 21.1	<0.01	a < d, b < c < d
Total cholesterol (mg/dL)	199.0 ± 36.0	200.6 ± 34.7	203.7 ± 36.2	208.8 ± 38.6	<0.01	a < c < d, b < d
Triglycerides (mg/dL)	126.4 ± 67.2	130.7 ± 72.2	146.1 ± 83.4	171.8 ± 104.0	<0.01	a, b < c < d
Total protein (g/dL)	7.3 ± 0.4	7.3 ± 0.4	7.4 ± 0.4	7.5 ± 0.4	<0.01	a < c < d, b < d
Albumin (g/dL)	4.4 ± 0.2	4.4 ± 0.2	4.5 ± 0.2	4.5 ± 0.2	<0.01	a, b < c, d
GGT (I/U)	17.2 ± 15.7	17.8 ± 21.7	20.4 ± 28.2	24.8 ± 30.8	<0.01	a, b < c < d
SBP (mmHg)	122.3 ± 18.1	122.3 ± 17.7	123.6 ± 18.4	126.4 ± 18.0	<0.01	a, b, c < d
DBP (mmHg)	75.8 ± 10.3	75.9 ± 10.4	76.5 ± 10.4	77.7 ± 10.4	<0.01	a, b, c < d
Hypertension	545 (32.2)	569 (34.6)	748 (39.3)	1103 (53.7)	<0.01	a, b < c < d

Abbreviations: sUA, serum urate; WC, waist circumference; HC, hip circumference; BMI, body mass index; HDL-C, high-density lipoprotein cholesterol; lipoprotein cholesterol; GGT, gamma-glutamyl transferase; SBP, systolic blood pressure; DBP, diastolic blood pressure. * post-hoc analysis was conducted using Scheffe's test, a: quartile 1, b: quartile 2, c: quartile 3, d: quartile 4.

3.2. Anthropometric Characteristics of Premenopausal and Postmenopausal Women with Hypertension

In postmenopausal women, the mean sUA values were higher in the participants in whom hypertension development was observed than in those who did not develop hypertension (4.72 ± 1.25 vs. 4.35 ± 1.03, $p < 0.001$). In the premenopausal women who developed hypertension, the baseline SBP and TC were 136.00 ± 16.44 mg/dL and 197.40 ± 33.90 mg/dL, respectively, and these parameters were higher than in those who did not develop hypertension ($p < 0.001$ vs. $p = 0.002$). There were no

differences in smoking habits and the regular exercise status, between the pre- and postmenopausal women (Table 2).

Table 2. Anthropometric characteristics of premenopausal and postmenopausal women with hypertension.

Variables	Premenopausal Women (*N* = 1415)			Postmenopausal Women (*N* = 5879)		
	Without Hypertension (*N* = 1142)	With Hypertension (*N* = 273)	*p* *	Without Hypertension (*N* = 3187)	With Hypertension (*N* = 2692)	*p* *
sUA (mg/dL)	4.08 ± 0.87	4.42 ± 1.06	<0.01	4.35 ± 1.03	4.72 ± 1.25	<0.01
Age (year)	46.65 ± 3.98	48.57 ± 4.27	<0.01	63.08 ± 8.44	65.84 ± 7.82	<0.01
BMI (kg/m^2)	24.10 ± 2.96	26.19 ± 3.55	<0.01	23.94 ± 3.15	25.06 ± 3.33	<0.01
WC (cm)	79.25 ± 8.18	84.26 ± 8.65	<0.01	82.35 ± 8.96	85.25 ± 8.98	<0.01
HC (cm)	93.45 ± 6.31	96.14 ± 6.39	<0.01	92.06 ± 6.48	93.54 ± 6.83	<0.01
SBP (mmHg)	111.00 ± 11.22	136.00 ± 16.44	<0.01	116.30 ± 12.00	136.90 ± 17.92	<0.01
DBP (mmHg)	72.43 ± 8.41	88.61 ± 9.81	<0.01	73.14 ± 7.97	81.14 ± 11.08	<0.01
TC (mg/dL)	190.40 ± 32.36	197.40 ± 33.90	0.002	205.48 ± 36.26	206.94 ± 37.99	0.14
TG (mg/dL)	114.80 ± 70.17	146.30 ± 94.17	<0.01	141.50 ± 78.67	162.50 ± 95.14	<0.01
HDL (mg/dL)	47.55 ± 11.10	46.49 ± 11.40	0.16	46.48 ± 11.10	44.55 ± 10.74	<0.01
Total protein (g/dL)	7.34 ± 0.38	7.43 ± 0.42	0.01	7.35 ± 0.40	7.42 ± 0.42	<0.01
Albumin (g/dL)	4.42 ± 0.22	4.44 ± 0.26	0.22	4.43 ± 0.23	4.48 ± 0.24	<0.01
GGT (I/U)	18.34 ± 27.48	22.56 ± 27.39	0.03	19.06 ± 20.14	22.38 ± 29.59	<0.01
Fasting glucose (mg/dL)	92.93 ± 15.41	98.30 ± 23.20	<0.01	97.10 ± 20.58	101.70 ± 23.00	<0.01
Creatinine (mg/dL)	0.85 ± 0.08	0.85 ± 0.11	0.88	0.86 ± 0.10	0.90 ± 0.25	<0.01
Current smoking (%)	7 (1.2)	2 (1.4)	0.93	90 (2.8)	66 (2.5)	0.53
Current drinker (%)	448 (39.2)	108 (39.6)	0.89	886 (27.8)	638 (23.7)	0.001
Regular exercise (%)	459 (40.2)	107 (39.2)	0.76	849 (26.6)	785 (29.2)	0.11

Abbreviations: sUA, serum urate; BMI, body mass index; WC, waist circumference; HC, hip circumference; SBP, systolic blood pressure; DBP, diastolic blood pressure; TC, total cholesterol; TG, triglyceride; HDL, high-density lipoprotein; GGT, gamma-glutamyl transferase. *, *p*-value was calculated by *t*-test.

3.3. Univariate and Multivariate Logistic Regression Analyses

In univariate logistic regression analysis, we investigated that postmenopausal women with a high sUA and high BMI were 3.13 times more likely to have hypertension than those with a normal BMI and a lower sUA (OR 3.13, 95% CI 2.67–3.66 vs. OR 1.85, 95% CI 1.63–2.11, respectively) (Table 3). A higher sUA level and TG were positively and significantly associated with the development of hypertension in postmenopausal women (OR 2.63, 95% CI 2.25–3.08) (Table 3). Pre-menopausal women with high sUA and high BMI showed 5.20 times more likely to have hypertension than those with normal BMI and a lower sUA (OR 5.20, 95% CI 3.48–7.78) (Table 3).

Multivariate logistic analysis after adjusting for age, SBP and BMI showed that a higher sUA and a high WC was found to be significantly associated with a 1.62-fold increased risk of hypertension in postmenopausal women (Table 3). After adjusting for age, SBP and BMI showed that a high sUA and a high TG were found to be significantly associated with a 2.08-fold increased risk of hypertension in postmenopausal women (OR 2.08, 95% CI 1.72–2.53) (Table 3).

Table 3. Effect of interactions between anthropometric parameters on hypertension in premenopausal and postmenopausal women.

Variables	Crude Odds Ratio (95% CI)		Adjusted Odds Ratio (95% CI)	
	Premenopausal (*N* = 1415)	Postmenopausal (*N* = 5879)	Premenopausal (*N* = 1415)	Postmenopausal (*N* = 5879)
Drinking status × sUA †				
Never, past drinker and sUA < 5 mg/dL	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Current drinker and sUA < 5 mg/dL	0.95 (0.69–1.30)	0.82 (0.71–0.95)	0.92 (0.60–1.42)	0.79 (0.66–0.95)
Never, past drinker and sUA ≥ 5 mg/dL	**2.16 (1.44–3.24)**	**1.96 (1.71–2.23)**	1.57 (0.88–2.77)	**1.91 (1.62–2.25)**
Current drinker and sUA ≥ 5 mg/dL	**2.66 (1.67–4.23)**	**1.42 (1.17–1.71)**	**2.09 (1.10–4.00)**	**1.44 (1.14–1.81)**
BMI × urate ‡				
Normal and sUA < 5 mg/dL	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Obesity and sUA < 5 mg/dL	**3.11 (2.27–4.25)**	**1.85 (1.63–2.11)**	1.15 (0.69–1.92)	**1.61 (1.33–1.94)**
Normal and sUA ≥ 5 mg/dL	**2.32 (1.34–4.00)**	**1.82 (1.56–2.12)**	1.93 (0.92–4.06)	**1.86 (1.53–2.25)**
Obesity and sUA ≥ 5 mg/dL	**5.20 (3.48–7.78)**	**3.13 (2.67–3.66)**	**2.13 (1.14–3.99)**	**3.18 (2.54–3.96)**
Waist circumference × urate †				
Normal and sUA < 5 mg/dL	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Abdominal obesity and sUA < 5 mg/dL	**2.58 (1.87–3.56)**	**1.60 (1.41–1.82)**	0.77 (0.43–1.27)	0.86 (0.71–1.03)
Normal and sUA ≥ 5 mg/dL	**2.27 (1.46–3.53)**	**1.87 (1.59–2.18)**	**1.97 (1.08–3.60)**	**1.84 (1.51–2.24)**
Abdominal obesity and sUA ≥ 5 mg/dL	**4.59 (2.98–7.06)**	**2.74 (2.35–3.20)**	1.27 (0.64–2.52)	**1.62 (1.30–2.01)**
Triglyceride × sUA †				
Normal and sUA < 5 mg/dL	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Hypertriglyceridemia and sUA < 5 mg/dL	**2.03 (1.43–2.88)**	**1.50 (1.32–1.72)**	1.26 (0.78–2.06)	1.14 (0.96–1.34)
Normal and sUA ≥ 5 mg/dL	**2.19 (1.45–3.31)**	**1.78 (1.53–2.07)**	1.65 (0.92–2.96)	**1.82 (1.51–2.20)**
Hypertriglyceridemia and sUA ≥ 5 mg/dL	**3.88 (2.49–6.04)**	**2.63 (2.25–3.08)**	**2.38 (1.28–4.43)**	**2.08 (1.72–2.53)**

Abbreviations: sUA, serum urate; BMI, body mass index. † Adjusted for age, systolic blood pressure and BMI. ‡ Adjusted for age, systolic blood pressure and waist circumference. Statistically significant results were presented by bold type.

3.4. Hypertension Risk of Participants according to Genotypes of Xanthine Dehydrogenase (XDH)

We found that participants with *XDH* rs206847 CC genotype had a risk of hypertension with statistical significance (OR=3.63). Multivariate logistic regression analysis showed a significantly lower risk of hypertension in women with rs206860AG genotypes than those with the wild type (AA) (OR = 0.26, 95% CI 0.08–0.89, p = 0.03) (Table 4). After adjusting for age, smoking status, alcohol consumption, and regular exercise, rs206826 AC genotype was associated with a decreased risk of hypertension (OR = 0.28 95 % CI 0.08–0.98). However, women with rs206847 CC and rs207425 GA genotypes had an elevated risk of hypertension.

Table 4. Association of *XDH* genetic variants and risk of hypertension.

SNP		Crude	p	Model 1	p	Model 2	p
Number	Type						
rs206847	AA	1.000 (ref)		1.000 (ref)		1.000 (ref)	
	AC	2.467 (0.677–8.997)	0.17	2.519 (0.690–9.202)	0.16	3.174 (0.807–12.491)	0.1
	CC	3.631 (0.863–15.287)	0.08	3.755 (0.887–15.900)	0.07	5.076 (1.044–24.682)	**0.04**
rs206860	AA	1.000 (ref)		1.000 (ref)		1.000 (ref)	
	AG	0.287 (0.091–0.905)	0.03	0.282 (0.089–0.893)	0.03	0.263 (0.078–0.888)	**0.03**
	GG	0.661 (0.183–2.386)	0.53	0.646 (0.178–2.342)	0.51	0.451 (0.109–1.866)	0.27
rs207425	GG	1.000 (ref)		1.000 (ref)		1.000 (ref)	
	GA	2.065 (0.770–5.537)	0.15	2.044 (0.758–5.509)	0.16	3.160 (1.015–9.841)	**0.05**
	AA	-	0.99	-	0.99	-	0.99
rs1884725	GG	1.000 (ref)		1.000 (ref)		1.000 (ref)	
	GA	1.052 (0.338–3.279)	0.93	1.011 (0.323–3.162)	0.99	0.957 (0.282–3.244)	0.94
	AA	3.668 (0.805–16.725)	0.09	3.935 (0.854–18.127)	0.08	11.162 (1.525–81.706)	**0.02**

Table 4. *Cont.*

SNP		Crude	p	Model 1	p	Model 2	p
Number	Type						
rs3769616	GG	1.000 (ref)		1.000 (ref)		1.000 (ref)	
	GA	-	0.99	-	0.99	-	0.99
	AA	1.580 (0.207–12.037)	0.66	1.562 (0.203–12.013)	0.67	9.732 (0.972–97.400)	0.05
rs206826	AA	1.000 (ref)		1.000 (ref)		1.000 (ref)	
	AC	0.351 (0.110–1.123)	0.08	0.355 (0.110–1.139)	0.08	0.282 (0.081–0.979)	**0.05**
	CC	0.907 (0.283–2.913)	0.87	0.905 (0.280–2.918)	0.87	0.588 (0.158–2.180)	0.43

Abbreviations: SNP, single nucleotide polymorphism; A, adenine, C, cytosine, G, guanine. Model 1 was adjusted for age, smoking status, alcohol consumption, regular exercise. Model 2 was adjusted for Model 1 and additionally adjusted for systolic blood pressure, total cholesterol and baseline body mass index. Statistically significant results were presented by bold type.

4. Discussion

In the present study, an elevated sUA was observed to be positively associated with an increased risk of hypertension and postmenopausal women with high sUA and BMI were 3.13 times more likely to have hypertension than those who had low sUA and BMI. Furthermore, *XDH* rs206860 AG genotype was found to be associated with the decreased risk of hypertension.

The present study showed that hyperuricemia was significantly associated with elevated risk of hypertension after adjusting for known confounders. Several potential mechanisms might explain the association between sUA and hypertension [5,26]. The first involves insulin resistance. Elevated insulin levels cause low urinary ammonium levels and predispose the precipitation of sUA [5,26]. Furthermore, insulin resistance is known to contribute to the developments of several metabolic disorders that influence the development of coronary artery disease [5,27], and to increase postmenopausal sUA concentrations. The second involves the detrimental effect of an elevated sUA concentration on renal function. Hyperuricemia leads to hypertension and renal injury via a crystal-independent mechanism by stimulating the renin–angiotensin system and inhibiting neuronal nitric oxide synthase [5,28]. Also, as sUA in rats induces sodium excretion to decrease by the epithelial sodium channel, it therefore contributes hypertension [29]. The third mechanism involves endothelial dysfunction. Experimental evidence suggests a potentially causal role for urate in the pathogenesis of hypertension and atherosclerosis [28]. It was reported that sUA concentrations are higher in postmenopausal than in premenopausal women [30]. As far as we know, in postmenopausal women who did not receive estrogen hormone therapy, sUA concentration tended to increase due to a shortage of the uricosuric effect of estrogen.

Although the association between sUA and hypertension has been already demonstrated in epidemiological and clinical studies, the nature of the interaction between sUA and hypertension remains debatable. The sUA might not be an independent risk factor of hypertension after controlling for other risk factors, but some studies have reported sUA is predictive of hypertension and renal disease development after controlling for associated risk factors. However, a recent meta-analysis including 25 studies of 97,824 participants has shown that high sUA significantly predicts systemic hypertension [16].

In fact, urate found to have several beneficial and potentially detrimental biologic effects [31]. Our results suggest hyperuricemia dose-dependently predicts higher risks of hypertension as proportions of subjects with hypertension increased significantly with sUA quartile, which is consistent with the findings of previous studies. Interestingly, we found genetic variations of the rs206860 polymorphism of *XDH* gene might lower the risk of hypertension. To the best of our knowledge, the association between polymorphism of rs206860 and hypertension was not widely reported. rs206860 was investigated in the studies of advanced liver disease [32] and anti-tuberculosis drug-induced hepatotoxicity [33]. As far as we know, other SNPs were not directly reported regarding association between heterogeneity of SNPs and hypertension.

Other SNPs of *XDH* gene associated with hypertension were also reported. Recently, Scheepers et al. reported that mean arterial pressure and DBP increased approximately 1 mmHg less in carriers of minor alleles of *XDH* rs2043013 in a European population [19]. Yang J et al. reported multivariate logistic regression analysis showed a significant association between the three SNPs of *XDH* at rs2043013 and hypertension in men: 47686C>T and 69901A>C in the recessive model, and 67873A>C (N1109T) in the dominant model [34]. Wu B et al. showed the *XDH* gene polymorphisms rs1042039, rs1054889, and rs2073316 might be associated with hypertension in the rural Han Chinese population [21].

The present study has some limitations that warrant considerations. First, subjects were not analyzed over a follow-up period, and therefore, we could not individually evaluate whether the association between sUA concentration and new-onset hypertension was relevant over a longer period. Second, the genome wide association study (GWAS) population included a smaller number of postmenopausal women with hypertension than our basic subjects for each parameter analysis. Third, our findings might be differently applied to other populations, especially younger age groups of different ethnicities. As far as we know, *XDH* rs206860 and its associations with hypertension have not been studied. rs206860 of *XDH* is unknown genetic variant so far and further studies in gene discovery and function needs to be verified. Also, cross-sectional nature of study has limited power to make conclusions about causality.

In order to investigate an association between genetic variants in *XDH* and risk of hypertension, we examined confounding factors in multivariate logistic regression analysis gradually. We identified the optimal model adjusted for age, smoking status, alcohol consumption, regular exercise, systolic blood pressure, total cholesterol, and baseline body mass index. However, we recognized that women with hypertension were few and have very wide confidence intervals because some women do not have the genotype of *XDH*. Also, it is not possible to identify whether high urate leads to hypertension or not from a cross-sectional study

However, this study showed a significant association between sUA and hypertension risk in postmenopausal women. sUA might be a useful marker to predict disease modality and progression of chronic metabolic diseases such as hypertension in clinical practice.

5. Conclusions

This study suggests that sUA concentrations might be associated with an increased risk of hypertension in postmenopausal women and that the rs206860 polymorphism of the *XDH* gene might be associated with a low risk of hypertension in Koreans. It seems that sUA may be a cost-effective, applicable parameter to evaluate hypertension risk of postmenopausal women. Further well-designed, large-scale studies in other populations are warranted to validate our results.

Author Contributions: Conceptualization, J.-H.L. and J.R.C.; methodology, J.-H.L., T.H.G. and J.R.C.; software, T.H.G. and S.J.; validation, J.-H.L., T.H.G. and J.R.C.; formal analysis, J.-H.L., T.H.G. and J.R.C.; investigation, J.-H.L. and J.R.C.; resources, J.-H.L. and J.R.C.; data curation, J.-H.L., T.H.G. and J.R.C.; writing—original draft preparation, J.-H.L.; writing—review and editing, J.-H.L. and J.R.C.; visualization, J.-H.L., T.H.G. and J.R.C.; supervision, S.-H.L., J.K., J.H.H., J.Y.K., D.R.K. and S.-B.K.; project administration, J.-H.L.; funding acquisition, J.R.C. and S.-B.K.

Funding: This study was supported by the Korea Centers for Disease Control and Prevention (2005-E71013-00, 2006-E71002-00, 2007-E71013-00, 2008-E71004-00, 2009-E71006-00, and 2010-E71003-00). This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2017R1D1A3B03034119). This study was provided with bioresources from National Biobank of Korea, the Centers for Disease Control and Prevention, Republic of Korea.

Acknowledgments: The authors would like to express their gratitude to members of the Institute of Genomic Cohort in Wonju Severance Christian Hospital for helpful discussions.

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

Abbreviations

sUA: serum urate; BMI, body mass index; WC, waist circumference; HC, hip circumference; HDL-C, high density lipoprotein cholesterol; GGT, gamma-glutamyl transferase; SBP, systolic blood pressure; DBP, diastolic blood pressure.

References

1. Chang, C.C.; Wu, C.H.; Liu, L.K.; Chou, R.H.; Kuo, C.S.; Huang, P.H.; Chen, L.K.; Lin, S.J. Association between serum uric acid and cardiovascular risk in nonhypertensive and nondiabetic individuals: The Taiwan I-Lan Longitudinal Aging Study. *Sci. Rep.* **2018**, *8*, 5234. [CrossRef]
2. Orlando, A.; Cazzaniga, E.; Giussani, M.; Palestini, P.; Genovesi, S. Hypertension in Children: Role of Obesity, Simple Carbohydrates, and Uric Acid. *Front. Public Health* **2018**, *6*, 129. [CrossRef]
3. Krishnan, E.; Kwoh, C.K.; Schumacher, H.R.; Kuller, L. Hyperuricemia and incidence of hypertension among men without metabolic syndrome. *Hypertension (Dallas Tex. 1979)* **2007**, *49*, 298–303. [CrossRef]
4. Jossa, F.; Farinaro, E.; Panico, S.; Krogh, V.; Celentano, E.; Galasso, R.; Mancini, M.; Trevisan, M. Serum uric acid and hypertension: The Olivetti heart study. *J. Hum. Hypertens.* **1994**, *8*, 677–681. [PubMed]
5. Chen, Q.; Yin, Y.J.; Chen, W.Y.; Wu, J.N.; Huang, X. Assessment of the association between serum uric acid levels and the incidence of hypertension in nonmetabolic syndrome subjects: A prospective observational study. *Medicine* **2018**, *97*, e9765. [CrossRef] [PubMed]
6. Wei, F.; Sun, N.; Cai, C.; Feng, S.; Tian, J.; Shi, W.; Xu, W.; Wang, Y.; Yang, X.; Li, W.D. Associations between serum uric acid and the incidence of hypertension: A Chinese senior dynamic cohort study. *J. Transl. Med.* **2016**, *14*, 110. [CrossRef]
7. Amos, C.I.; Wu, X.; Broderick, P.; Gorlov, I.P.; Gu, J.; Eisen, T.; Dong, Q.; Zhang, Q.; Gu, X.; Vijayakrishnan, J.; et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat. Genet.* **2008**, *40*, 616–622. [CrossRef]
8. Nieto, F.J.; Iribarren, C.; Gross, M.D.; Comstock, G.W.; Cutler, R.G. Uric acid and serum antioxidant capacity: A reaction to atherosclerosis? *Atherosclerosis* **2000**, *148*, 131–139. [CrossRef]
9. Sciacqua, A.; Perticone, M.; Tassone, E.J.; Cimellaro, A.; Miceli, S.; Maio, R.; Sesti, G.; Perticone, F. Uric acid is an independent predictor of cardiovascular events in post-menopausal women. *Int. J. Cardiol.* **2015**, *197*, 271–275. [CrossRef] [PubMed]
10. Cheng, D.; Du, R.; Wu, X.Y.; Lin, L.; Peng, K.; Ma, L.N.; Xu, Y.; Xu, M.; Chen, Y.H.; Bi, Y.F.; et al. Serum Uric Acid is Associated with the Predicted Risk of Prevalent Cardiovascular Disease in a Community-dwelling Population without Diabetes. *Biomed. Environ. Sci.* **2018**, *31*, 106–114. [CrossRef]
11. Kawamoto, R.; Ninomiya, D.; Senzaki, K.; Kumagi, T. Interaction between body mass index and serum uric acid in relation to blood pressure in community-dwelling Japanese men. *Clin. Hypertens.* **2018**, *24*, 1. [CrossRef] [PubMed]
12. Jin, M.; Yang, F.; Yang, I.; Yin, Y.; Luo, J.J.; Wang, H.; Yang, X.F. Uric acid, hyperuricemia and vascular diseases. *Front. Biosci. (Landmark Ed.)* **2012**, *17*, 656–669. [CrossRef]
13. Corry, D.B.; Eslami, P.; Yamamoto, K.; Nyby, M.D.; Makino, H.; Tuck, M.L. Uric acid stimulates vascular smooth muscle cell proliferation and oxidative stress via the vascular renin-angiotensin system. *J. Hypertens.* **2008**, *26*, 269–275. [CrossRef]
14. Kang, D.H.; Park, S.K.; Lee, I.K.; Johnson, R.J. Uric acid-induced C-reactive protein expression: Implication on cell proliferation and nitric oxide production of human vascular cells. *J. Am. Soc. Nephrol. JASN* **2005**, *16*, 3553–3562. [CrossRef]
15. Mazzali, M.; Hughes, J.; Kim, Y.G.; Jefferson, J.A.; Kang, D.H.; Gordon, K.L.; Lan, H.Y.; Kivlighn, S.; Johnson, R.J. Elevated uric acid increases blood pressure in the rat by a novel crystal-independent mechanism. *Hypertension (Dallas Tex. 1979)* **2001**, *38*, 1101–1106. [CrossRef]
16. Wang, J.; Qin, T.; Chen, J.; Li, Y.; Wang, L.; Huang, H.; Li, J. Hyperuricemia and risk of incident hypertension: A systematic review and meta-analysis of observational studies. *PLoS ONE* **2014**, *9*, e114259. [CrossRef]
17. Hak, A.E.; Choi, H.K. Menopause, postmenopausal hormone use and serum uric acid levels in US women–the Third National Health and Nutrition Examination Survey. *Arthritis Res. Ther.* **2008**, *10*, R116. [CrossRef] [PubMed]

18. Sumino, H.; Ichikawa, S.; Kanda, T.; Nakamura, T.; Sakamaki, T. Reduction of serum uric acid by hormone replacement therapy in postmenopausal women with hyperuricaemia. *Lancet (Lond. Engl.)* **1999**, *354*, 650. [CrossRef]
19. Scheepers, L.E.; Wei, F.F.; Stolarz-Skrzypek, K.; Malyutina, S.; Tikhonoff, V.; Thijs, L.; Salvi, E.; Barlassina, C.; Filipovsky, J.; Casiglia, E.; et al. Xanthine oxidase gene variants and their association with blood pressure and incident hypertension: A population study. *J. Hypertens.* **2016**, *34*, 2147–2154. [CrossRef]
20. Battelli, M.G.; Bolognesi, A.; Polito, L. Pathophysiology of circulating xanthine oxidoreductase: New emerging roles for a multi-tasking enzyme. *Biochim. Biophys. Acta* **2014**, *1842*, 1502–1517. [CrossRef]
21. Wu, B.; Hao, Y.; Shi, J.; Geng, N.; Li, T.; Chen, Y.; Sun, Z.; Zheng, L.; Li, H.; Li, N.; et al. Association between xanthine dehydrogenase tag single nucleotide polymorphisms and essential hypertension. *Mol. Med. Rep.* **2015**, *12*, 5685–5690. [CrossRef]
22. Verdecchia, P.; Schillaci, G.; Reboldi, G.; Santeusanio, F.; Porcellati, C.; Brunetti, P. Relation between serum uric acid and risk of cardiovascular disease in essential hypertension. The PIUMA study. *Hypertension (Dallas Tex. 1979)* **2000**, *36*, 1072–1078. [CrossRef]
23. Choi, J.R.; Ahn, S.V.; Kim, J.Y.; Koh, S.B.; Choi, E.H.; Lee, G.Y.; Jang, Y.E. Comparison of various anthropometric indices for the identification of a predictor of incident hypertension: The ARIRANG study. *J. Hum. Hypertens.* **2018**, *32*, 294–300. [CrossRef]
24. Kim, J.; Yoon, D.W.; Lee, S.K.; Lee, S.; Choi, K.M.; Robert, T.J.; Shin, C. Concurrent presence of inflammation and obstructive sleep apnea exacerbates the risk of metabolic syndrome: A KoGES 6-year follow-up study. *Medicine* **2017**, *96*, e4488. [CrossRef] [PubMed]
25. Choi, J.R.; Jang, Y.; Kim Yoon, S.; Park, J.K.; Sorn, S.R.; Park, M.Y.; Lee, M. The Impact of CDH13 Polymorphism and Statin Administration on TG/HDL Ratio in Cardiovascular Patients. *Yonsei Med. J.* **2015**, *56*, 1604–1612. [CrossRef] [PubMed]
26. Bonora, E.; Capaldo, B.; Perin, P.C.; Del Prato, S.; De Mattia, G.; Frittitta, L.; Frontoni, S.; Leonetti, F.; Luzi, L.; Marchesini, G.; et al. Hyperinsulinemia and insulin resistance are independently associated with plasma lipids, uric acid and blood pressure in non-diabetic subjects. The GISIR database. *Nutr. Metab. Cardiovasc. Dis. NMCD* **2008**, *18*, 624–631. [CrossRef] [PubMed]
27. Ninomiya, T.; Kubo, M.; Doi, Y.; Yonemoto, K.; Tanizaki, Y.; Rahman, M.; Arima, H.; Tsuryuya, K.; Iida, M.; Kiyohara, Y. Impact of metabolic syndrome on the development of cardiovascular disease in a general Japanese population: The Hisayama study. *Stroke* **2007**, *38*, 2063–2069. [CrossRef]
28. Kanellis, J.; Kang, D.H. Uric acid as a mediator of endothelial dysfunction, inflammation, and vascular disease. *Semin. Nephrol.* **2005**, *25*, 39–42. [CrossRef]
29. Xu, J.; Lloyd, D.J.; Hale, C.; Stanislaus, S.; Chen, M.; Sivits, G.; Vonderfecht, S.; Hecht, R.; Li, Y.S.; Lindberg, R.A.; et al. Fibroblast growth factor 21 reverses hepatic steatosis, increases energy expenditure, and improves insulin sensitivity in diet-induced obese mice. *Diabetes* **2009**, *58*, 250–259. [CrossRef]
30. Techatraisak, K.; Kongkaew, T. The association of hyperuricemia and metabolic syndrome in Thai postmenopausal women. *Climacteric J. Int. Menopause Soc.* **2017**, *20*, 552–557. [CrossRef]
31. Johnson, R.J.; Kang, D.H.; Feig, D.; Kivlighn, S.; Kanellis, J.; Watanabe, S.; Tuttle, K.R.; Rodriguez-Iturbe, B.; Herrera-Acosta, J.; Mazzali, M. Is there a pathogenetic role for uric acid in hypertension and cardiovascular and renal disease? *Hypertension (Dallas Tex. 1979)* **2003**, *41*, 1183–1190. [CrossRef] [PubMed]
32. Roberts, K.E.; Kawut, S.M.; Krowka, M.J.; Brown, R.S., Jr.; Trotter, J.F.; Shah, V.; Peter, I.; Tighiouart, H.; Mitra, N.; Handorf, E.; et al. Genetic risk factors for hepatopulmonary syndrome in patients with advanced liver disease. *Gastroenterology* **2010**, *139*, 130–139.e24. [CrossRef]
33. Li, Y.; Tang, H.; Qi, H.; Shen, C.; Sun, L.; Li, J.; Xu, F.; Jiao, W.; Yang, X.; Shen, A. rs1800796 of the IL6 gene is associated with increased risk for anti-tuberculosis drug-induced hepatotoxicity in Chinese Han children. *Tuberculosis* **2018**, *111*, 71–77. [CrossRef] [PubMed]
34. Yang, J.; Kamide, K.; Kokubo, Y.; Takiuchi, S.; Horio, T.; Matayoshi, T.; Yasuda, H.; Miwa, Y.; Yoshii, M.; Yoshihara, F.; et al. Associations of hypertension and its complications with variations in the xanthine dehydrogenase gene. *Hypertens. Res. Off. J. Jpn. Soc. Hypertens.* **2008**, *31*, 931–940. [CrossRef] [PubMed]

Journal of
Clinical Medicine

MDPI

Article

Preoperative Health-Related Quality of Life Predicts Minimal Clinically Important Difference and Survival after Surgical Resection of Hepatocellular Carcinoma

Chong-Chi Chiu [1,2,3], King-Teh Lee [4,5], Jhi-Joung Wang [6], Ding-Ping Sun [2], Hao-Hsien Lee [1], Chien-Cheng Huang [7,8] and Hon-Yi Shi [5,9,10,*]

1 Department of General Surgery, Chi Mei Medical Center, Liouying 73657, Taiwan; chiuchongchi@yahoo.com.tw (C.-C.C.); hao_hsien@hotmail.com (H.-H.L.)
2 Department of General Surgery, Chi Mei Medical Center, Tainan 71004, Taiwan; sdp0127@gmail.com
3 Department of Electrical Engineering, Southern Taiwan University of Science and Technology, Tainan 71005, Taiwan
4 Division of Hepatobiliary Surgery, Department of Surgery, Kaohsiung Medical University Hospital, Kaohsiung 80708, Taiwan; ktlee@kmu.edu.tw
5 Department of Healthcare Administration and Medical Informatics, Kaohsiung Medical University, Kaohsiung 80708, Taiwan
6 Department of Medical Research, Chi Mei Medical Center, Tainan 71004, Taiwan; 400002@mail.chimei.org.tw
7 Department of Emergency Medicine, Chi Mei Medical Center, Tainan 71004, Taiwan; honyi6165@yahoo.com
8 Department of Senior Services, Southern Taiwan University of Science and Technology, Tainan 71005, Taiwan
9 Department of Business Management, National Sun Yat-sen University, Kaohsiung 80424, Taiwan
10 Department of Medical Research, Kaohsiung Medical University Hospital, Kaohsiung 80708, Taiwan
* Correspondence: hshi@kmu.edu.tw; Tel.: +886-7-3121101 (ext. 2648)

Received: 23 February 2019; Accepted: 22 April 2019; Published: 27 April 2019

Abstract: Despite the growing use of minimal clinically important difference (MCID) as a cancer outcome measure, no study has reported clinically significant outcomes in cancer patients. We defined MCID and evaluated the use of preoperative HRQoL for predicting MCID and survival after surgical resection of hepatocellular carcinoma (HCC). In total, 369 patients completed the Functional Assessment of Cancer Therapy-Hepatobiliary (FACT-Hep) and the SF-36 at baseline and at two years post-operative at three tertiary academic hospitals. The corresponding MCID values were 3.6 (SF-36 physical component summary), 4.2 (SF-36 mental component summary), 5.4 (FACT-General total score), and 6.7 (FACT-Hep total score). The predictors of achieving postoperative MCID were significantly higher in patients who had low preoperative HRQoL score, advanced age, high education level, and high BMI ($p < 0.05$). However, patients with a high preoperative HRQoL score, high education level, high BMI, and low Charlson comorbidity index score were significantly associated with survival ($p < 0.05$). Preoperative HRQoL scores were predictive of MCID and overall survival after surgical resection of HCC. The findings of this study may be useful for managing the preoperative expectations of candidates for HCC resection and for developing shared decision-making procedures for patients undergoing surgical resection of HCC.

Keywords: hepatocellular carcinoma; health-related quality of life; minimal clinically important difference; survival

1. Introduction

Hepatic resection is the mainstay curative treatment for patients with hepatocellular carcinoma (HCC), even in some patients with early-stage HCC [1–3]. Health-related quality of life (HRQoL) is a recognized indicator of healthcare outcomes and, since the 1990s, evaluations of cancer treatment

outcomes have increasingly emphasized assessment of HRQoL [4,5]. Disease-specific and generic HRQoL measures are often reported together and provide complementary assessments of patient well-being before and after an intervention. It is important not to mix up the concept of quality of life with a recently growing area of HRQoL. Quality of life is an essential concept in the field of international development since it allows analysis of development on a measure broader than the standard of living [6,7]. Within development theory, however, there are varying ideas concerning what constitutes desirable change for a particular society, and the different ways that institutions define the quality of life, therefore, shapes how these organizations work for its improvement as a whole. The Functional Assessment of Cancer Therapy-Hepatobiliary (FACT-Hep) measure is one of the most widely used patient-reported questionnaires for measuring HRQoL in cancer research [8]. Unfortunately, research has shown that most of the studies that have used the FACT-Hep lack any reporting of clinical significance, even though guidelines for assessing clinical significance do exist.

A clinically important difference is a change that a patient or clinician would consider meaningful or worthwhile, such that an intervention or treatment would be considered worthy of repeating or such that patients would consider the change as an improvement in HRQoL. As such, measures of clinical significance such as the minimal clinically important difference (MCID) are increasingly used as a standard clinical outcome measure [6,7]. The MCID is defined as the smallest outcome change that the patient perceives as clinically important [9]. Despite the growing use of MCID as a cancer outcome measure, no study has reported clinically significant outcomes after surgical resection of HCC.

For cancer patients and their families, clinical data for HRQoL outcomes provide a useful indicator of the expected course of recovery and the expected effects of treatment. Thus, HRQoL data can help them make informed treatment decisions [10,11]. Baseline assessments of HRQoL have proven useful for predicting survival in various cancers, including colorectal, esophageal, breast, oropharyngeal, and lung cancers [10–13]. For varying severity of cancer, HRQoL has shown higher sensitivity compared to conventional prognostic indicators and compared to physician assessments [11,14]. Gotay et al. assessed the use of patient-reported HRQoL as a prognostic indicator of cancer outcomes [15]. Out of 39 clinical trials reviewed by the authors, 36 reported at least one HRQoL domain that was a significant predictor of survival. However, comparisons of results published in the literature are difficult because studies differ in the HRQoL measures applied and studied populations differ in patient attributes such as the type, site, and stage of disease [14–17]. Given the variability in reported overall survival rates and overall study heterogeneity, the evidence base for overall survival and determinants of overall survival after surgical resection of HCC are still evolving, and continued investigation is warranted.

A growing body of evidence indicates that preoperative functional status or HRQoL are important determinants of cancer surgery outcomes [14–17]. To the best of our knowledge, no prior study has systematically evaluated the role of preoperative HRQoL in achieving MCID and in overall survival after surgical resection of HCC. Therefore, the purpose of this study was to investigate the use of preoperative HRQoL scores for predicting achievement of MCID and for predicting overall survival after surgical resection of HCC.

2. Materials and Methods

2.1. Subjects and Data Collection

This study recruited all patients who had received surgical resection of HCC performed at one of three southern Taiwan medical centers between February 2013 and February 2017. For accurate assessment of postoperative outcome measures, the analysis was limited to patients who had received surgical resection performed by a director of surgery in a medical institution or by a senior attending doctor specializing in HCC surgery or treatment. Inclusion criteria were the following: (1) a histologic or combined radiographic and laboratory diagnosis of HCC, (2) ability to communicate in Chinese or Taiwanese, and (3) agreement to participate in a questionnaire survey performed in the hospital ward or by telephone. Major exclusion criteria included concurrent malignancy or participation in another

quality-of-life study that might have interfered with this study. Figure 1 shows that, during the sample selection period, 496 subjects were eligible for participation. Of these, 62 were excluded due to benign tumor or cognitive impairment. Therefore, 369 subjects were assessed preoperatively (baseline) and at 2 years postoperatively. Baseline demographic and clinical data were collected through questionnaire surveys and medical records reviews. This study was approved by the Institutional Review Board of Chi Mei Medical Center (10002-L01).

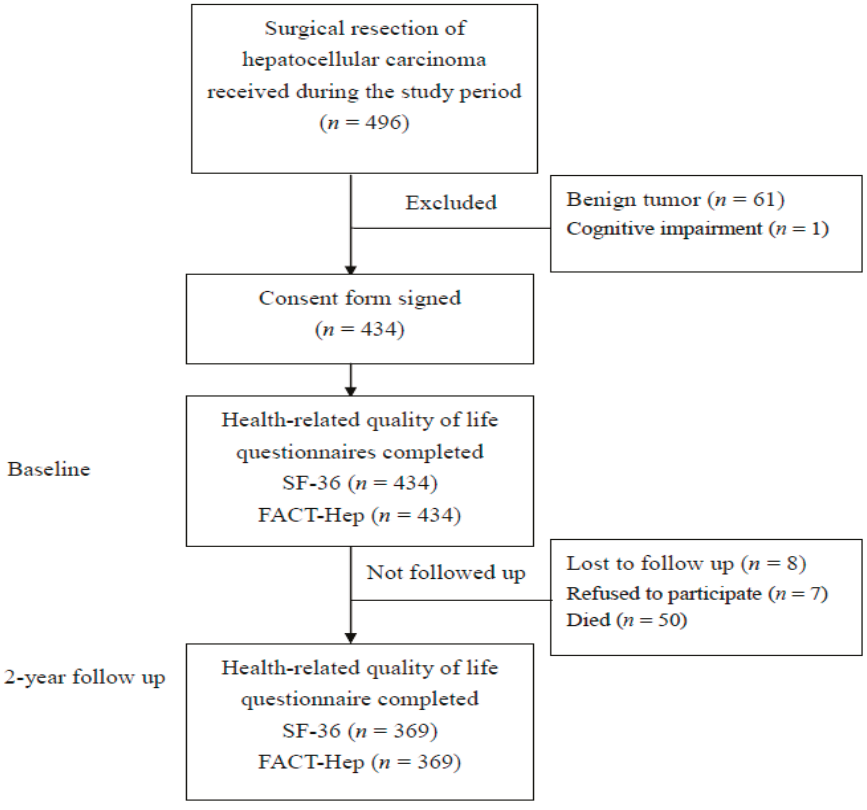


Figure 1. Flow chart showing population changes during the study, including subjects who met initial exclusion criteria, those who later declined to participate and those who lost to follow-up. SF-36: 36-Item Short Form Survey; FACT-Hep: Functional Assessment of Cancer Therapy-Hepatobiliary.

2.2. Study Protocol

Patients were asked to complete the questionnaires during follow-up visits at our outpatient clinic. To maximize compliance and minimize volunteer bias, a research assistant was available to help patients complete the questionnaires during each outpatient session. All HRQoL data were collected by the same research assistant. Patients were informed that their questionnaire responses would not be revealed to their attending surgeons and, hence, would not affect their treatment.

2.3. Measures of HRQoL

The Short Form-36 (SF-36) Health Survey measures eight dimensions: physical function, role limitation due to physical health, bodily pain, general health, vitality, social function, role limitation due to emotional health, and mental health. To compare the overall physical and mental functioning

of the study population with those in the general Taiwan population, physical component summary scores (PCS) and mental component summary scores (MCS) were calculated by norm-based scoring methods and used as dependent variables [18]. Based on a previous study [19], the PCS and MCS were computed in comparison with the general population of Taiwan. Values below 50 indicated that the examined PCS or MCS were below the average values for the general Taiwan population, and vice versa.

The FACT-Hep measure contains five dimensions: physical well-being, social/family well-being, functional well-being, emotional well-being, and additional concerns. The subscales for the physical well-being, social/family well-being, and functional well-being dimensions each contained seven items with a subscale score range of 0–28 points; the subscale for emotional well-being contained six items with a subscale score range of 0–24 points; the subscale for additional concerns about HCC contained 18 items with a subscale score range of 0–72 points [8]. The Functional Assessment of Cancer Therapy-General (Fact-G) and additional concerns for HCC scores were summed to obtain the FACT-Hep total score, which ranged from 0 to 180. Higher scores on all FACT-Hep dimensions were interpreted as better HRQoL and fewer symptoms.

2.4. Statistical Analysis

Studies show that a distribution-based method can reliably derive MCID calculated as one-half the standard deviation (SD) in outcome score change from baseline to the two-year follow up for a given instrument in a patient cohort [9,20]. Therefore, this methodology was used to determine MCID values for the SF-36 PCS, SF-36 MCS, FACT-G total score, and FACT-Hep total score.

Multivariable logistic regression models were used to identify predictors of the achievement of MCID after surgical resection of HCC. A Cox multivariable proportional hazard regression model was also used to evaluate how other prognostic factors affect survival. Survival distribution was estimated by Kaplan–Meier method. Significant differences in survival probability were stratified by a log-rank test. Hazard ratios (HRs) and 95% confidence intervals (CIs) were estimated from regression coefficients.

Variables included in the multivariable analyses were gender, age, marital status, education, body mass index (BMI), Charlson co-morbidity index (CCI) score, co-residence with family, smoking, drinking, tumor stage, chemotherapy, radiotherapy, and average length of stay (ALOS). Multivariable analyses also included preoperative HRQoL score. Variables that fell out of the model were excluded from the tables of multivariable results. Statistical analyses were performed using SPSS software (IBM SPSS Statistics for Windows, Version 20.0, Armonk, NY, USA). All statistical tests were two tailed with a significance level of 0.05.

3. Results

3.1. Patient Demographics

The SF-36 and FACT-Hep measures were completed by 369 HCC surgery patients preoperatively and at two years postoperatively. We compared the patients who remained in the study throughout the two-year period with those who were lost or dead to follow up between the baseline and the second year after discharge. There was no difference in terms of gender, age, marital status, education, BMI, CCI score, co-residence with family, smoking, drinking, tumor stage, chemotherapy, radiotherapy, ALOS, or in any preoperative HRQoL parameters mentioned above (data not shown). Table 1 presents their demographic and clinical characteristics.

Table 1. Demographic and clinical characteristics of 369 patients with hepatic resection for hepatocellular carcinoma.

Variable		N (%) or Mean ± SD
Gender	Male	271 (73.4)
	Female	98 (26.6)
Age, years		60.2 ± 10.8
Marital status	Married	335 (90.8)
	Divorced or widowed	34 (9.2)
Education		8.7 ± 3.6
	No formal education	26 (7.1)
	Primary school	122 (33.1)
	Junior high school	75 (20.3)
	Senior high school	92 (24.9)
	College or above	54 (14.6)
Body mass index, kg/m^2		25.0 ± 3.5
	Normal (18.5~24.9 kg/m^2)	218 (59.1)
	Overweight (25.0~29.9 kg/m^2)	124 (33.6)
	Obese (≥30.0 kg/m^2)	27 (7.3)
Charlson co-morbidity index, score		1.6 ± 1.3
Co-residence with family	Yes	358 (97.0)
	No	11 (3.0)
Smoking	Yes	71 (19.2)
	No	298 (80.8)
Drinking	Yes	78 (21.1)
	No	291 (78.8)
Tumor stage	I	216 (58.6)
	II	102 (27.6)
	III	51 (13.8)
Chemotherapy	Yes	11 (3.0)
	No	358 (97.0)
Radiotherapy	Yes	5 (1.4)
	No	364 (98.6)
Average length of stay, days		13.0 ± 6.6

SD: standard deviation.

3.2. HRQoL Outcomes

The patients had a mean age of 60.2 ± 10.8 years, and 73.4% (271) patients were male. Table 2 shows that mean patient-reported HRQoL scores at two years after surgery were significantly higher than those before surgery ($p < 0.001$). The MCID values were 3.6 for the SF-36 PCS; 4.2 for the SF-36 MCS; 5.4 for the FACT-G total score; and 6.7 for the FACT-Hep total score.

Table 2. Mean ± standard deviation for SF-36 and Functional Assessment of Cancer Therapy-Hepatobiliary (FACT-Hep) before and after resection for hepatocellular carcinoma ($n = 369$) *.

Variable	Preoperative	2 Years Postoperative	2 Years Postoperative-Preoperative	*p* Value
SF-36 PCS	56.0 ± 8.7	61.7 ± 9.5	5.8 ± 7.2	$p < 0.001$
SF-36 MCS	48.5 ± 8.1	57.1 ± 9.8	8.5 ± 8.4	$p < 0.001$
FACT-G total	91.2 ± 10.4	98.7 ± 10.8	7.5 ± 10.9	$p < 0.001$
FACT-Hep total	156.9 ± 14.2	165.6 ± 15.8	9.7 ± 13.3	$p < 0.001$

* Both PCS and MCS scores were converted to obtain a mean of 50 and a standard deviation of 10 compared to the normal (nationwide) group. SF-36, 36-Item Short Form Survey; PCS, physical component summary; MCS, mental component summary; FACT-G, Functional Assessment of Cancer Therapy-General; FACT-Hep, Functional Assessment of Cancer Therapy-Hepatobiliary.

3.3. Multivariable Analyses

Multivariable analyses of HRQoL and survival were performed to identify predictors of the achievement of MCID after surgical resection for HCC. For each HRQoL measure, a high preoperative score negatively predicted achievement of MCID ($p < 0.001$) (Table 3). According to the SF-36 PCS data, the odds of achieving MCID were lower in males than in females (odds ratio (OR), 0.31; 95% CI, 0.12–0.83) but were higher in patients with advanced age (OR, 1.05; 95% CI, 1.01, 1.10), high education level (OR, 1.14; 95% CI, 1.01, 1.30), and high BMI (OR, 1.11; 95% CI, 1.09, 1.12) compared to their counterparts with young age, low education level, and low BMI, respectively. According to the SF-12 MCS data, the odds of achieving MCID were lower in patients with high BMI compared to those with low BMI (OR, 0.91; 95% CI, 0.84, 0.99); however, the odds of achieving MCID were higher in those with advanced age (OR, 1.01; 95% CI, 1.01, 1.02) and high CCI score (OR, 1.53; 95% CI, 1.13, 1.94) compared to their counterparts with young age and low CCI score, respectively. According to the FACT-G total data, the odds of achieving MCID were higher in patients with advanced age (OR, 1.04; 95% CI, 1.01, 1.07), high education level (OR, 1.12; 95% CI, 1.04, 1.21), and high BMI (OR, 1.19; 95% CI, 1.09, 1.29) compared to their counterparts with young age, low education level, and low BMI, respectively. According to the FACT-Hep total data, the odds of achieving MCID were higher in patients with advanced age (OR, 1.05; 95% CI, 1.02, 1.07), high education level (OR, 1.11; 95% CI, 1.03, 1.19), and high BMI (OR, 1.09; 95% CI, 1.01, 1.18) compared to their counterparts with young age, low education level, and low BMI.

Table 3. Odds of achieving minimal clinical important difference (MCID) in health-related quality of life according to multivariate logistic regression model *.

Variables	Odds Ratio (95% CI)	*p* Value
SF-36 PCS		
Preoperative SF-36 PCS score	0.90 (0.84, 0.96)	<0.001
Age	1.05 (1.01, 1.10)	0.041
Gender (male vs. female)	0.31 (0.12, 0.83)	0.019
Education	1.14 (1.01, 1.30)	0.040
Body mass index	1.11 (1.09, 1.12)	0.045
SF-36 MCS		
Preoperative SF-36 MCS score	0.80 (0.73, 0.88)	<0.001
Age	1.01 (1.01, 1.02)	<0.001
Body mass index	0.91 (0.84, 0.99)	0.034
Charlson co-morbidity index	1.53 (1.13, 1.94)	<0.001
FACT-G total		
Preoperative FACT-G total score	0.92 (0.90, 0.95)	<0.001
Age	1.04 (1.01, 1.07)	0.007
Education	1.12 (1.04, 1.21)	0.002
Body mass index	1.19 (1.09, 1.29)	<0.001
FACT-Hep total		
Preoperative FACT-Hep total score	0.97 (0.95, 0.98)	<0.001
Age	1.05 (1.02, 1.07)	0.001
Education	1.11 (1.03, 1.19)	0.007
Body mass index	1.09 (1.01, 1.18)	0.020

* The full model was adjusted for preoperative functional status, gender, age, marital status, education, body mass index, Charlson co-morbidity index, co-residence with family, smoking, drinking, tumor stage, chemotherapy, radiotherapy, and average length of stay. SF-36, 36-Item Short Form Survey; PCS, physical component summary; MCS, mental component summary; FACT-G, Functional Assessment of Cancer Therapy-General; FACT-Hep, Functional Assessment of Cancer Therapy-Hepatobiliary.

Multivariate analyses of each pre-operative HRQoL score showed that a high score was a positive predictor of overall survival (Table 4). Lower pre-operative HRQoL scores were significantly associated with post-operative morbidity ($p < 0.05$). Education level, BMI, and CCI also showed significant associations with overall survival ($p < 0.001$).

Table 4. Overall survival analysis by Cox multivariable proportional hazard regression model *.

Variable	HR (95% CI)	*p* Value
Preoperative SF-36 PCS score	1.05 (1.03, 1.08)	<0.001
Preoperative SF-36 MCS score	1.03 (1.01, 1.05)	<0.001
Preoperative SF-36 physical function	1.06 (1.01, 1.10)	<0.001
Preoperative SF-36 role physical	1.03 (1.00, 1.05)	<0.001
Preoperative SF-36 bodily pain	1.02 (1.01, 1.03)	0.008
Preoperative SF-36 general health	1.07 (1.01, 1.14)	<0.001
Preoperative SF-36 vitality	1.02 (1.01, 1.04)	0.001
Preoperative SF-36 social function	1.02 (1.01, 1.03)	0.003
Preoperative SF-36 role emotional	1.04 (1.01, 1.06)	<0.001
Preoperative SF-36 mental health	1.03 (1.00, 1.05)	<0.001
Preoperative FACT physical well-being	1.04 (1.00, 1.07)	<0.001
Preoperative FACT social/family well-being	1.01 (1.01, 1.02)	0.010
Preoperative FACT functional well-being	1.03 (1.01, 1.06)	<0.001
Preoperative FACT emotional well-being	1.03 (1.01, 1.05)	<0.001
Preoperative FACT additional concerns	1.02 (1.00, 1.04)	<0.001
Preoperative FACT-G total score	1.07 (1.01, 1.14)	<0.001
Preoperative FACT-Hep total score	1.10 (1.02, 1.19)	<0.001
Education	1.10 (1.02, 1.18)	0.012
Body mass index	1.02 (1.01, 1.04)	0.002
Charlson co-morbidity index	0.83 (0.70, 0.99)	0.040

* The full model was adjusted for preoperative functional status, gender, age, marital status, education, body mass index, Charlson co-morbidity index, co-residence with family, smoking, drinking, tumor stage, chemotherapy, radiotherapy, and average length of stay. SF-36, 36-Item Short Form Survey; PCS, physical component summary; MCS, mental component summary; FACT-G, Functional Assessment of Cancer Therapy-General; FACT-Hep, Functional Assessment of Cancer Therapy-Hepatobiliary; HR, hazard ratio; CI, confidence interval.

4. Discussion

This study investigated how patient-reported preoperative HRQoL affects two outcomes of surgical resection of HCC: MCID and overall survival. Multivariate analyses showed that each preoperative HRQoL score was predictive of both MCID and overall survival. Low preoperative HRQoL score, advanced age, high education level, and high BMI were significantly associated with achievement of postoperative MCID ($p < 0.05$). Additionally, high preoperative HRQoL score, high education level, high BMI, and low CCI had significant positive associations overall survival ($p < 0.05$). It demonstrates that, at baseline, preoperative HRQoL scores relates to postoperative mortality. Lower scores in physical and functional domains are associated with an increased risk of postoperative mortality. The importance of preoperative HRQoL scores for predicting outcomes of surgical resection in HCC patients is now well recognized [9,21]. The current study found that, for a given HRQoL outcome measure, a high preoperative score was significantly for not achieving a postoperative MCID in the outcome measure. The likely explanation for this finding is that patients who already have high HRQoL scores before surgery and less potential for achieving a HRQoL score improvement that meets the criteria for an MCID.

This study aimed to calculate and report the MCID value of commonly used HRQoL scales. Changes in HRQoL by time and/or treatment may not correlate with the direction (positive vs. negative) as well as the magnitude of clinical improvements in outcomes as they are perceived by the patients. Furthermore, cancer treatment has a more significant impact on HRQoL among HCC surgical patients. The MCID for (SF-36 PCS, SF-36 MCS, and FACT-G total score) value changes differed across domains, and they differed for perceived improvement and deterioration. We knew that domain scores related to physical function diminished from pre-treatment to on- or immediately after treatment, and emotional

function improved. Additionally, the initial anxiety of the diagnosis and treatment initiation period may have been ameliorated by subsequent familiarity and supportive psychosocial care provided by the clinical service teams after patients finished the initial questionnaire. Thus, this might explain why those patients with a low preoperative score did achieve a postoperative MCID in the outcome measure and regarded as "efficacy" when compared with that achieved by those patients who had full pre-operative familiarity and mental support and thus noted with a high preoperative score.

The significant associations revealed by the HRQoL instruments investigated in this study underscore the relationship between HRQoL measures and medical outcomes. The study showed that preoperative HRQoL scores accurately predict postoperative MCID and overall survival, which is consistent with the literature [11,15,17]. Therefore, counseling is essential for apprising HCC resection candidates of expected postoperative improvements and impairments. If medical outcomes are considered benchmarks, then the preoperative HRQoL score, which is an important predictor of postoperative MCID and overall survival, is crucial.

Until now, no studies have described the implications of significant changes in disease-specific and generic HRQoL outcome measures in patients who have undergone HCC surgery. Steel et al. evaluated the clinical meaningfulness of FACT-Hep scores in HCC patients [22]. The authors combined distribution-based analyses with cross-sectional anchor-based analyses to obtain minimally important differences (MIDs) in FACT-G subscale scores (2–3 points), FACT-G total scores (6–7 points), Hepatobiliary Cancer Subscale scores (MID 5–6 points), and FACT-Hep scores (MID 8–9 points). However, data for clinically significant improvements in HCC surgery outcomes, particularly patient-reported outcomes, are very limited [4–6]. For patients who undergo surgical resection of HCC, the current study obtained MCID values of 3.6 for SF-36 PCS; 4.2 for SF-36 MCS; 5.4 for FACT-G total score; and 6.7 for FACT-Hep total score. These data are a novel addition to the literature and provide a useful reference for further studies of HRQoL outcomes after surgical resection of HCC. Notably, this study used patient-reported data for the period from before surgery until two years after surgery. A distribution-based method was used to calculate MCID. Since MCID values may change depending on the time point studied and the psychometric method used for analysis, further studies are needed to investigate MCID after surgical resection of HCC using different time points and mixed anchor/distribution-based methods of deriving MCID values.

The evidence base for overall survival after surgical resection of HCC is growing but is still relatively limited. The overall survival rates reported in the literature are somewhat variable, and the studied populations have been heterogeneous. Quinten et al. investigated the prognostic relationship between HRQoL and survival in a dataset for 30 randomized controlled trials performed by the European Organization for Research and Treatment of Cancer [11]. Their study showed that, in each cancer site, at least one HRQoL domain had an additive prognostic value that exceeded the prognostic values of clinical and sociodemographic variables. A systematic literature review by Quinten et al. confirmed that baseline HRQoL and at least one HRQoL domain were significantly associated with overall survival [11]. The current study found that preoperative HRQoL scores, education level, and BMI had significant positive associations with overall survival ($p < 0.05$) whereas CCI score had a significant negative association with overall survival ($p < 0.05$).

The HRQoL factors identified in this study varied from those in previous studies [9,20,22]. One possible explanation is differences in study populations. Our study focused on patients in both early and advanced stages of HCC whereas previous studies have only focused on patients in advanced stages of the disease. Secondly, patients with different cultural backgrounds may have different perceptions of HRQoL. Thirdly, patients in recent studies have more treatment options compared to patients in earlier studies, which can result in different perceptions of the implications of HCC and thus different perceptions of HRQoL. Fourthly, even studies that use the same HRQoL measure may have very different data analysis methodologies.

The findings of this study have important implications for preoperative counseling of patients, management of patient expectations, and stratification of outcomes. Healthcare providers increasingly

emphasize shared decision making and are now using predictive modeling to inform surgery patients about potential clinical outcomes and the likelihood of success [11,22]. The results of the present study suggest that healthcare providers should consider routinely administering HRQoL instruments preoperatively as screening tools and for informing shared decision-making strategies. Patients with low HRQoL scores can be referred for counseling to modify their outcome expectations or referred for targeted interventions to optimize their HCC resection outcomes.

Certain limitations of this study are noted. Firstly, the patient data were derived from a multi-institutional HCC registry containing data contributed by multiple surgeons. As such, other than institutional best practices, surgical techniques and rehabilitation protocols were not standardized. However, since patient-reported HRQoL outcomes were obtained at the time of each clinical encounter, the reports of functional status are assumedly accurate. Another limitation is that sensitivity of the MCID values was not analyzed. The MCID can be calculated according to a consensus of or by using an anchor-based method or a distribution-based method [9,23]. Each methodology for deriving MCID has its associated pitfalls, and none has consistently proven to be superior. The applied methodology should be selected according to the characteristics of the data and the disease under study [23]. Additionally, the role of the FACT-Hep for assessing outcomes after surgical resection of HCC has not been robustly studied or validated. As such, the responsiveness of the HRQoL outcomes for this population subset is not clear; for example, the reliability (and change over time) of emotional, psychological, and social responses after surgical resection of HCC needs further study.

5. Conclusions

In conclusion, this study revealed that HRQoL scores are independent predictors of MCID and overall survival after surgical resection of HCC. The MCID values were 3.6 for the SF-36 PCS; 4.2 for the SF-36 MCS; 5.4 for the FACT-G total score; and 6.7 for the FACT-Hep total score. The findings of this study may be useful for preoperative management of patient expectations and for developing shared decision-making measures for patients undergoing surgical resection of HCC.

Author Contributions: Study concept and design: C.-C.C., and H.-Y.S.; acquisition of data: K.-T.L., J.-J.W., D.-P.S., H.-H.L., and C.-C.H.; analysis and interpretation of data: C.-C.C., and H.-Y.S.; drafting of the manuscript: C.-C.C., and H.-Y.S.; revised the manuscript: K.-T.L., J.-J.W., D.-P.S., H.-H.L., and C.-C.H.; approved the manuscript: all authors.

Funding: This study was supported by grants from the Chi Mei Medical Center, Liouying, Research Grant (CLFHR 10301 and 10409) and from the Chi-Mei Medical Center and the Kaohsiung Medical University Research Foundation (107CM-KMU).

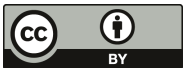
Acknowledgments: We thank all the individuals who participated in this study, without which this study would not have been possible.

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

References

1. Dhir, M.; Melin, A.A.; Douaiher, J.; Lin, C.; Zhen, W.K.; Hussain, S.M.; Geschwind, J.F.; Doyle, M.B.; Abou-Alfa, G.K.; Are, C. A Review and update of treatment options and controversies in the management of hepatocellular carcinoma. *Ann. Surg.* **2016**, *263*, 1112–1125. [CrossRef]
2. Robertson, F.P.; Fuller, B.J.; Davidson, B.R. An evaluation of ischaemic preconditioning as a method of reducing ischaemia reperfusion injury in liver surgery and transplantation. *J. Clin. Med.* **2017**, *6*, 69. [CrossRef] [PubMed]
3. Glantzounis, G.K.; Paliouras, A.; Stylianidi, M.C.; Milionis, H.; Tzimas, P.; Roukos, D.; Pentheroudakis, G.; Felekouras, E. The role of liver resection in the management of intermediate and advanced stage hepatocellular carcinoma. A systematic review. *Eur. J. Surg. Oncol.* **2018**, *44*, 195–208. [CrossRef]
4. Nishikawa, H.; Enomoto, H.; Yoh, K.; Iwata, Y.; Sakai, Y.; Kishino, K.; Ikeda, N.; Takashima, T.; Aizawa, N.; Takata, R.; et al. Health-related quality of life after treatment for bladder cancer in England. *J. Clin. Med.* **2018**, *7*, 553. [CrossRef] [PubMed]

5. Chiu, C.C.; Lee, K.T.; Wang, J.J.; Sun, D.P.; Lee, H.H.; Shi, H.Y. Health-related quality of life before and after surgical resection of hepatocellular carcinoma: A prospective study. *Asian Pac. J. Cancer Prev.* **2018**, *19*, 65–72.
6. Shapiro, C.L. Highlights of recent findings on quality-of-life management for patients with cancer and their survivors. *JAMA Oncol.* **2016**, *2*, 1401–1402. [CrossRef]
7. Costa, D.S. Reflective, causal, and composite indicators of quality of life: A conceptual or an empirical distinction? *Qual. Life Res.* **2015**, *24*, 2057–2065. [CrossRef]
8. Heffernan, N.; Cella, D.; Webster, K.; Odom, L.; Martone, M.; Passik, S.; Bookbinder, M.; Fong, Y.; Jarnagin, W.; Blumgart, L. Measuring health-related quality of life in patients with hepatobiliary cancers: The functional assessment of cancer therapy-hepatobiliary questionnaire. *J. Clin. Oncol.* **2002**, *20*, 2229–2239. [CrossRef] [PubMed]
9. Jayadevappa, R.; Cook, R.; Chhatre, S. Minimal important difference to infer changes in health-related quality of life-a systematic review. *Clin. Epidemiol.* **2017**, *89*, 188–198. [CrossRef]
10. Quinten, C.; Martinelli, F.; Coens, C.; Sprangers, M.A.; Ringash, J.; Gotay, C.; Bjordal, K.; Greimel, E.; Reeve, B.B.; Maringwa, J.; et al. A global analysis of multitrial data investigating quality of life and symptoms as prognostic factors for survival in different tumor sites. *Cancer* **2014**, *120*, 302–311. [CrossRef]
11. Quinten, C.; Coens, C.; Mauer, M.; Comte, S.; Sprangers, M.A.; Cleeland, C.; Osoba, D.; Bjordal, K.; Bottomley, A.; EORTC Clinical Groups. Baseline quality of life as a prognostic indicator of survival: A meta-analysis of individual patient data from EORTC clinical trials. *Lancet Oncol.* **2009**, *10*, 865–871. [CrossRef]
12. Grande, G.E.; Farquhar, M.C.; Barclay, S.I.; Todd, C.J. Quality of life measures (EORTC QLQ-C30 and SF-36) as predictors of survival in palliative colorectal and lung cancer patients. *Palliat. Support. Care* **2009**, *7*, 289–297. [CrossRef] [PubMed]
13. Dunne, D.F.; Jack, S.; Jones, R.P.; Jones, L.; Lythgoe, D.T.; Malik, H.Z.; Poston, G.J.; Palmer, D.H.; Fenwick, S.W. Randomized clinical trial of prehabilitation before planned liver resection. *Br. J. Surg.* **2016**, *103*, 504–512. [CrossRef] [PubMed]
14. Quinten, C.; Maringwa, J.; Gotay, C.C.; Martinelli, F.; Coens, C.; Reeve, B.B.; Flechtner, H.; Greimel, E.; King, M.; Osoba, D.; et al. Patient self-reports of symptoms and clinician ratings as predictors of overall cancer survival. *J. Natl. Cancer Inst.* **2011**, *103*, 1851–1858. [CrossRef] [PubMed]
15. Robinson, D.W.J.; Eisenberg, D.F.; Cella, D.; Zhao, N.; de Boer, C.; DeWitte, M. The prognostic significance of patient-reported outcomes in cancer clinical trials. *J. Clin. Oncol.* **2008**, *26*, 1355–1363.
16. Lin, S.; Hoffmann, K.; Schemmer, P. Treatment of hepatocellular carcinoma: A systematic review. *Liver Cancer* **2012**, *1*, 144–158. [CrossRef] [PubMed]
17. Pompili, C.; Salati, M.; Refai, M.; Berardi, R.; Onofri, A.; Mazzanti, P.; Brunelli, A. Preoperative quality of life predicts survival following pulmonary resection in stage I non-small-cell lung cancer. *Eur. J. Cardiothorac. Surg.* **2013**, *43*, 905–910. [CrossRef]
18. Ware, J.E.J.; Sherbourne, C.D. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med. Care* **1992**, *30*, 473–483. [CrossRef]
19. Huang, I.C.; Wu, A.W.; Frangakis, C. Do the SF-36 and WHOQOL-BREF measure the same constructs? Evidence from the Taiwan population. *Qual. Life Res.* **2006**, *5*, 15–24. [CrossRef]
20. Blondel, B.; Schwab, F.; Ungar, B.; Smith, J.; Bridwell, K.; Glassman, S.; Shaffrey, C.; Farcy, J.P.; Lafage, V. Impact of magnitude and percentage of global sagittal plane correction on health-related quality of life at 2-years follow-up. *Neurosurgery* **2012**, *71*, 341–348. [CrossRef] [PubMed]
21. Doll, K.M.; Snavely, A.C.; Kalinowski, A.; Irwin, D.E.; Bensen, J.T.; Bae-Jump, V.; Boggess, J.F.; Soper, J.T.; Brewster, W.R.; Gehrig, P.A. Preoperative quality of life and surgical outcomes in gynecologic oncology patients: A new predictor of operative risk? *Gynecol. Oncol.* **2014**, *133*, 546–551. [CrossRef] [PubMed]
22. Steel, J.L.; Eton, D.T.; Cella, D.; Olek, M.C.; Carr, B.I. Clinically meaningful changes in health-related quality of life in patients diagnosed with hepatobiliary carcinoma. *Ann. Oncol.* **2006**, *17*, 304–312. [CrossRef] [PubMed]
23. McGlothlin, A.E.; Lewis, R.J. Minimal clinically important difference: Defining what really matters to patients. *JAMA* **2014**, *312*, 1342–1343. [CrossRef] [PubMed]



Journal of
Clinical Medicine

MDPI

Article

Re-Evaluating the Protective Effect of Hemodialysis Catheter Locking Solutions in Hemodialysis Patients

Chang-Hua Chen [1,2,3,4,*], Yu-Min Chen [5], Yu Yang [6], Yu-Jun Chang [7], Li-Jhen Lin [2] and Hua-Cheng Yen [8]

1 Division of Infectious Disease, Department of Internal Medicine, Changhua Christian Hospital, Changhua 500, Taiwan
2 Center for Infection Prevention and Control, Changhua Christian Hospital, Changhua 500, Taiwan; 3344@cch.org.tw
3 Program in Translational Medicine, National Chung Hsing University, Taichung County 402, Taiwan
4 Rong Hsing Research Center for Translational Medicine, National Chung Hsing University, Taichung County 402, Taiwan
5 Department of Pharmacy, Changhua Christian Hospital, Changhua 500, Taiwan; 30855@cch.org.tw
6 Division of Nephrology, Department of Internal Medicine, Changhua Christian Hospital, Changhua 500, Taiwan; 2219@cch.org.tw
7 Epidemiology and Biostatistics Center, Changhua Christian Hospital, Changhua 500, Taiwan; 83686@cch.org.tw
8 Department of Neurosurgery, Changhua Christian Hospital, Changhua 500, Taiwan; 90211@cch.org.tw
* Correspondence: chenchanghuachad@gmail.com; Tel.: +886-4-723-8595-4080

Received: 5 March 2019; Accepted: 20 March 2019; Published: 25 March 2019

Abstract: Catheter-related bloodstream infections (CRBSIs) and exit-site infections (ESIs) are common complications associated with the use of central venous catheters for hemodialysis. The aim of this study was to analyze the impact of routine locking solutions on the incidence of CRBSI and ESI, in preserving catheter function, and on the rate of all-cause mortality in patients undergoing hemodialysis. We selected publications (from inception until July 2018) with studies comparing locking solutions for hemodialysis catheters used in patients undergoing hemodialysis. A total of 21 eligible studies were included, with a total of 4832 patients and 318,769 days of catheter use. The incidence of CRBSI and ESI was significantly lower in the treated group (citrate-based regimen) than in the controls (heparin-based regimen). No significant difference in preserving catheter function and all-cause mortality was found between the two groups. Our findings demonstrated that routine locking solutions for hemodialysis catheters effectively reduce the incidence of CRBSIs and ESIs, but our findings failed to show a benefit for preserving catheter function and mortality rates. Therefore, further studies are urgently needed to conclusively evaluate the impact of routine locking solutions on preserving catheter function and improving the rates of all-cause mortality.

Keywords: effect; protection; catheter; hemodialysis; meta-analysis; trial sequential analysis

1. Introduction

1.1. Variety of New Strategies for Locking Solutions to Avoid Catheter Infection and Catheter Malfunction in Hemodialysis Patients

Infections are widely prevalent in patients on chronic hemodialysis, and mortality from infection account for 10% of deaths observed in patients undergoing hemodialysis [1]. The use of central venous catheters in hemodialysis has been associated with catheter-related bloodstream infections (CRBSIs) and exit-site infections (ESIs) [2–4]. Although recent efforts have minimized the use of catheters, the proportion patients with end-stage renal disease undergoing dialysis using central

venous catheters has not yet declined [5]. Protective strategies against CRBSI and catheter malfunction are necessary [6], and to this end, the use of heparin as a routine locking solution for central venous catheters has become an accepted clinical practice [4]. However, heparinized locking solutions might cause unintended complications, such as systemic anticoagulation effects, bleeding episodes, heparin-induced thrombocytopenia, and susceptibility to bacterial biofilm formation [7–9]. A variety of new locking solutions have been developed; this includes citrate, which has antimicrobial properties [10–12]. However, the disadvantages of citrate compared with heparin have been raised and included the ability of avoiding catheter malfunction, citrate toxicity, and induction of cardiac arrhythmia [13]. Weijmer et al. showed that a 30% citrate solution was superior to heparin in preventing CRBSI [14]. In contrast, other studies have reported that the use of citrate does not have an advantage over heparin in preventing CRBSI [4,12]. Currently, the findings of the studies comparing citrate with heparin locking solutions are inconclusive for protecting against CRBSI and ESI and preserving the catheter function. Clinicians question if locking solutions should be considered a modifiable risk factor for CRBSIs in patients undergoing hemodialysis. Furthermore, the recommended locking solution for the routine care of patients undergoing hemodialysis continue to remain questionable.

1.2. Rationale for Re-Evaluating the Protective Effect of Hemodialysis Catheter Locking Solutions in Hemodialysis Patients

Routine locking solutions for hemodialysis catheters are recommended with category II evidence according to the guideline by the Healthcare Infection Control Practices Advisory Committee in 2011 [15]; however, there are some limitations of the studies providing the current and update evidence. Mostly, conclusions of meta-analysis could be influenced by the heterogeneity between individual studies and insufficient information size. Quantification of the required information size [16] is important to ensure the reliability of the data. In addition, current meta-analyses lack information size calculation [17–24]. Additionally, the incidence of CRBSI is difficult to evaluate because of their subjectivity for case finding, lack of specificity, and high inter-observer variability. CRBSI is associated with high morbidity and mortality in patients undergoing hemodialysis [1], and the prevention of CRBSI and ESI is becoming increasingly essential. Given these limitations, we performed a meta-analysis and trial sequential analysis to assess the impact of routine locking solutions on the incidence of CRBSI and ESI, in preserving catheter function, and on the rate of all-cause mortality in patients undergoing hemodialysis. We grouped the eligible publications according to combination regimen, antimicrobial activity, and concentration of the locking solutions; thereafter, we grouped according to the study design to assess its potential effect on the reported outcomes.

2. Experimental Section

2.1. Search Strategy and Inclusion Criteria

The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the institutional review board of Changhua Christian Hospital (CCH IRB No. 180801). From the earliest record to July 2018, we searched PubMed, Scopus, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, ClinicalTrials.gov, Embase, and Web of Science databases for studies on locking solutions for central venous catheters used in hemodialysis of patients. Full search strategies for each database are available in the Appendix A. The reference lists of the eligible publications were manually reviewed for relevant studies. Articles published in languages other than English or those with no available full text were excluded.

We included all trials and studies that provided data on one or more of our target outcomes for both the treated group and control group: CRBSIs and ESIs. Two investigators (CHC and YMC) independently reviewed potential trials and studies for inclusion. Disagreements were resolved by consensus. We also tried to contact the corresponding authors of selected papers to provide clarifications and missing data where needed.

2.2. Definition of Study Outcomes

Based on the original studies, the treated group comprised of patients undergoing hemodialysis using citrate as the locking solution for central venous catheters; for the control group heparin was used as the locking solution (Table 1). The outcomes of the original studies were included in this meta-analysis. The primary outcomes included (1) CRBSI, defined as bacteremia caused by an intravenous catheter, and (2) ESI, defined as the development of a purulent redness around the exit site that did not result from residual stitches. The secondary outcomes included (1), the need to remove the catheter due to catheter malfunction; and (2) the need for thrombolytic treatments due to catheter malfunction; and (3) all-cause mortality at any timeframe. Incidence was presented as the number of episodes per catheter or per patient depending on the available data.

2.3. Data Extraction and Quality Assessment

Two reviewers examined all retrieved articles and extracted data using a pre-determined form, recording the name of the first author, year of publication, country where the study was conducted, study design (RCT or observational studies), demographic and disease characteristics of participants, number of participants enrolled, and quality assessment of each study. Each reviewer independently evaluated the quality of the eligible studies, using Jadad scoring [25] for the RCTs and the Newcastle-Ottawa quality assessment scale [26] for the comparative experimental studies.

2.4. Data Synthesis and Analysis

The outcomes were measured by determining the odds ratios (ORs). A random effects model was used to pool individual ORs. Analyses were performed with the Comprehensive Meta-Analysis software version 3.0 (Biostat, Englewood, NJ, USA). Between-trial heterogeneity was determined using I^2 tests; values > 50% were regarded as considerable heterogeneity [27]. Funnel plots and Egger's test were used to examine potential publication bias [27]. Statistical significance was defined as $p < 0.05$, except for the determination of publication bias where $p < 0.10$ was considered significant. This study was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Table S1) [28].

In trial sequential analyses, the inconsistence of heterogeneity (I^2) adjusted by determining the required information size. The required information size was calculated with an intervention effect of a 10% relative risk reduction, an overall 5% risk of a type I error, and a 20% risk of a type II error. All trial sequential analyses were performed using TSA version 0.9 Beta (www.ctu.dk/tsa/, Copenhagen Trial Unit, Copenhagen, Denmark).

Table 1. Summary of the retrieved trials investigating experimental group and control group.

Author, Year, Country, Reference	RCT	Total *N*	Treated (*N*)	Control (*N*)	QA
Buturovic et al., 1998, SI, [29]	No	30	4% CiT (20)	1666 U/mL HpR (10)	3 #
Dogra et al., 2002, AU, [30]	Yes	79	26.7 mg/mL GM + 1.04% CiT (42)	5000 U/mL HpR (37)	8 *
Betjes et al., 2004, NL, [31]	No	58	1.35% TRD +4% CiT (37)	5000 U/mL HpR (39)	3 #
Weijmer et al., 2005, NL, [14]	Yes	291	30% CiT (148)	5000 U/mL HpR (143)	8 *
Nori et al., 2006, USA, [32]	No	40	4 mg/mL GM + 3.13% CiT (41)	5000 U/mL HpR (21)	3 #
Lok et al., 2007, CA, [13]	No	250	4% CiT (129)	5000 U/mL HpR (121)	3 #
MacRae et al., 2008, CA, [12]	No	61	4% CiT (32)	5000 U/mL HpR (29)	3 #
Power et al., 2009, UK, [4]	Yes	232	46.7% CiT (132)	5000 U/mL HpR (100)	8 *
Solomon et al., 2010, UK, [33]	Yes	107	1.35% TRD + 4% CiT (53)	5000 U/mL HpR (54)	8 *
Filiopoulos et al., 2011, GR, [34]	Yes	117	1.35% TRD + 4% CiT (119)	5000 U/mL HpR (58)	8 *
Maki et al., 2011, USA, [6]	Yes	407	7.0% CiT + MMP (206)	5000 U/mL HpR (201)	8 *
Moran et al., 2012, USA, [8]	No	303	320 µg/mL GM + 4% CiT (155)	1000 U/mL HpR (148)	3 #
Chen et al., 2014, CH, [35]	Yes	72	10% NaCl (36)	3125 U/mL HpR (36)	8 *

Table 1. *Cont.*

Author, Year, Country, Reference	RCT	Total *N*	Treated (*N*)	Control (*N*)	QA
Souweine et al., 2015, FR, [19]	Yes	1460	60% *w*/*w* EtOH (730)	0.9% NaCl (730)	8 *
Moghaddas et al., 2015, IR, [18]	Yes	87	10 mg/mL TMP/SMX + 2500 U/mL HpR (46)	2500 U/mL HpR (41)	8 *
Kanaa et al., 2015, UK, [17]	Yes	115	4% EDTA (59)	5000 U/mL HpR (56)	8 *
Zwiech et al., 2016, PL, [21]	Yes	50	4% CiT (26)	5000 U/mL HpR (24)	8 *
Chu et al., 2016, AU, [20]	Yes	100	1000 U/mL HpR (52)	5000 U/mL HpR (48)	8 *
Correa Barcellos et al., 2017, BZ, [22]	Yes	464	30% CiT (231)	5000 U/mL HpR (233)	8 *
Sofroniadou et al., 2017, GR, [23]	Yes	103	70% *w*/*w* EtOH + UFH 2000 U/mL (52)	2000 U/mL HpR (51)	8 *
Winnicki et al., 2018, Au, [24]	No	406	1.35% TRD + 4% CiT + HpR (52)	4% CiT (54)	3 #

Abbreviations: AU, Australia; Au, Austria; BZ, Brazil; CA, Canada; CH, China; CiT, citrate; EDTA, tetra-sodium ethylenediaminetetraacetic acid; EtOH, ethanol; FR, France; GM, gentamicin; GR, Greece; HpR, heparin; IR, Iran; MMP, 0.15% methylene blue + 0.15% methylparaben + 0.015% propylparaben; NaCl, sodium chloride; N, number; NL, Netherlands; PL, Poland; QA, quality assessment; RCT, randomized controlled trial; SI, Slovenia; TMP/SMX, cotrimoxazole (=trimethoprim/sulfamethoxazole); TRD, taurolidine ; UFH, unfractionated heparin; UK, United Kingdom; US, United States. #, the study was evaluated using Jadad scale. *, the study was assessed using the Newcastle-Ottawa scale.

3. Results

3.1. Eligible Studies

The literature search yielded 458 potentially eligible articles. By screening the abstracts, we removed 350 irrelevant articles. The remaining 100 articles were assessed further by full-text reading, of which 79 were excluded (Figure 1). Thus, 21 selected articles comparing citrate with heparin lockings for central venous catheters used in hemodialysis were included in this meta-analysis [4,6,8,12–14,17–24,29–35].

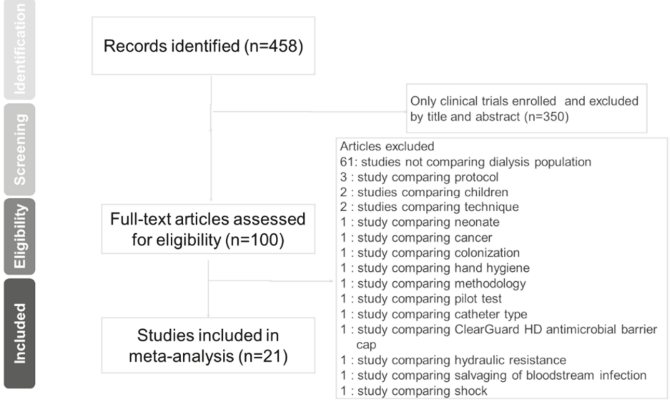


Figure 1. Preferred reporting items for systematic reviews and meta-analyses (PRISMA) flow diagram for the search and identification of the included studies.

The studies published in the selected articles were conducted from the earliest record to July 2018, with a total of 4832 patients and 318,769 total days of catheter use. Six studies compared citrate alone with heparin lockings; 14 studies tested regimens of citrate and other antimicrobials (gentamicin, taurolidine, methylparaben, methylene blue, and propylparaben) with heparin lockings; and two studies compared ethanol or combination solution (citrate, heparin and taurolidine) with non-heparin locking. Studies were conducted in North America (5 studies), South America (1), Europe (12), and Asia (3). A variety of end points were used in these studies. Most studies reported on CRBSI (17 studies [4,6,8,12–14,17,19,21–24,30–34]), followed by ESI (11 studies [4,8,12,14,17–19,24,30,31,33]),

catheter removal for poor flow (9 studies [6,8,12,14,18,24,29,31,33]), thrombolytic treatment (8 studies [4,8,14,17,18,32,33]), and mortality (5 studies [6,14,19,32,33]). The characteristics of the studies fulfilling the inclusion criteria are listed in Table 1. Thirteen studies were identified as RCT, and 6 studies were not double-blinded (Table 1).

3.2. Pooled Odds for Primary Outcomes and Subgroup Analysis

3.2.1. Catheter-Related Bloodstream Infection (CRBSI)

Seventeen studies (1731 patients; 217,128 catheter days) reported on CRBSI. The incidence of CRBSI was significantly lower in the treated group compared with the control group (OR, 0.424; 95% CI, 0.267–0.673; $p < 0.001$) (Figure 2). CRBSI subgroup analysis showed that the OR appeared to have a tendency to favor the treatment groups with either the combined regimen (OR, 0.206; 95% CI, 0.058–0.730; $p = 0.027$), the single regimen (OR, 0.289; 95% CI, 0.083–0.365; $p = 0.037$), a regimen containing antibiotics (OR, 0.136; 95% CI, 0.051–0.365; $p = 0.002$), or a low concentration of a major regimen (OR, 0.421; 95% CI, 0.186–0.956; $p = 0.039$; Table 2).

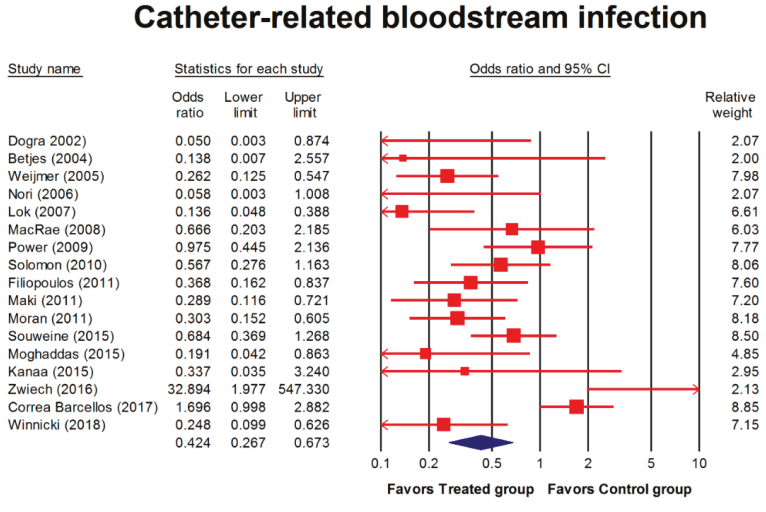


Figure 2. Forest plot of the overall odds ratios for catheter-related bloodstream infection in the treated group versus the control group. The random model of overall odds ratio showed a significant overall effect of interventions in reducing the risk for developing catheter-related bloodstream infections as compared with the control condition (OR, 0.424; 95% CI, 0.267–0.673; $p < 0.001$).

3.2.2. Exit-Site Infection (ESI)

A total of 11 RCTs (2,425 patients; 231,086 catheter days) described ESI. The incidence of ESI was significantly lower in the treated group compared with the control group (OR, 0.627; 95% CI, 0.441–0.893; $p = 0.001$; Figure 3). Further focusing at exit-site infection (Table 3), the subgroup analysis (combined regimen, regimen containing antibiotic, and concentration of regimen for exit-site infection) disclosed no significant differences between any groups except for combined regimen.

Table 2. Subgroup analysis of odds ratio based on study designs, combined regimen, regimen containing antibiotic, and concentration of regimen for CRBSI.

Subgroup	Odds Ratio	95% Confidence Interval
combined regimen		
RCT	0.606	0.298–1.230
Not RCT	0.206	0.058–0.730
Not combined regimen		
RCT	0.417	0.192–0.905
Not RCT	0.289	0.083–0.365
Regimen containing antibiotic		
RCT	0.191	0.023–1.564
Not RCT	0.136	0.051–0.365
Regimen Not containing antibiotic		
RCT	0.546	0.314–0.949
Not RCT	0.342	0.191–0.614
High Concentration of major regimen		
RCT	0.644	0.155–2.671
Low Concentration of major regimen		
RCT	0.421	0.186–0.956
Not RCT	0.260	0.135–0.497

Abbreviation: RCT, randomized controlled trial.

Table 3. Subgroup analysis of odds ratio based on study designs, combined regimen, regimen containing antibiotic, and concentration of regimen for exit site infection.

Subgroup	Odds Ratio	95% Confidence Interval
combined regimen		
RCT	0.849	0.358–2.011
Not RCT	0.706	0.307–1.62
Not combined regimen		
RCT	0.503	0.276–0.918
Not RCT	0.620	0.113–3.389
Regimen containing antibiotic		
RCT	0.571	0.189–1.725
Not RCT	0.735	0.284–1.905
Regimen Not containing antibiotic		
RCT	0.599	0.334–1.071
Not RCT	0.650	0.246–1.722
High Concentration of major regimen		
RCT	0.631	0.214–1.862
Low Concentration of major regimen		
RCT	0.805	0.282–2.297
Not RCT	0.692	0.35–1.368

Abbreviation: RCT, randomized controlled trial.

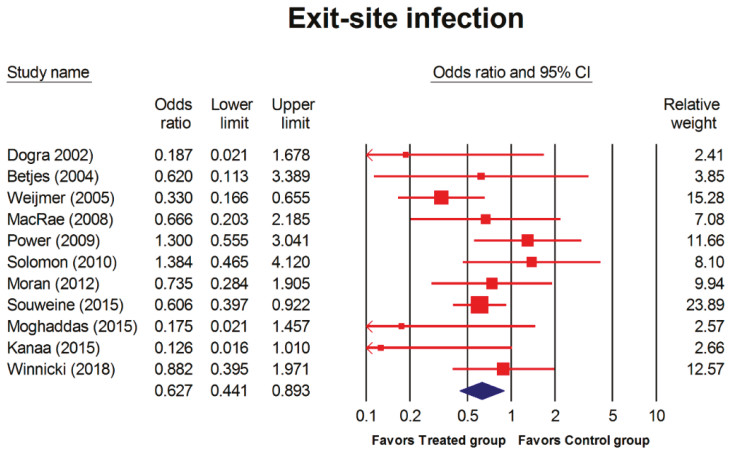


Figure 3. Forest plot of the overall odds ratios for exit-site infection in treated group versus the control group. The random model of overall odds ratio for exit-site infection showed a significant overall effect of interventions in reducing the risk for developing exit-site infection as compared with the control condition (OR, 0.627; 95% CI, 0.441–0.893; $p = 0.001$).

3.3. Pooled Odds for Secondary Outcomes and Subgroup Analysis

3.3.1. Catheter Withdrawal Due to Malfunction

Nine studies (1826 patients; 205,163 catheter days) reported catheters being removed for poor blood flow. As shown in Figure 4, no difference was identified between the two groups (OR, 0.696; 95% CI, 0.397–1.223; $p = 0.208$). Further subgroup analysis (combined regimen, regimen containing antibiotic, and concentration of regimen for catheter removal due to catheter malfunction) failed to reveal any differences between any groups (Table 4).

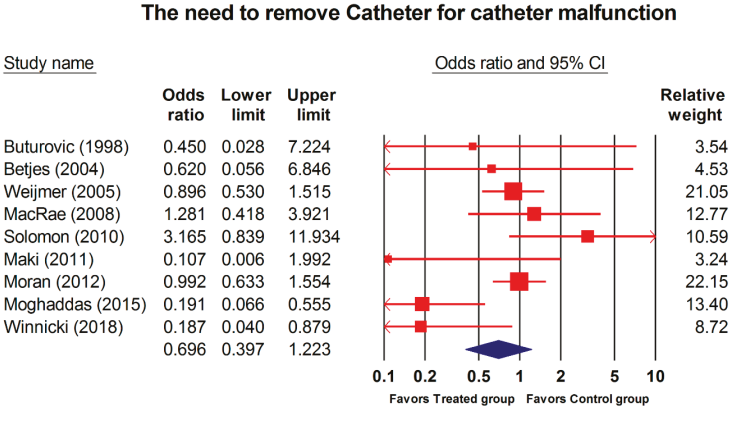


Figure 4. Forest plot of the overall odds ratios for catheter removal due to catheter malfunction in the treated group vs. the control group. The random model of overall odds ratio for the need to remove the catheter for malfunction showed a significant overall effect of the interventions in reducing the risk for catheter removal compared with the control condition (OR, 0.696; 95% CI, 0.397–1.223; $p = 0.208$).

Table 4. Subgroup analysis of odds ratio based on study designs, combined regimen, regimen containing antibiotic, and concentration of regimen for catheter removal due to catheter malfunction.

Subgroup	Odds Ratio	95% Confidence Interval
Combined regimen		
RCT	0.520	0.086–3.15
Not RCT	0.977	0.628–1.518
Not combined regimen		
RCT	0.434	0.068–2.786
Not RCT	1.106	0.392–3.124
Regimen containing antibiotic		
RCT	0.741	0.087–6.287
Not RCT	0.992	0.633–1.554
Regimen not containing antibiotic		
RCT	0.329	0.051–2.138
Not RCT	1.010	0.39–2.619
High concentration of major regimen		
RCT	0.896	0.029–27.554
Low concentration of major regimen		
RCT	0.479	0.051–4.537
Not RCT	0.995	0.663–1.494

Abbreviation: RCT, randomized controlled trial.

3.3.2. Thrombolytic Treatment Due to Catheter Malfunction

Overall, in eight RCTs (2092 patients; 220,460 catheter days) included in this meta-analysis the patients underwent thrombolytic treatment [4,8,14,17,18,32,33]. The incidence of thrombolytic treatment was not significantly lower in the treated group compared with the control group using the random-effects model (OR, 1.105; 95% CI, 0.655–1.573; $p = 0.946$; Figure 5). Thrombolytic treatment subgroup analysis showed no differences in the OR between the two groups (Table 5).

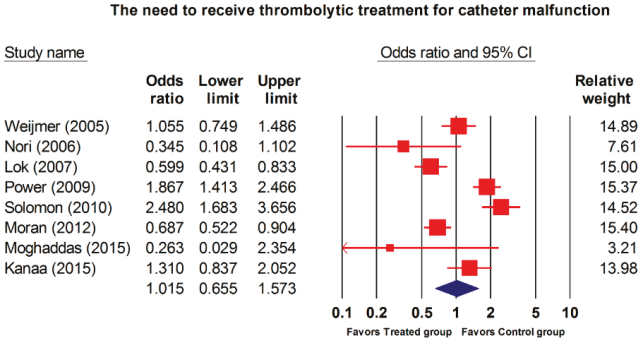


Figure 5. Forest plot of the overall odds ratios for thrombolytic treatments for catheter malfunction in the treated group versus the control group. The random model of overall odds ratio for the need to administer thrombolytic treatment for catheter malfunction showed a significant overall reduced risk for receiving thrombolytic treatments with interventions as compared with the control condition (OR, 1.105; 95% CI, 0.655–1.573; $p = 0.946$).

Table 5. Subgroup analysis of odds ratio based on study designs, combined regimen, regimen containing antibiotic, and concentration of regimen for the need of thrombolytic treatment for catheter malfunction.

Subgroup	Odds Ratio	95% Confidence Interval
Combined regimen		
RCT	2.480	1.214–5.066
Not RCT	0.620	0.382–1.004
Not combined regimen		
RCT	1.320	0.888–1.961
Not RCT	0.599	0.344–1.043
Regimen containing antibiotic		
RCT	1.969	0.944–4.107
Not RCT	0.620	0.382–1.004
Regimen not containing antibiotic		
RCT	1.385	0.893–2.149
Not RCT	0.345	0.108–1.102
High concentration of major regimen		
RCT	1.415	0.784–2.554
Low concentration of major regimen		
RCT	2.480	1.042–5.902
Not RCT	0.637	0.518–0.783

Abbreviation: RCT, randomized controlled trial.

3.3.3. All-Cause Mortality

The meta-analysis included five RCTs (2,327 patients) comparing all-cause mortality rate between the two groups; no significant difference was identified (OR, 0.909; 95% CI, 0.580–1.423; $p = 0.676$; Figure 6). The corresponding subgroup analysis (combined regimen, regimen containing antibiotic, and concentration of regimen for all-cause mortality) showed no apparent differences between the two groups (Table 6).

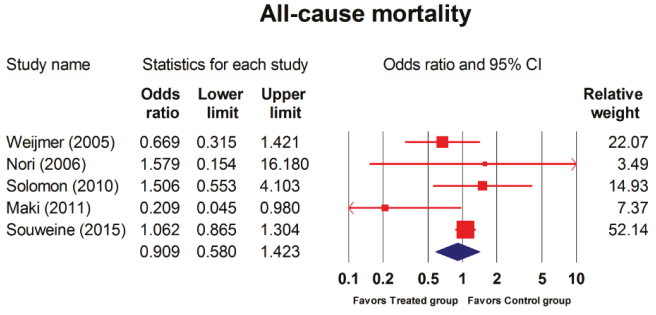


Figure 6. Forest plot of the overall odds ratios for all-cause mortality in the treated group versus the control group. The random model of overall odds ratio for all-cause mortality rate showed a significant overall effect of the interventions in reducing mortality rate as compared with the control condition (OR, 0.909; 95% CI, 0.580–1.423; $p = 0.676$).

Table 6. Subgroup analysis of odds ratio based on study designs, combined regimen, regimen containing antibiotic, and concentration of regimen for all-cause mortality.

Subgroup	Odds Ratio	95% Confidence Interval
Combined regimen		
RCT	0.725	0.237–2.211
Not RCT	1.579	0.154–16.18
Not combined regimen		
RCT	0.884	0.404–1.933
Regimen containing antibiotic		
RCT	1.506	0.388–5.838
Not RCT	1.579	0.154–16.18
Regimen not containing antibiotic		
RCT	0.723	0.367 – 1.425
High concentration of major regimen		
RCT	0.669	0.054–8.324
Low Concentration of major regimen		
RCT	0.615	0.09–4.22
Not RCT	1.579	0.154–16.18

Abbreviation: RCT, randomized controlled trial.

3.4. Pooled Odds for Outcomes in Trial Sequential Analysis

In trial sequential analysis between the treated and control groups, the overall OR of CRBSI was 0.439 (95% CI, 0.290–0.668; $p < 0.001$; Figure 7a), the OR of ESI was 0.644 (95% CI, 0.469–0.883; $p = 0.006$; Figure 7b), the OR of the need to remove the catheter for catheter malfunction was 0.746 (95% CI, 0.431–1.293; $p = 0.151$; Figure 7c), the OR of the need to receive thrombolytic treatment for catheter malfunction was 1.015 (95% CI, 0.655–1.573; $p = 0.461$; Figure 7d), and the OR of all-cause mortality was 0.976 (95% CI, 0.663–1.439; $p = 0.296$; Figure 7e).

(a) Trial sequential analysis of catheter-related bloodstream infection

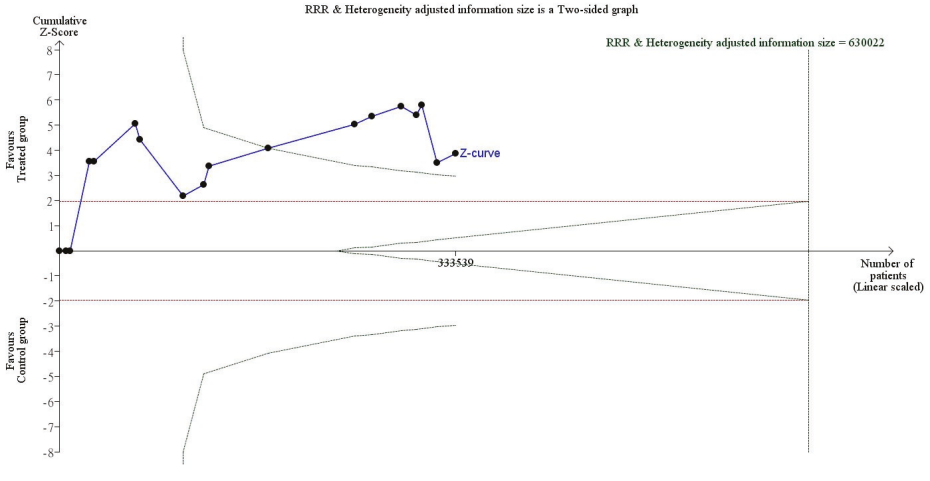
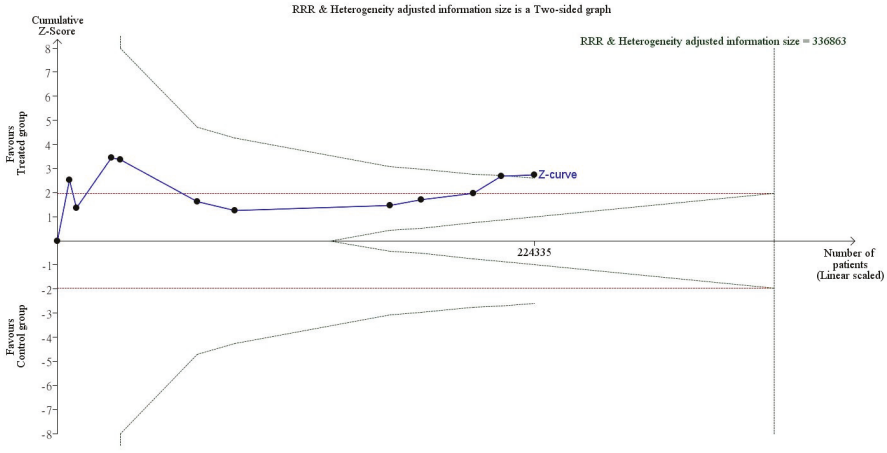
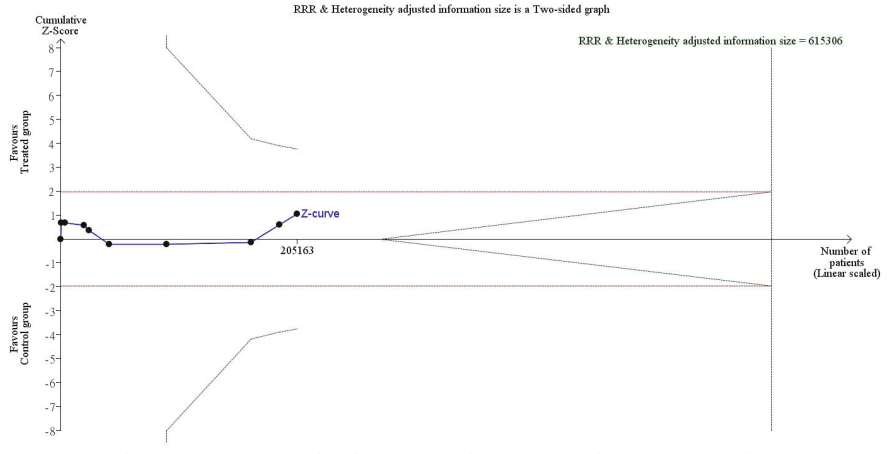


Figure 7. *Cont.*

(b) Trial sequential analysis of exit-site infection



(c) Trial sequential analysis of catheter removal for catheter malfunction



(d) Trial sequential analysis of thrombolytic treatments for catheter malfunction

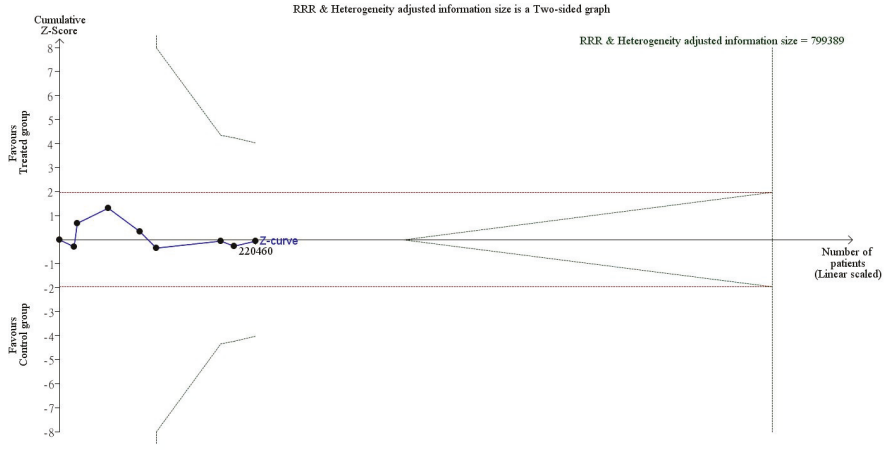


Figure 7. *Cont.*

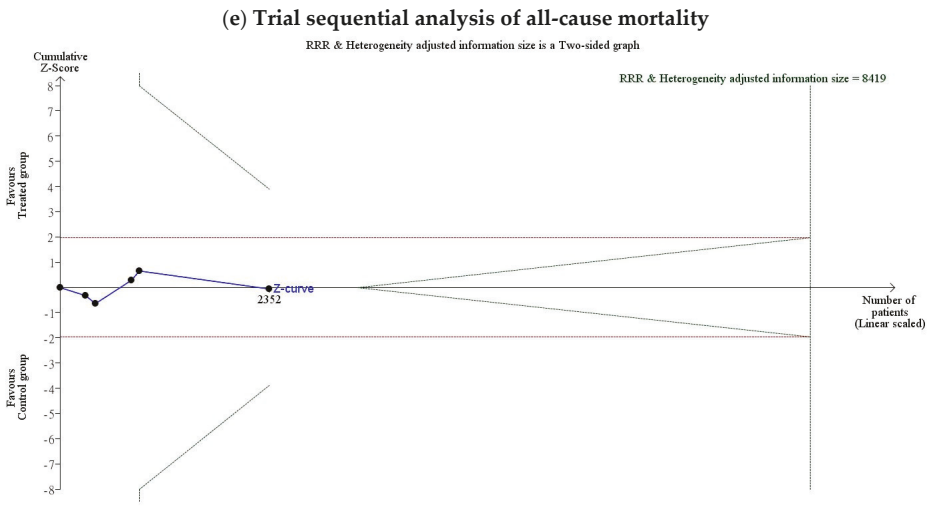


Figure 7. Trial sequential analysis of the odds ratio for evaluation event: (**a**) Trial sequential analysis of catheter-related bloodstream infection. Trial sequential analysis of 17 studies with a lower risk of bias in reporting catheter-related bloodstream infection, with a control event proportion of 17%, diversity of 45%, type I error of 5%, power of 80%, and relative risk reduction of 30%. The required information size of 630,022 was not reached and none of the boundaries for benefit, harm, or futility were crossed, leaving the meta-analysis inconclusive at a 30% relative risk reduction. The overall OR of CRBSI was 0.439 (95% CI, 0.290–1.668; $p < 0.001$); (**b**) trial sequential analysis of exit-site infection. Trial sequential analysis of eleven studies with low risk of bias reporting exit-site infection, with a control event proportion of 17%, diversity of 30%, type I error of 5%, power of 80%, and relative risk reduction of 30%. The required information size of 336,863 was not reached and none of the boundaries for benefit, harm, or futility were crossed, leaving the meta-analysis inconclusive at a 30% relative risk reduction. The OR of ESI was 0.644 (95% CI, 0.469–0.883; $p = 0.006$); (**c**) trial sequential analysis of nine studies with a lower risk of bias reporting the need to remove the catheter for catheter malfunction, with a control event proportion of 17%, diversity of 71%, type I error of 5%, power of 80%, and relative risk reduction of 30%. The required information size of 625,306 were not reached and none of the boundaries for benefit, harm, or futility were crossed, leaving the meta-analysis inconclusive at a 30% relative risk reduction. The OR of the need to remove the catheter for catheter malfunction was 0.746 (95% CI, 0.431–1.293; $p = 0.151$); (**d**) trial sequential analysis of thrombolytic treatments for catheter malfunction. Trial sequential analysis of nine studies with low risk of bias reporting the need to receive thrombolytic treatment for catheter malfunction, with a control event proportion of 17%, diversity of 91%, type I error of 5%, power of 80%, and relative risk reduction of 30%. The required information size of 615,306 were not reached and none of the boundaries for benefit, harm, or futility were crossed, leaving the meta-analysis inconclusive at a 30% relative risk reduction. The OR of the need to receive thrombolytic treatment for catheter malfunction was 1.015 (95% CI, 0.655–1.573; $p = 0.461$); (**e**) trial sequential analysis of all-cause mortality. Trial sequential analysis of five studies with a lower risk of bias reporting all-cause mortality, with a control event proportion of 17%, diversity of 78%, type I error of 5%, power of 80%, and relative risk reduction of 30%. The required information size of 8419were not reached and none of the boundaries for benefit, harm, or futility were crossed, leaving the meta-analysis inconclusive at a 30% relative risk reduction. The OR of all-cause mortality was 0.976 (95% CI, 0.663–1.439; $p = 0.296$). Notes: The solid blue line is the cumulative Z-curve. The vertical black dashed line is required information size. The green dashed lines represent the trial sequential monitoring boundaries and the futility boundaries.

3.5. Funnel Plot for the Overall OR of the Included Studies among Four Outcomes

We examined possible sources of underlying heterogeneity across studies. With regards to OR heterogeneity, the I^2 value was calculated in both the overall studies included. In the funnel plot of the OR for evaluation event, the I^2 value of CRBSI was 70.1% (p = 0.303, Figure 8a), ESI was 28.0% (p = 0.010; Figure 8b), the need to remove the catheter for catheter malfunction was 55.9% (p = 0.208; Figure 8c), the need to receive thrombolytic treatment for catheter malfunction was 88.69% (p = 0.946; Figure 8d), and all-cause mortality was 88.6% (p = 0.804; Figure 8e).

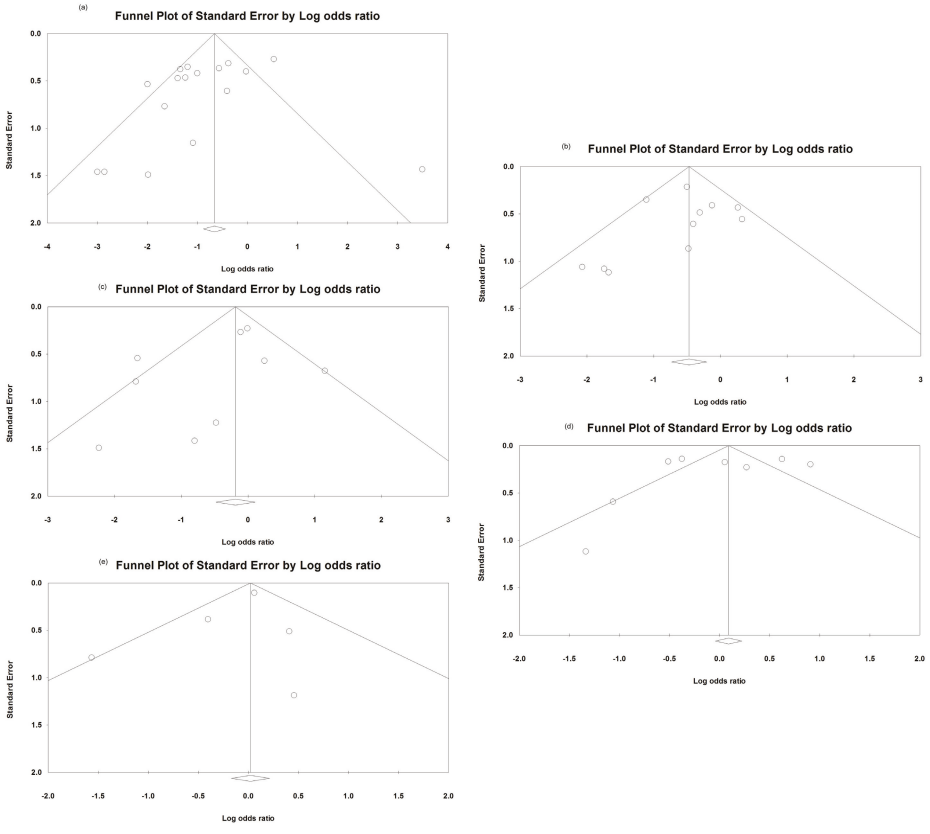


Figure 8. Funnel plot of the odds ratio for evaluation event: (**a**) Funnel plot of the odds ratio of catheter-related bloodstream infection. I^2 value, 70.1%; p = 0.303; (**b**) funnel plot of the odds ratio of exit-site infection. I^2 value, 28.0%; p = 0.010; (**c**) funnel plot of the odds ratio of catheter removal for catheter malfunction. I^2 value, 55.9%; p = 0.208; (**d**) funnel plot of the odds ratio of thrombolytic treatments for catheter malfunction. I^2 value, 88.69%; p = 0.946; (**e**) funnel plot of the odds ratio of all-cause mortality. I^2 value, 88.6%; p = 0.804. Regarding odds ratio heterogeneity, the I^2 value in both the overall studies included is indicated for each case. Egger's test revealed the existence of significant publication bias regarding the overall odds ratios, p-value is indicated for each case.

4. Discussion

Our meta-analysis and trial sequential analysis shows that routine locking solutions for hemodialysis catheters could effectively reduce the incidence of CRBSI and ESI. Our current meta-analysis, based on 21 selected studies with a total of 6118 participants, showed that the incidence of CRBSI and ESI significantly decreased in the treated group relative to the control group,

that is less infections when using citrate or citrate mixtures versus heparin. Moreover, we found no significant difference in preserving catheter function, including in the need for catheter withdrawal or for thrombolytic treatment due to catheter malfunction, between the treated and control groups. We found no significant alteration in all-cause mortality between the two groups. The lack of statistical significance may not only be due to the heterogeneity and underlying variance in the outcomes of each regimen, but also due to inadequate required information sizes, as revealed by the trial sequential analysis. Regular locking care with citrate is standard practice for patients undergoing hemodialysis in many healthcare institutes, but not in some countries including Taiwan. Our updated review suggests that the role of routine locking solutions in preventing CRBSI and ESI in hemodialysis patients is robust. However, it does not show a benefit in preserving catheter function in hemodialysis patients, including in the need to remove catheters or in the need for thrombolytic treatment for catheter malfunction.

The current study shows that the incidence of CRBSI significantly decreased in the treated group relative to the control group, which is consistent with previous studies [36,37]. Subgroup analyses based on the type of locking solutions for hemodialysis catheters revealed that the usage of citrate-base regimens was associated with a lower incidence of CRBSI [4,14]. Our subgroup analysis for the concentration of citrate used showed that the incidence of CRBSI was similar in treated group, although the American Society of Diagnostic and Interventional Nephrology and the European Renal Best Practice recommend 4% citrate to be used as a catheter locking solution [38,39]. In some countries, including Taiwan, 4% citrate is still not routinely used in locking solutions for hemodialysis catheters. The current meta-analysis emphasizes that 4% citrate shows a benefit and could be routinely used as a locking solution for hemodialysis catheters.

Our current study shows that the incidence of ESI is significantly decreased in the treated group compared with the control group. Our result is in agreement with previous studies [14,19]. In some studies, patients received additional antibiotic ointments at the exit site during dressing changes, which could reduce the incidence of ESI [8,14,40]. After subgroup categorization, there is no significant difference between two groups except for combined regimen, which could result from the heterogeneity of the included studies and inadequate information size.

We found no significant difference in preserving catheter function between the treated and control groups, including the need to remove catheters or the need for thrombolytic treatment. However, Yahav et al. reported that citrate reduced catheter removals [41]. This incongruity may arise from the following: (1) Variation in enrollment criteria and definitions for the spectrum of catheter removal and (2) the number of cases is still limited because the meta-analysis information size does not meet the required information size. Concerning thrombolytic treatments and thrombosis episodes, our report is similar to previous studies [41,42]. Focusing on the need to remove catheters and to receive thrombolytic treatment for catheter malfunction, further large-scale RCTs are necessary to elucidate this issue for preserving catheter function.

The possible association between the two groups and all-cause mortality was not statistically significant in the current study (OR, 0.909; 95% CI, 0.580–1.423; $p = 0.676$). Subgroup analysis showed no difference in all-cause mortality. Mortality due to CRBSIs or ESIs account around one-tenth of all hemodialysis patient deaths [1–4]. Protective strategies with locking solutions to prevent CRBSIs and ESIs in hemodialysis patients still cannot decrease the mortality rate. Further large-scale RCTs are necessary to elucidate modifiable risk factor for decreasing morality in hemodialysis patients.

Guidelines for the Prevention of Intravascular Catheter-Related Infections has been published by the Center for Disease Control and Prevention [15], which recommends using prophylactic antimicrobial locking solution in patients undergoing hemodialysis who have a history of multiple CRBSI, despite optimal maximal adherence to aseptic techniques (Category II). This recommendation has been embraced by some dialysis centers due to the low execution rate of locking solutions in preventing CRBSI in hemodialysis patients. In fact, many challenges persist in managing daily care in dialysis centers, such as a lack of safety locking solutions for hemodialysis catheters, lack of a designated health-care workers to perform locking care, limited training on catheter care among

health-care workers of dialysis centers, potential hemodialysis patients' noncompliance due to discomfort, as well as health-care workers being unable to maintain high adherence rates in conducting care procedures.

The current study has several limitations. Firstly, the enrolled trials and studies included in the primary analysis dealt with different indications for outcome measures by randomizing a variety of patient groups in different clinical settings. Thus, there is the risk of introducing potentially heterogeneity. Additionally, it is difficult to perform a subgroup analysis based on conditions, such as catheter type, heparin dosage, and other differences in individual unit practices. Secondly, differences in the study individuals, disease severity, setting, and type of infections between individual studies made the study population highly heterogeneous. The I^2 value for OR heterogeneity ranged from 25% to 50%, and this heterogeneity would impact the findings of this meta-analysis. Thus, the influence of measurement precision was considered when reporting treatment effectiveness using ORs. Due to the lack of adjusted data in our selected trials, we compiled the unadjusted ORs. We therefore suggest that future similar trials should record serial changes in catheter function and infection status to provide a more accurate indication of clinical effectiveness. Regardless of aforementioned limitations, we have minimized bias throughout the process by our methods of study identification, data selection, and statistical analysis, as well as in our control of publication bias. These steps should strengthen the stability and accuracy of the meta-analysis. Our findings of this meta-analysis are reliable to provide suggestions for improving clinical care.

5. Conclusions

In conclusion, our study demonstrated that routine locking solutions for hemodialysis catheters could effectively reduce the incidence of CRBSI and ESI. Our findings showed no benefit of routine locking solutions for hemodialysis catheters in decreasing all-cause mortality as well as preserving catheter function, including in the need to remove catheters and in the need to receive thrombolytic treatment, both due to catheter malfunction. The latter results lack statistical significance and the comparisons are limited due to the heterogeneity of the included trials and inadequate information size. Therefore, further well-conducted observational studies and randomized controlled trials are urgently needed to conclusively evaluate the impact of routine locking solutions on preserving catheter function and improving the rates of all-cause mortality.

Supplementary Materials: The following are available online at http://www.mdpi.com/2077-0383/8/3/412/s1, Table S1: PRISMA 2009 Checklist.

Author Contributions: Conceptualization, C.-H.C., Y.-M.C., Y.Y., Y.-J.C., L.-J.Y. and H.-C.Y.; Methodology, C.-H.C., Y.-M.C. and Y.-J.C.; Software, C.-H.C., Y.-M.C. and Y.-J.C.; Validation, C.-H.C., Y.-M.C., Y.Y. and Y.-J.C.; Formal Analysis, C.-H.C., Y.-M.C. and Y.-J.C.; Investigation, X.X.; Data Curation, C.-H.C.; Writing—Original Draft Preparation, C.-H.C.; Writing—Review and Editing, C.-H.C., Y.-M.C., Y.Y. and Y.-J.C.; Writing—Approval, C.-H.C., Y.-M.C., Y.Y., Y.-J.C., L.-J.Y. and H.-C.Y.; Visualization, C.-H.C. and Y.-M.C.; Project Administration, C.-H.C.; Funding Acquisition, C.-H.C.

Funding: The present work was partially supported by a grant obtained from the Changhua Christian Hospital (106-CCHIRP-097).

Acknowledgments: The authors thank Ping-Tao Tseng for technical analysis. The authors thank the staff at the Epidemiology and Biostatistics Center, the Department of Pharmacology, and the Department of Computer at Changhua Christian Hospital for the literature findings. This research project would not have been possible without the support of many people. The authors wish to express their gratitude to the staffs of the Department of Critical Care, Epidemiology and Biostatistics Center, the Division of Nephrology, the Division of Infectious Diseases, the Department of Pharmacology, the Department of Nursing, and the Department of Healthcare Quality at Changhua Christian Hospital who were extremely helpful and provided invaluable assistance and support.

Conflicts of Interest: All authors declare that they have no competing interests. The sponsors had no role in the design, execution, interpretation, or writing of the study.

Appendix A

Supplement Search strategy in PubMed

#1 lock
#2 filling solution
#3 #1 or #2
#4 End-Stage Kidney Disease or Disease, End-Stage Kidney, End Stage Kidney Disease, End-Stage Chronic Kidney Failure, End-Stage Renal Disease, End-Stage Renal Disease, Chronic Chronic Renal Failure, or ESRD
#5 Renal Dialyses, Renal Dialysis, Hemodialyses, Extracorporeal Dialysis or Renal replacement therapy
#6 Catheter Related Infections, Catheter-Related Infection
#7 #4 or #5 and #6
#8 #3 and #7

References

1. Liu, K.D.; Chertow, G.M. Dialysis in the Treatment of Renal Failure. In *Harrison's Principles of Internal Medicine, 20e*; Jameson, J.L., Fauci, A.S., Kasper, D.L., Hauser, S.L., Longo, D.L., Loscalzo, J., Eds.; McGraw-Hill Education: New York, NY, USA, 2018.
2. Schwab, S.J.; Beathard, G. The hemodialysis catheter conundrum: Hate living with them, but can't live without them. *Kidney Int.* **1999**, *56*, 1–17. [CrossRef] [PubMed]
3. Zhao, Y.; Li, Z.; Zhang, L.; Yang, J.; Yang, Y.; Tang, Y.; Fu, P. Citrate versus heparin locking for hemodialysis catheters: A systematic review and meta-analysis of randomized controlled trials. *Am. J. Kidney Dis.* **2014**, *63*, 479–490. [CrossRef] [PubMed]
4. Power, A.; Duncan, N.; Singh, S.K.; Brown, W.; Dalby, E.; Edwards, C.; Lynch, K.; Prout, V.; Cairns, T.; Griffith, M.; et al. Sodium citrate versus heparin catheter lockings for cuffed central venous catheters: A single-center randomized controlled trial. *Am. J. Kidney Dis.* **2009**, *53*, 1034–1041. [CrossRef] [PubMed]
5. Allon, M. Current management of vascular access. *Clin. J. Am. Soc. Nephrol.* **2007**, *2*, 786–800. [CrossRef]
6. Maki, D.G.; Ash, S.R.; Winger, R.K.; Lavin, P.; Investigators, A.T. A novel antimicrobial and antithrombotic locking solution for hemodialysis catheters: A multi-center, controlled, randomized trial. *Crit. Care Med.* **2011**, *39*, 613–620. [CrossRef] [PubMed]
7. Yevzlin, A.S.; Sanchez, R.J.; Hiatt, J.G.; Washington, M.H.; Wakeen, M.; Hofmann, R.M.; Becker, Y.T. Concentrated heparin locking is associated with major bleeding complications after tunneled hemodialysis catheter placement. In *Seminars in Dialysis*; Blackwell Publishing Ltd.: Oxford, UK, 2007; Volume 20, pp. 351–354.
8. Moran, J.; Sun, S.; Khababa, I.; Pedan, A.; Doss, S.; Schiller, B. A randomized trial comparing gentamicin/citrate and heparin lockings for central venous catheters in maintenance hemodialysis patients. *Am. J. Kidney Dis.* **2012**, *59*, 102–107. [CrossRef]
9. Shanks, R.M.; Donegan, N.P.; Graber, M.L.; Buckingham, S.E.; Zegans, M.E.; Cheung, A.L.; O'Toole, G.A. Heparin stimulates *Staphylococcus aureus* biofilm formation. *Infect. Immun.* **2005**, *73*, 4596–4606. [CrossRef]
10. Shanks, R.M.; Sargent, J.L.; Martinez, R.M.; Graber, M.L.; O'Toole, G.A. Catheter locking solutions influence staphylococcal biofilm formation on abiotic surfaces. *Nephrol. Dial. Transplant.* **2006**, *21*, 2247–2255. [CrossRef]
11. Ibberson, C.B.; Parlet, C.P.; Kwiecinski, J.; Crosby, H.A.; Meyerholz, D.K.; Horswill, A.R. Hyaluronan Modulation Impacts *Staphylococcus aureus* Biofilm Infection. *Infect. Immun.* **2016**, *84*, 1917–1929. [CrossRef] [PubMed]
12. Macrae, J.M.; Dojcinovic, I.; Djurdjev, O.; Jung, B.; Shalansky, S.; Levin, A.; Kiaii, M. Citrate 4% versus heparin and the reduction of thrombosis study (CHARTS). *Clin. J. Am. Soc. Nephrol.* **2008**, *3*, 369–374. [CrossRef]
13. Lok, C.E.; Appleton, D.; Bhola, C.; Khoo, B.; Richardson, R.M. Trisodium citrate 4%—An alternative to heparin capping of haemodialysis catheters. *Nephrol. Dial. Transplant.* **2007**, *22*, 477–483. [CrossRef] [PubMed]

14. Weijmer, M.C.; van den Dorpel, M.A.; Van de Ven, P.J.; ter Wee, P.M.; van Geelen, J.A.; Groeneveld, J.O.; van Jaarsveld, B.C.; Koopmans, M.G.; le Poole, C.Y.; Schrander-Van der Meer, A.M.; et al. Randomized, clinical trial comparison of trisodium citrate 30% and heparin as catheter-lockinging solution in hemodialysis patients. *J. Am. Soc. Nephrol.* **2005**, *16*, 2769–2777. [CrossRef]
15. O'grady, N.P.; Alexander, M.; Burns, L.A.; Dellinger, E.P.; Garland, J.; Heard, S.O.; Lipsett, P.A.; Masur, H.; Mermel, L.A.; Pearson, M.L. Guidelines for the prevention of intravascular catheter-related infections. *Clin. Infect. Dis.* **2011**, *52*, e162–e193. [CrossRef] [PubMed]
16. Thorlund, K.; Imberger, G.; Walsh, M.; Chu, R.; Gluud, C.; Wetterslev, J.; Guyatt, G.; Devereaux, P.J.; Thabane, L. The number of patients and events required to limit the risk of overestimation of intervention effects in meta-analysis—A simulation study. *PLoS ONE* **2011**, *6*, e25491. [CrossRef] [PubMed]
17. Kanaa, M.; Wright, M.J.; Akbani, H.; Laboi, P.; Bhandari, S.; Sandoe, J.A. Cathasept Line Locking and Microbial Colonization of Tunneled Hemodialysis Catheters: A Multicenter Randomized Controlled Trial. *Am. J. Kidney Dis.* **2015**, *66*, 1015–1023. [CrossRef] [PubMed]
18. Moghaddas, A.; Abbasi, M.R.; Gharekhani, A.; Dashti-Khavidaki, S.; Razeghi, E.; Jafari, A.; Khalili, H. Prevention of hemodialysis catheter-related blood stream infections using a cotrimoxazole-locking technique. *Future Microbiol.* **2015**, *10*, 169–178. [CrossRef]
19. Souweine, B.; Lautrette, A.; Gruson, D.; Canet, E.; Klouche, K.; Argaud, L.; Bohe, J.; Garrouste-Orgeas, M.; Mariat, C.; Vincent, F.; et al. Ethanol locking and risk of hemodialysis catheter infection in critically ill patients. A randomized controlled trial. *Am. J. Respir. Crit. Care Med.* **2015**, *191*, 1024–1032. [CrossRef]
20. Chu, G.; Fogarty, G.M.; Avis, L.F.; Bergin, S.; McElduff, P.; Gillies, A.H.; Choi, P. Low dose heparin locking (1000 U/mL) maintains tunnelled hemodialysis catheter patency when compared with high dose heparin (5000 U/mL): A randomised controlled trial. *Hemodial. Int.* **2016**, *20*, 385–391. [CrossRef]
21. Zwiech, R.; Adelt, M.; Chrul, S. A Taurolidine-Citrate-Heparin Locking Solution Effectively Eradicates Pathogens from the Catheter Biofilm in Hemodialysis Patients. *Am. J. Ther.* **2016**, *23*, e363–e368. [CrossRef]
22. Correa Barcellos, F.; Pereira Nunes, B.; Jorge Valle, L.; Lopes, T.; Orlando, B.; Scherer, C.; Nunes, M.; Araujo Duarte, G.; Bohlke, M. Comparative effectiveness of 30% trisodium citrate and heparin locking solution in preventing infection and dysfunction of hemodialysis catheters: A randomized controlled trial (CITRIM trial). *Infection* **2017**, *45*, 139–145. [CrossRef] [PubMed]
23. Sofroniadou, S.; Revela, I.; Kouloubinis, A.; Makriniotou, I.; Zerbala, S.; Smirloglou, D.; Kalocheretis, P.; Drouzas, A.; Samonis, G.; Iatrou, C. Ethanol combined with heparin as a lockinging solution for the prevention of catheter related blood stream infections in hemodialysis patients: A prospective randomized study. *Hemodial. Int.* **2017**, *21*, 498–506. [CrossRef] [PubMed]
24. Winnicki, W.; Herkner, H.; Lorenz, M.; Handisurya, A.; Kikic, Z.; Bielesz, B.; Schairer, B.; Reiter, T.; Eskandary, F.; Sunder-Plassmann, G.; et al. Taurolidine-based catheter locking regimen significantly reduces overall costs, infection, and dysfunction rates of tunneled hemodialysis catheters. *Kidney Int.* **2018**, *93*, 753–760. [CrossRef]
25. Jadad, A.R.; Moore, R.A.; Carroll, D.; Jenkinson, C.; Reynolds, D.J.M.; Gavaghan, D.J.; McQuay, H.J. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Control. Clin. Trials* **1996**, *17*, 1–12. [CrossRef]
26. Wells, G.; Shea, B.; O'Connell, D.; Peterson, J.; Welch, V.; Losos, M. *Newcastle-Ottawa Quality Assessment Scale*; Ottawa Hospital Research Institute: Ottawa, ON, Canada, 2013.
27. Higgins, J.P.; Thompson, S.G. Quantifying heterogeneity in a meta-analysis. *Stat. Med.* **2002**, *21*, 1539–1558. [CrossRef]
28. 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]
29. Buturovic, J.; Ponikvar, R.; Kandus, A.; Boh, M.; Klinkmann, J.; Ivanovich, P. Filling hemodialysis catheters in the interdialytic period: Heparin versus citrate versus polygeline: A prospective randomized study. *Artif. Organs* **1998**, *22*, 945–947. [CrossRef]
30. Dogra, G.K. Prevention of Tunneled Hemodialysis Catheter-Related Infections Using Catheter-Restricted Filling with Gentamicin and Citrate: A Randomized Controlled Study. *J. Am. Soc. Nephrol.* **2002**, *13*, 2133–2139. [CrossRef]

31. Betjes, M.G.; van Agteren, M. Prevention of dialysis catheter-related sepsis with a citrate-taurolidine-containing locking solution. *Nephrol. Dial. Transplant.* **2004**, *19*, 1546–1551. [CrossRef] [PubMed]
32. Nori, U.S.; Manoharan, A.; Yee, J.; Besarab, A. Comparison of low-dose gentamicin with minocycline as catheter locking solutions in the prevention of catheter-related bacteremia. *Am. J. Kidney Dis.* **2006**, *48*, 596–605. [CrossRef] [PubMed]
33. Solomon, L.R.; Cheesbrough, J.S.; Ebah, L.; Al-Sayed, T.; Heap, M.; Millband, N.; Waterhouse, D.; Mitra, S.; Curry, A.; Saxena, R.; et al. A randomized double-blind controlled trial of taurolidine-citrate catheter lockings for the prevention of bacteremia in patients treated with hemodialysis. *Am. J. Kidney Dis.* **2010**, *55*, 1060–1068. [CrossRef]
34. Filiopoulos, V.; Hadjiyannakos, D.; Koutis, I.; Trompouki, S.; Micha, T.; Lazarou, D.; Vlassopoulos, D. Approaches to prolong the use of uncuffed hemodialysis catheters: Results of a randomized trial. *Am. J. Nephrol.* **2011**, *33*, 260–268. [CrossRef]
35. Chen, F.K.; Li, J.J.; Song, Y.; Zhang, Y.Y.; Chen, P.; Zhao, C.Z.; Gong, H.Y.; Yao, D.F. Concentrated sodium chloride catheter locking solution—A new effective alternative method for hemodialysis patients with high bleeding risk. *Ren. Fail.* **2014**, *36*, 17–22. [CrossRef]
36. Jaffer, Y.; Selby, N.M.; Taal, M.W.; Fluck, R.J.; McIntyre, C.W. A meta-analysis of hemodialysis catheter lockinging solutions in the prevention of catheter-related infection. *Am. J. Kidney Dis.* **2008**, *51*, 233–241. [CrossRef]
37. Labriola, L.; Crott, R.; Jadoul, M. Preventing haemodialysis catheter-related bacteraemia with an antimicrobial locking solution: A meta-analysis of prospective randomized trials. *Nephrol. Dial. Transplant.* **2008**, *23*, 1666–1672. [CrossRef]
38. Moran, J.E.; Ash, S.R. Lockinging solutions for hemodialysis catheters; heparin and citrate—A position paper by ASDIN. *Semin. Dial.* **2008**, *21*, 490–492. [CrossRef] [PubMed]
39. Vanholder, R.; Canaud, B.; Fluck, R.; Jadoul, M.; Labriola, L.; Marti-Monros, A.; Tordoir, J.; Van Biesen, W. Catheter-related blood stream infections (CRBSI): A European view. *Nephrol. Dial. Transplant.* **2010**, *25*, 1753–1756. [CrossRef]
40. Jenks, M.; Craig, J.; Green, W.; Hewitt, N.; Arber, M.; Sims, A. Tegaderm CHG IV Securement Dressing for Central Venous and Arterial Catheter Insertion Sites: A NICE Medical Technology Guidance. *Appl. Health Econ. Health Policy* **2016**, *14*, 135–149. [CrossRef]
41. Yahav, D.; Rozen-Zvi, B.; Gafter-Gvili, A.; Leibovici, L.; Gafter, U.; Paul, M. Antimicrobial locking solutions for the prevention of infections associated with intravascular catheters in patients undergoing hemodialysis: Systematic review and meta-analysis of randomized, controlled trials. *Clin. Infect. Dis.* **2008**, *47*, 83–93. [CrossRef] [PubMed]
42. Yon, C.K.; Low, C.L. Sodium citrate 4% versus heparin as a locking solution in hemodialysis patients with central venous catheters. *Am. J. Health Syst. Pharm.* **2013**, *70*, 131–136. [CrossRef] [PubMed]

Journal of
Clinical Medicine

MDPI

Article

Predictors of Discordance in the Assessment of Skeletal Muscle Mass between Computed Tomography and Bioimpedance Analysis

Min Ho Jo [1], Tae Seop Lim [1,2], Mi Young Jeon [1,2], Hye Won Lee [1,2], Beom Kyung Kim [1,2,3], Jun Yong Park [1,2,3], Do Young Kim [1,2,3], Sang Hoon Ahn [1,2,3], Kwang-Hyub Han [1,2,3] and Seung Up Kim [1,2,3,*]

[1] Department of Internal Medicine, Institute of Gastroenterology, Yonsei University College of Medicine, Seoul 03722, Korea; zns-1@hanmail.net (M.H.J.); TSLIM21@yuhs.ac (T.S.L.); HYUK4385@yuhs.ac (M.Y.J.); LORRY-LEE@yuhs.ac (H.W.L.); BEOMKKIM@yuhs.ac (B.K.K.); DRPJY@yuhs.ac (J.Y.P.); DYK1025@yuhs.ac (D.Y.K.); AHNSH@yuhs.ac (S.H.A.); GIHANKHYS@yuhs.ac (K.-H.H.)
[2] Yonsei Liver Center, Severance Hospital, Seoul 03722, Korea
[3] Institute of Gastroenterology, Severance Hospital Yonsei University College of Medicine, Seoul 03722, Korea
* Correspondence: KSUKOREA@yuhs.ac; Tel.: +82-2-2228-1944

Received: 20 February 2019; Accepted: 1 March 2019; Published: 7 March 2019

Abstract: Computed tomography (CT) and bioimpedance analysis (BIA) can assess skeletal muscle mass (SMM). Our objective was to identify the predictors of discordance between CT and BIA in assessing SMM. Participants who received a comprehensive medical health check-up between 2010 and 2018 were recruited. The CT and BIA-based diagnostic criteria for low SMM are as follows: Defined CT cutoff values (lumbar skeletal muscle index (LSMI) <1 standard deviation (SD) and means of 46.12 cm^2/m^2 for men and 34.18 cm^2/m^2 for women) and defined BIA cutoff values (appendicular skeletal muscle/$height^2$ <7.0 kg/m^2 for men and <5.7 kg/m^2 for women). A total of 1163 subjects were selected. The crude and body mass index (BMI)-adjusted SMM assessed by CT were significantly associated with those assessed by BIA (correlation coefficient = 0.78 and 0.68, respectively; $p < 0.001$). The prevalence of low SMM was 15.1% by CT and 16.4% by BIA. Low SMM diagnosed by CT was significantly associated with advanced age, female gender, and lower serum albumin level, whereas low SMM diagnosed by BIA was significantly associated with advanced age, female gender, and lower BMI (all $p < 0.05$). Upon multivariate analysis, age >65 years, female and BMI <25 kg/m^2 had significantly higher risks of discordance than their counterparts (all $p < 0.05$). We found a significant association between SMM assessed by CT and BIA. SMM assessment using CT and BIA should be interpreted cautiously in older adults (>65 years of age), female and BMI <25 kg/m^2.

Keywords: sarcopenia; bioimpedance analysis; computed tomography; discordance

1. Introduction

Sarcopenia is a syndrome characterized by the loss of skeletal muscle mass, strength, and performance [1–3] that results in an increased risk of fracture, dysfunction, reduced quality of life, and increased mortality [4,5]. Due to the varying diagnostic cutoff values for muscle mass and the varying diagnostic tools used in previous studies, the reported prevalence of sarcopenia has been inconsistent [6,7]. Several studies of sarcopenia have been performed, and multiple guidelines have been proposed; these have enhanced our knowledge of the condition. Sarcopenia is now officially recognized as a disorder in some countries, with an ICD-10-MC diagnostic code [8].

The measurement of skeletal muscle mass (SMM) is of paramount importance to diagnose sarcopenia. Several imaging techniques that can assess SMM are currently available: Dual-energy

X-ray absorptiometry (DXA), computed tomography (CT), magnetic resonance imaging (MRI), and bioimpedance analysis (BIA) [2]. DXA has several advantages over the other methods, such as safety, accuracy, and non-invasiveness, but it can overestimate muscle mass in cases of muscle edema or intramuscular fat deposition [9]. CT can accurately measure the quantity and quality of SMM, but it is costly and exposes the patient to radiation [2,10]. MRI has no radiation exposure for the patient and accurately measures SMM, but its clinical application is significantly limited due to its high cost [11]. BIA has been recognized as a rapid, inexpensive, portable, and safe methodology but, because BIA measures the resistance to a current that is applied through a body of water, the assessment of SMM may be inaccurate if the patients are dehydrated, overhydrated fluid status or obese [12]. BIA tends to overestimate SMM because it cannot discriminate among appendicular, non-appendicular fat, and non-fat mass [13].

To date, measurement of SMM by BIA has typically been performed using the Kyle, Jassen, Ergi, and Scafoglieri prediction models [14–17]. SMM can now be assessed directly using vertical, eight-point analyzers. Several studies have reported the accuracy and reproducibility of direct segmental multi-frequency BIA and the strong correlation between its results and SMM measured by DXA [18–23]. CT provides an accurate measurement of SMM, with a significant correlation to whole-body muscle mass [24,25]. Accordingly, CT has been considered to be the gold standard for measuring SMM [2,10]. Despite its drawbacks, BIA is more easily applied in clinical practice. Thus, investigating the prevalence of discordance in the assessment of SMM between CT and BIA and identifying predictors of this discordance is valuable. This investigation can ultimately help physicians select the optimal candidates for each modality to diagnose low SMM and interpret the results appropriately.

The primary aim of this study was to identify predictors of discordance between SMM measured by BIA and by CT. The secondary aims were to investigate the prevalence of low SMM by CT and BIA and the correlation between SMM measured by CT and BIA in apparently healthy subjects undergoing comprehensive medical health check-ups.

2. Methods

2.1. Study Subjects

A total of 1191 subjects who visited the health promotion center in Severance Hospital, a university-affiliated tertiary care hospital, for a comprehensive medical health check-up from June 2010 to April 2018 were included, see Figure 1. Severance Hospital is a 2000-bed academic referral hospital in Northwestern Seoul, Republic of Korea. Severance Hospital is supported by Yonsei University College of Medicine. Exclusion criteria were as follows: (1) no BIA data, (2) limited access to BIA data due to personal privacy, (3) poor CT quality, and (4) major operation in the lumbar area.

The study's protocol adhered to the tenets of the Declaration of Helsinki and was approved by the Institutional Review Board of Severance Hospital. Informed consents were waived due to the retrospective nature of the study.

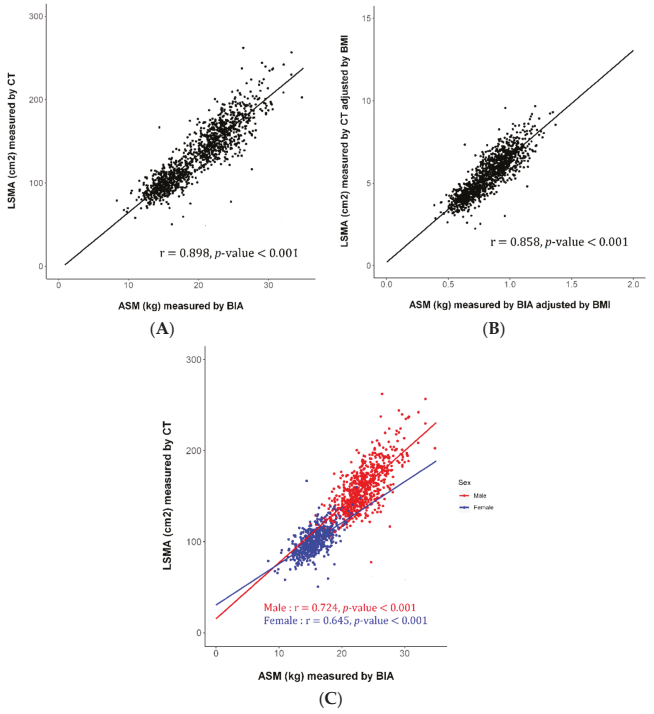


Figure 1. Correlation between SMM assessed by CT and BIA. Crude (**A**) and BMI-adjusted SMM (**B**) assessed by CT were significantly correlated with those by BIA (all $p < 0.001$, correlation coefficient = 0.898 and 0.858, respectively). The correlation between crude SMM assessed by CT and BIA was significant in men and women (**C**) (all $p < 0.001$, correlation coefficient = 0.724 in men and 0.645 in women, respectively). LSMA (cm^2) = −4.366 + 6.920 * ASM (kg), Standard error = 0.099 LSMI adjusted by BMI = 0.212 + 6.424 * (ASM adjusted by BMI), Standard error = 0.113. SMM, skeletal muscle mass; LSMA, lumbar skeletal mass area; AMS, appendicular skeletal mass; CT, computed tomography; BIA, bioimpedance analysis; BMI, body mass index; LMSI, lumbar skeletal muscle index. Regression equations and standard error are as follows.

2.2. Data Collection

A medical health check-up was performed, and collected data included age, gender, height, body weight, body mass index (BMI), and laboratory test results. Histories of hypertension, diabetes, and viral hepatitis were collected from the medical record and individual questionnaires.

2.3. Fibrosis-4 Index Calculation

Recent studies have shown that fibrotic burden in the liver is independently associated with sarcopenia. Therefore, the fibrosis-4 index (FIB-4) was calculated using the following formula: Age (years) × aspartate aminotransferase (AST) (U/L)/(platelets (10^9/L) × alanine aminotransferase (ALT) (U/L))$^{1/2}$ [26].

2.4. Measurements of Skeletal Muscle Area

Skeletal muscle area was measured at the mid-body level of the L3 vertebra in a supine position by a dual-source 128-slice CT scanner (Somatom Definition Flash, Siemens Healthcare, Forchheim, Germany), a 64-slice CT scanner (Somatom Sensation 64, Siemens Healthcare), a Discovery 710 PET-CT 128-slice scanner (General Electric Medical Systems, Milwaukee, WI, USA), a Biograph

TruePoint 40 PET-CT 40-slice scanner (Siemens Medical Solutions, Hoffman Estates, IL, USA), or a Discovery 600 PET-CT 16-slice scanner (General Electric Medical Systems, Milwaukee, WI, USA). The muscle area was identified using attenuation values between −29 to −150 Hounsfield units. Total lumbar skeletal muscle area (psoas, erector spinae, quadratus lumborum, transversus abdominus, external and internal obliques, and rectus abdominus) (cm^2) was defined as a region with density ranging from −29 to −150 Hounsfield units using Aquarius Intuition Viewer software, version 4.4.12 (Terarecon, San Mateo, CA, USA). Boundaries were corrected manually, as necessary. To minimize measurement error, the CT instruments are periodically tested and calibrated for spatial resolution, length measurement, alignment, and linearity of attenuation (CT number) using a standard phantom (AAPM CT Performance Phantom, 76-410). All tests are performed in compliance with the regulations of the Korean Institute for Accreditation of Medical Imaging. The lumbar skeletal muscle index (LSMI) was defined as 10,000 × lumbar skeletal muscle area (LSMA, cm^2)/height^2 (m^2). Based on previous studies [2,10], we assumed that measurement of SMM by CT is more accurate.

The InBody 770 (Biospace Co., Seoul, Korea) measured body composition. Participants fasted for 12 h prior to testing. Participants wore a t-shirt and short pants on the day of testing, and provided their age, gender, and height at the time of measurement. Testing was conducted according to the manufacturer's instructions. Data were uploaded to the electronic medical record. The measurement was comprised of two combinations: z-axis at frequencies of 1, 5, 50, 250, and 500 kHz for impedance, and x-axis at frequencies of 5, 50, and 250 kHz for reactance. Impedance was measured for five body segments: Trunk, right and left arms, and right and left legs. We reviewed the medical records and measured appendicular skeletal muscle (ASM) (kg) through direct segmental multi-frequency BIA [27].

Muscle mass was determined by measuring electrical resistance [28] using four surface tactile electrodes placed on the dorsal surface of the hand and foot. Whole-body resistance (R_{sumx}) was calculated by summing the segmental resistances at frequency x, according to the following equation:

$$R_{sumx} = R_{RA} + R_{LA} + R_T + R_{RL} + R_{LL} \quad (1)$$

The index of R_{sumx} (RI_{sumx}) is calculated by using the following equation:

$$RI_{sumx}\ \text{Height (cm)}^2/R_{sumx}\ (\Omega) \quad (2)$$

$$\text{Appendicular muscle mass} = 0.236 \times \text{Height}^2/R_{RA} + 0.0109 \times \text{Hright}^{2/}R_T + 0.121 \times \text{Hright}^2/R_{RL} + 1.554 \quad (3)$$

Using the formula above, the muscle mass is automatically calculated in InBody.

2.5. Definition of Low SMM

The CT diagnostic criterion for low SMM was a lumbar skeletal mass index (LSMI) <1 standard deviation (SD) below the sex-specific mean of the study group. The BIA diagnostic criterion for low SMM was adopted from the Asian Working Group of Sarcopenia [6]: ASM/height^2 <7.0 kg/m^2 for men and <5.7 kg/m^2 for women [29].

We also used additional CT and BIA diagnostic criteria for low SMM. The additional criterion for CT was an LSMI $\leq$52.4 cm^2/m^2 for men and $\leq$38.5 cm^2/m^2 for women [30]. The additional criterion for BIA was adopted from The Foundation for the National Institutes of Health: ASM/BMI <0.79 for men and <0.51 for women [7]. We attached the relevant analysis using Supplementary Data.

2.6. Statistical Analysis

Statistical analyses were performed using Statistical Package for the Social Science (SPSS) version 23.0 for Windows (IBM Corp., Armonk, NY, USA). Continuous and categorical variables were expressed as mean ± standard deviation and *n* (%), respectively. *p*-Value < 0.05 was considered statistically significant. Simple and partial correlation analyses were used to analyze the relationship between CT and BIA muscle mass. The distribution between muscle mass by BIA and quartile

stratification of muscle mass by CT was evaluated using the Mann-Whitney U test. The comparison between subjects with and without low SMM was performed using the chi-square test for categorical variables and Student's t-test for continuous variables. Multivariate analysis using binary logistic regression analysis was performed on variables that showed a p-value <0.05 and was used to determine the predictors of discordance in defining low SMM between CT and BIA.

3. Results

3.1. Patients

A total of 1191 subjects who underwent a comprehensive medical health check-up were considered eligible. However, 19 subjects were excluded due to a lack of BIA data, and an additional nine subjects were excluded due to poor-quality CT scans and a history of a major operations around the lumbar or appendicular skeletal muscle area. As a result, 1163 subjects were included in the statistical analysis, see Supplementary Figure S1.

Baseline characteristics of the study population (641 men and 521 women) are summarized in Table 1. The mean age of the patients was 57 years; 41.0% were over 60 years of age. The mean BMI of the patients was 24.0 kg/m^2. Of the study population, 41.0% of subjects (n = 488) had hypertension, 29.4% (n = 314) had diabetes, and 4.9% (n = 57) had viral hepatitis. Using CT scans, the mean whole-body fat-free mass and LSMI were 45.3 kg and 46.9 cm^2/m^2, respectively. Using BIA, the mean ASM, ASM index, and ASM/BMI ratio were 20.1 kg, 7.1 kg/m^2, and 0.82, respectively. The mean FIB-4 was 1.17.

Table 1. Baseline characteristics (n = 1163).

Variables	All
Demographic parameters	
Age, years	57 (18–92)
<40	59 (5.0)
40–49	182 (15.6)
50–59	445 (38.2)
60–69	296 (25.4)
>70	181 (15.5)
Female gender	521 (43.7)
Body mass index, kg/m^2	24.0 (15.4–43.9)
Hypertension	488 (41.0)
Diabetes mellitus	314 (26.4)
Viral hepatitis	57 (4.9)
Laboratory parameters	
Fasting glucose, mg/dL	96 (58–340)
Aspartate aminotransferase, IU/L	21 (8–140)
Alanine aminotransferase, IU/L	19 (3–196)
Serum albumin, mg/dL	4.3 (3.4–5.3)
Total bilirubin, mg/dL	0.7 (0.2–4.0)
Gamma glutamyl-transpeptidase, IU/L	23 (6–539)
Serum creatinine, mg/dL	0.8 (0.4–7.3)
Platelet count, 10^9/L	231 (89–846)
Prothrombin time, INR	0.9 (0.7–2.3)
Total cholesterol, mg/dL	187 (83–392)
Triglycerides, mg/dL	103 (31–815)
High-density lipoprotein cholesterol, mg/dL	48 (23–115)
Low-density lipoprotein cholesterol, mg/dL	109 (27–299)
HbA1c, %	5.8 (4.4–13.4)
Fibrosis-4 index	1.17 (0.20–5.47)
Muscle mass parameters	
By computed tomography	
Whole body fat-free mass, kg	45.3 (21.2–84.7)
Lumbar skeletal muscle index, cm^2/m^2	46.9 (20.0–85.6)
By bioimpedance analysis	
ASM, kg	20.1 (8.3–34.7)
ASM index, kg/m^2	7.1 (3.2–28.9)
ASM/body mass index	0.82 (0.43–1.23)

Variables are expressed as median (interquartile range) or n (%). INR, international normalized ratio; ASM, appendicular skeletal muscle mass; ASMI, appendicular skeletal mass index.

3.2. Association between SMM Assessed Using CT and BIA

The crude and BMI-adjusted SMM assessed by CT were significantly associated with those assessed by BIA ($p < 0.001$, correlation coefficient = 0.898 for crude SMM; $p < 0.001$, correlation coefficient = 0.858 for BMI-adjusted SMM), see Figure 1A. The association between crude SMM assessed by CT and BIA was statistically significant, regardless of gender ($p < 0.001$, correlation coefficient = 0.724 in men; $p < 0.001$, correlation coefficient = 0.645 in women), as shown in Figure 1. Linear regression results comparing CT and BIA assessed SMM were added to the Supplementary Table S6.

We divided the patients into four groups according to quartiles of SMM assessed by CT and BIA. SMM as assessed by BIA significantly increased according to the CT-assessed SMM quartile ($p < 0.001$), see Supplementary Figure S2.

3.3. Comparison between Subjects with and without Low SMM Assessed by CT

The baseline characteristics of subjects with and without CT-defined low SMM in Table 2. The cutoff value of low SMM was defined as less than one standard deviations sex-specific mean value of the participants. The sex-specific cut-off values of LSMI were 46.12 cm^2/m^2 in men and 34.18 c^2/m^2 in women.

Table 2. Comparison between subjects with and without low SMM assessed by CT.

Variables	without Low SMM (*n* = 988, 84.9%)	with Low SMM (*n* = 176, 15.1%)	*p*-Value
Demographic parameters			
Age, years	57 (19–92)	63 (18–92)	0.001
Female gender	435 (44.1)	86 (48.8)	0.017
Body mass index, kg/m^2	24.2 (16.5–43.8)	22.4 (15.4–28.9)	0.584
Hypertension	411 (41.5)	77 (43.7)	0.436
Diabetes mellitus	268 (27.1)	46 (26.1)	0.780
Viral hepatitis	45(4.5)	12 (6.8)	0.152
Laboratory parameters			
Fasting glucose, mg/dL	96 (58–340)	96 (65–325)	0.820
Aspartate aminotransferase, IU/L	21 (8–140)	20 (11–69)	0.964
Alanine aminotransferase, IU/L	19 (3–196)	18 (4–58)	0.756
Serum albumin, mg/dL	4.3 (3.4–5.3)	4.2 (3.5–4.8)	0.025
Total bilirubin, mg/dL	0.7 (0.2–4.0)	0.7 (0.2–2.8)	0.441
Gamma glutamyl-transpeptidase, IU/L	23 (7–398)	22 (6–539)	0.407
Serum creatinine, mg/dL	0.81 (0.38–7.3)	0.74 (0.41–2.74)	0.828
Platelet count, $10^9/L$	232 (89–846)	229 (122–438)	0.654
Prothrombin time, INR	0.93 (0.78–2.28)	0.94 (0.73–2.15)	0.574
Total cholesterol, mg/dL	188 (83–392)	177 (98–302)	0.007
Triglycerides, mg/dL	105 (31–684)	84.5 (43–815)	0.825
High-density lipoprotein cholesterol, mg/dL	48 (24–100)	50 (23–115)	0.027
Low-density lipoprotein cholesterol, mg/dL	111 (27–299)	100 (43–213)	0.038
HbA1c, %	5.8 (4.4–13.4)	5.8 (4.7–12.4)	0.436
Fibrosis-4 index	1.15 (0.20–5.47)	1.31 (0.37–3.39)	0.825
Muscle mass parameters			
By computed tomography			
Whole body fat-free mass, kg	46.4 (23.9–84.7)	40.8 (21.2–79.3)	0.530
Lumbar skeletal muscle index, cm^2/m^2	48.7 (34.2–85.6)	39.9 (20.0–46.0)	<0.001
By bioimpedance analysis			
ASM, kg	20.3 (10.3–34.7)	18.1 (8.3–27.6)	0.973
ASM index, kg/m^2	7.30 (4.62–10.58)	6.39 (3.24–8.64)	<0.001
ASM/body mass index	0.81 (0.45–1.23)	0.82 (0.43–1.17)	0.044

Variables are expressed as median (interquartile range) or *n* (%). SMM, skeletal muscle mass; CT, computed tomography; INR, international normalized ratio; ASM, appendicular skeletal muscle mass; ASMI, appendicular skeletal mass index. * CT cutoff indicates <1 standard deviation (SD), sex–specific mean value of the participants.

When CT-defined cutoff values were used, subjects with low SMM were significantly older (median 63 vs. 57 years) and female gender (48.8% vs. 44.1%). Subjects with low SMM had significantly lower serum albumin levels (median 4.2 vs. 4.3 mg/dL), lower total cholesterol (median 177 vs.

188 mg/dL), higher high-density lipoprotein (HDL) cholesterol (median 50 vs. 48 mg/dL) and lower low-density lipoprotein (LDL) cholesterol (median 100 vs. 111 mg/dL) than those of subjects without low SMM (all $p < 0.05$). In addition, various muscle indexes were unfavorable in subjects with CT-defined low SMM.

We also analyzed additional diagnostic criteria for low SMM defined by CT ($\leq$52.4 cm^2/m^2 for men and $\leq$ 38.5 cm^2/m^2 for women), see Supplementary Table S1.

3.4. Comparison between Subjects with and without Low SMM Assessed by BIA

The baseline characteristics of subjects with and without BIA-defined low SMM are shown in Table 3. The cutoff value of low SMM was defined previous study [6]. The Asian Working Group of Sarcopenia defined cutoff values appendicular lean mass (ALM)/height^2 of <7.0 kg/m^2 in men and <5.7 kg/m^2 in women.

Table 3. Comparison between subjects with and without low SMM assessed by BIA.

Variables	without Low SMM (*n* = 972, 83.6%)	with Low SMM (*n* = 191, 16.4%)	*p*-Value
Demographic parameters			
Age, years	57 (19–92)	60 (18–92)	<0.001
Female gender	392 (40.0)	129 (67.5)	<0.001
Body mass index, kg/m^2	24.2 (17.1–43.8)	21.8 (15.4–27.8)	0.005
Hypertension	411 (42.2)	77 (40.3)	0.474
Diabetes mellitus	273 (28.0)	41 (21.4)	0.339
Viral hepatitis	48 (4.9)	9 (4.7)	0.757
Laboratory parameters			
Fasting glucose, mg/dL	97 (58–340)	94 (65–265)	0.701
Aspartate aminotransferase, IU/L	21 (8–140)	20 (11–69)	0.604
Alanine aminotransferase, IU/L	20 (3–196)	17 (5–66)	0.683
Serum albumin, mg/dL	4.3 (3.4–5.2)	4.2 (3.5–5.3)	0.135
Total bilirubin, mg/dL	0.7 (0.2–4.0)	0.7 (0.3–2.5)	0.740
Gamma glutamyl-transpeptidase, IU/L	23 (6–398)	19 (7–539)	<0.001
Serum creatinine, mg/dL	0.82 (0.38–7.01)	0.69 (0.39–7.3)	0.016
Platelet count, $10^9/L$	230 (89–846)	241 (122–458)	0.256
Prothrombin time, INR	0.93 (0.73–2.28)	0.94 (0.78–2.15)	0.027
Total cholesterol, mg/dL	185 (83–392)	194 (98–300)	0.918
Triglycerides, mg/dL	106 (31–815)	88 (36–435)	0.915
High-density lipoprotein cholesterol, mg/dL	47 (23–98)	53 (29–115)	0.082
Low-density lipoprotein cholesterol, mg/dL	108 (27–299)	112 (43–213)	0.968
HbA1c, %	5.8 (4.7–13.4)	5.8 (4.4–10.5)	0.386
Fibrosis-4 index	1.16 (0.32–5.47)	1.29 (0.20–4.82)	<0.001
Muscle mass parameters			
By computed tomography			
Whole body fat-free mass, kg	47.7 (21.2~84.7)	34.4 (23.5~54.2)	0.131
Lumbar skeletal muscle index, cm^2/m^2	48.5 (20.0–85.6)	38.0 (27.0–58.1)	<0.001
By bioimpedance analysis			
ASM, kg	21.3 (12.8–34.7)	14.0 (8.3–21.6)	<0.001
ASM index, kg/m^2	7.45 (5.70–10.58)	5.56 (3.24–6.99)	<0.001
ASMI/body mass index	0.83 (0.45–1.23)	0.76 (0.43–1.08)	0.568

Variables are expressed as median (interquartile range) or *n* (%). BIA Cutoff indicates AWGS index. SMM, skeletal muscle mass; BIA, bioimpedance analysis; INR, international normalized ratio; FNIH, The Foundation for the National Institutes of Health; ALM, appendicular lean mass; ASMI, appendicular skeletal mass index; BMI, body mass index; AWGS, Asian Working Group of Sarcopenia.

When BIA-defined cutoff values were used, subjects with low SMM were significantly older (median 60 vs. 57 years) and had a higher proportion of female subjects (67.5% vs. 40.0%), lower BMI (median 21.8 vs. 24.2 kg/m^2) (all $p < 0.05$). In addition, various muscle indices were unfavorable in subjects with BIA-defined low SMM.

We also analyzed additional diagnostic criteria for low SMM defined by BIA (ALM/BMI <0.79 for men and <0.51 for women), see Supplementary Table S2.

3.5. Prevalence and Predictors of Discordance in Defining Low SMM Assessed by CT and BIA

The proportion of non-discordant and discordant subjects, when different measuring methods were applied (CT vs. BIA), is described in Table 4. The proportion of subjects without low SMM by both CT and BIA was 72.3%, and that of subjects with low SMM ranged was 3.9%. The overall proportion of non-discordant subjects was 76.2%. The results of analysis using additional diagnostic criteria for low SMM are given in Supplementary Table S3.

Table 4. Distribution of subjects with and without low SMM assessed by CT and BIA.

Muscle Mass Assessed by CT	Muscle Mass Assessed by BIA	
	* BIA Cutoff	
	without Low SMM	with Low SMM
	(*n* = 972, 83.6%)	(*n* = 191, 16.4%)
** CT cutoff		
Without low SMM (*n* = 987, 84.9%)	841 (72.3)	146 (12.6)
With low SMM (*n* = 176, 15.1%)	131 (11.3)	45 (3.9)

Variables are expressed as *n* (%). * BIA cutoff indicates AWGS index (ASMI, ALM/height^2) of <7.0 kg/m^2 in men and <5.7 kg/m^2. ** CT cutoff indicates <1 SD, sex-specific mean value of the participants. SMM, skeletal muscle mass; CT, computed tomography; BIA, bioimpedance analysis; FNIH, The Foundation for the National Institutes of Health; ALM, appendicular lean mass; ASMI, appendicular skeletal mass index; BMI, body mass index; AWGS, Asian Working Group of Sarcopenia.

To identify the predictors of discordant results by CT and BIA, univariate analysis was performed, see Table 5. Older age (HR = 1.05), female sex (HR = 1.48), lower BMI (HR = 0.73), lower serum albumin level (HR = 0.58), and higher GGT (HR = 1.01) were significantly predictive of discordance between CT- and BIA-defined low SMM ($p < 0.05$). The results of analyses using the additional diagnostic criteria for low SMM are listed in Supplementary Table S4. The results of basic demographic characteristics, specificity and sensitivity of the low SMM defined by the BIA compared to the low SMM defined by the CT as diagnostic standard criteria, added to Supplementary Table S5.

Among selected independent predictors of the presence of discordance, age, female gender, and BMI were selected for multivariate analysis. Thus, we stratified our study population into two groups according to these three independent variables to check the prevalence of discordance, see Figure 2. Older age (>65 years) (22.3% vs. 12.2%), female gender (20.9% vs. 9.8%), and lower BMI (<25 kg/m^2) (20.1%% vs. 3.5%) had a significantly higher risk of discordance than the counterparts (all $p < 0.001$). The results of analysis using additional diagnostic criteria for low SMM are given in Supplementary Figure S3.

Table 5. Predictors of discordance between CT and BIA-based low SMM.

Variables	Discordance between CT and BIA-Based Low SMM		
	Univariate	Multivariate	
	p-Value	*p*-Value	OR (95% CI)
Demographic parameters			
Age, years	<0.001	<0.001	1.050 (1.035–1.069)
Female gender	<0.001	0.044	1.480 (1.012–2.303)
Body mass index, kg/m^2	<0.001	<0.001	0.725 (0.668–0.790)
Hypertension	0.420	-	-
Diabetes mellitus	0.219	-	-
Viral hepatitis	0.875	-	-
Laboratory parameters			
Fasting glucose, mg/dL	0.728	-	-
Aspartate aminotransferase, IU/L	0.734	-	-
Alanine aminotransferase, IU/L	0.092	0.021	0.977 (0.955–0.996)
Serum albumin, mg/dL	0.020	0.097	0.581 (0.264–1.117)
Total bilirubin, mg/dL	0.691	-	-
Gamma glutamyl-transpeptidase, IU/L	0.002	<0.001	1.009 (1.004–1.014)
Serum creatinine, mg/dL	0.803	-	-
Platelet count, 10^9/L	0.759	-	-
Prothrombin time, INR	0.084	0.153	3.580 (0.629–19.312)
Total cholesterol, mg/dL	0.277	-	-
Triglycerides, mg/dL	0.096	0.128	1.003 (0.999–1.005)
High-density lipoprotein cholesterol, mg/dL	0.096	0.280	1.011 (0.993–1.023)
Low-density lipoprotein cholesterol, mg/dL	0.172	-	-
HbA1c, %	0.696	-	-
Fibrosis-4 index	0.615	-	-

CT cutoff indicates <1 SD, sex-specific mean value of the participants. BIA cutoff indicates AWGS index (ASMI, ALM/height^2) of <7.0 kg/m^2 in men and <5.7 kg/m^2. SMM, skeletal muscle mass; CT, computed tomography; BIA, bioimpedance analysis; OR, odds ratio; CI, confidence interval; INR, international normalized ratio; FNIH, The Foundation for the National Institutes of Health; ALM, appendicular lean mass; ASMI, appendicular skeletal mass index; BMI, body mass index; AWGS, Asian Working Group of Sarcopenia.

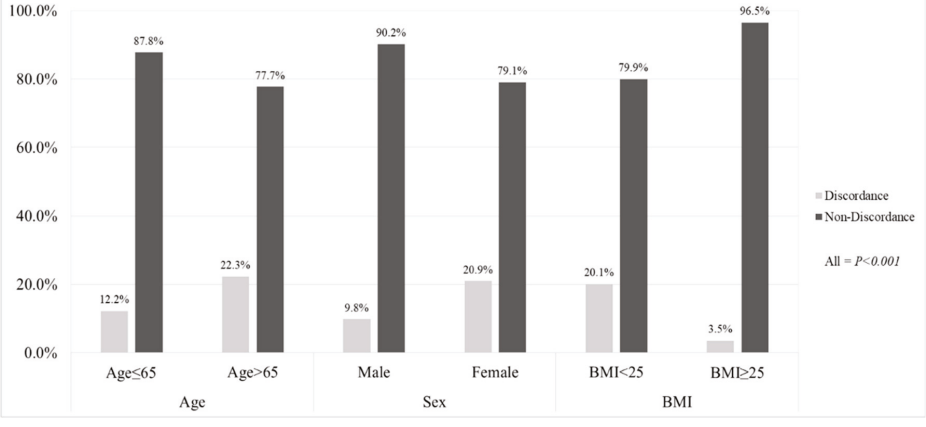


Figure 2. Percentage of subjects with non-discordance and those with discordance in diagnosing low SMM using CT and BIA according to identified independent predictors. Participants with age > 65 years, female gender and BMI < 25 kg/m^2 had a significantly higher proportion of discordance than the counterparts (all $p < 0.001$). CT cutoff indicates < 1 SD. BIA cutoff indicates AWGS index (ASMI, ALM/height^2) of <7.0 kg/m^2 in men and <5.7 kg/m^2. BMI, body mass index; ALM, appendicular lean mass; ASMI, appendicular skeletal mass index; AWGS, Asian Working Group of Sarcopenia.

4. Discussion

The diagnostic criteria for sarcopenia have not yet been definitively established, even though it is one of the most important public health concerns [31]. Varying diagnostic criteria for sarcopenia based on several assessment modalities, which include CT and BIA, are available [2,3,10], and the criteria are different between Asian and Western countries [6,7,32]. Ethnicity is an important factor for the diagnosis of sarcopenia [33]. Several recent research groups have published diagnostic guidelines for sarcopenia, which have emphasized the importance of ethnicity [34,35]. The BIA and CT diagnostic criteria differ according to ethnicity [3,6,7,35]. According to our knowledge, no comparison of BIA and SMM measured by CT at the L3 level has been performed. Therefore, our findings will facilitate the establishment of diagnostic cutoff values for Asian patients.

Our data show a significant association in crude and BMI-adjusted SMM assessed by CT and BIA, although the area assessed was different for each method. Similar to previous studies [3,11,32,36–39], the proportion of subjects with low SMM in our study varied from 15.1% to 16.4% when CT or BIA was used to assess SMM, and the risk factors for discordant results between the methodologies were advanced age, female gender, and lower BMI.

We believe the identified risk factors for discordant results can be explained in several ways: Total fat mass tends to be higher in older adults [40], and BIA can overestimate SMM when the subject has a high fat mass [41]; assessment of SMM using BIA can be overestimated in female subjects who have a higher probability of increased body fat [41]; and there is a weaker correlation between SMM in the limb and L3 area among subjects with a lower BMI [42,43]. All of these factors suggest that CT may be required for a more accurate assessment of SMM in subjects with advanced age, female gender, and low BMI.

Our study has several strengths. First, the overall sample size was over 1100, which ensures the statistical power and precision of our results. We adopted several cutoff values for CT and BIA when defining low SMM, and we found the three factors of age, gender, and BMI to be associated with discordance between CT- and BIA-based SMM assessments. Second, we focused on the general population instead of medically vulnerable subjects, such as only older adults, or those with liver cirrhosis or cancers for whom sarcopenia already showed clinical implications. Similar to our study, several recent studies proved the clinical significance of assessing sarcopenia in the general population and non-alcoholic fatty liver disease (NAFLD) subjects [44,45]. Thus, our study provides information that helps to identify optimal subjects for CT-based assessment of sarcopenia. Third, in contrast to most previous studies [7,11,27,46], we adopted several cutoff values for SMM assessed by CT and BIA. Although the predictors of discordance were not exactly the same, we obtained relatively consistent results regardless of the cutoff value used. Fourth, several studies have compared DXA and BIA, but few have directly compared CT and BIA to assess SMM [46,47]. In our study, SMM using CT and BIA was measured on the same day, in contrast to most previous studies [9,25,46]. As a result, any bias caused by different time points of SMM assessments may have been prevented. Lastly, because there are significant differences in SMM between Western and Asian populations, focusing on a single ethnicity may be important. Thus, the results of our study could be optimized for an Asian population.

Several issues remain unresolved in our study. First, although we adopted several known cutoff values for CT and BIA, the results of our study should be further validated based on existing diagnostic criteria for sarcopenia. Second, recent studies have insisted that other factors, such as muscle strength and walking speed, should be considered when diagnosing sarcopenia. However, our study was retrospectively performed based on the clinical information of the subjects who underwent a comprehensive medical health check-up, and we only used SMM to define sarcopenia. Further studies with additional markers of sarcopenia should validate our results. Third, our study only included subjects who were willing to receive and could afford a medical health check-up. The prevalence of hypertension (29.1%) and diabetes (11.3%) in the general Korean population in 2016 (Korean Center for Disease Control and Prevention; Ministry of Health and Welfare) [48], were lower than those in this study (41.0% and 26.4%, respectively). In Korea, individuals >40 years of age are eligible for

basic health check-ups; those with chronic diseases such as hypertension and diabetes receive health checkups more frequently. The mean age of our patients was 57 years, and 40.9% were >60 years of age. Because of this potential selection bias, our results may not be fully applicable to the general population, but this can be resolved in future studies. Fourth, SMM measured by BIA is affected by the hydration status [12]. The patients were admitted to the health check-up unit and fasted overnight. Thus, the hydration status of all of the patients should have been similar. Lastly, when discordant results between CT- and BIA-based SMM assessments were obtained, we did not know which diagnostic modality to accept. For patients with discordant CT and BIA results, it is important to decide which results should be used. However, a definitive diagnostic method for sarcopenia has not been established. This issue should be explored in future longitudinal follow-up studies that use solid end-points, such as mortality. This issue should be explored in future longitudinal follow-up studies that use solid end-points such as mortality, which might propose the right direction toward CT or BIA.

In conclusion, the significant association between CT and BIA for SMM assessment suggests that BIA could be used to assess sarcopenia in clinical practice. However, because advanced age, female gender, and low BMI were risk factors for discordant results between CT and BIA, BIA assessment should be interpreted cautiously in subjects with these risk factors and, if possible, CT or other modalities should be considered as an alternative diagnostic tool to assess SMM to define sarcopenia.

Supplementary Materials: The following are available online at http://www.mdpi.com/2077-0383/8/3/322/s1, Supplementary Figure S1: Flow chart depicting selection of the study population. Supplementary Figure S2: Distribution of skeletal muscle mass assessed by BIA according to quartile stratification of skeletal muscle mass assessed by CT. Supplementary Figure S3. Percentage of subjects with non-discordance and those with discordance in diagnosing low skeletal muscle mass assessed using CT and BIA according to identified independent predictors. Participants with Age>65 years (A), female gender (B) and BMI<25 kg/m^2 (C). Supplementary Table S1: Comparison between subjects with and without low skeletal muscle mass assessed by CT. Supplementary Table S2: Comparison between subjects with and without low skeletal muscle mass assessed by BIA. Supplementary Table S3: Distribution of subjects with and without low skeletal muscle mass according to CT and BIA. Supplementary Table S4: Predictors of discordance in defining low skeletal muscle mass assessed by CT and BIA Supplementary Table S5: Basic demographic characteristics, specificity and sensitivity of the low SMM defined by the BIA compared to the low SMM defined by the CT as diagnostic standard criteria. Supplementary Table S1: Linear regression results of comparing SMM evaluated by CT and BIA.

Author Contributions: Conceptualization, M.H.J. and S.U.K.; methodology, M.H.J., T.S.L., M.Y.J. and S.U.K.; participation in patient management and data collection; H.W.L., B.K.K., J.Y.P., D.Y.K., S.H.A., K.-H.H. and S.U.K.; contribution to the data acquisition, responsibility for writing the paper, and statistical analysis: M.H.J. and S.U.K.; All authors reviewed the paper and approved the final version.

Acknowledgments: This study was supported by the Liver Cirrhosis Clinical Research Center, in part by a grant from the Korea Healthcare technology R & D project, Ministry of Health and Welfare, Republic of Korea (no. HI10C2020). 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. Roubenoff, R. Origins and clinical relevance of sarcopenia. *Can. J. Appl. Physiol.* **2001**, *26*, 78–89. [CrossRef] [PubMed]
2. Cruz-Jentoft, A.J.; Baeyens, J.P.; Bauer, J.M.; Boirie, Y.; Cederholm, T.; Landi, F.; Martin, F.C.; Michel, J.P.; Rolland, Y.; Schneider, S.M.; et al. Sarcopenia: European consensus on definition and diagnosis: Report of the european working group on sarcopenia in older people. *Age Ageing* **2010**, *39*, 412–423. [CrossRef] [PubMed]
3. Janssen, I. The epidemiology of sarcopenia. *Clin. Geriatr. Med.* **2011**, *27*, 355–363. [CrossRef] [PubMed]
4. Cesari, M.; Pahor, M.; Lauretani, F.; Zamboni, V.; Bandinelli, S.; Bernabei, R.; Guralnik, J.M.; Ferrucci, L. Skeletal muscle and mortality results from the inchianti study. *J. Gerontol. A. Biol. Sci. Med. Sci.* **2009**, *64*, 377–384. [CrossRef] [PubMed]

5. Newman, A.B.; Kupelian, V.; Visser, M.; Simonsick, E.M.; Goodpaster, B.H.; Kritchevsky, S.B.; Tylavsky, F.A.; Rubin, S.M.; Harris, T.B. Strength, but not muscle mass, is associated with mortality in the health, aging and body composition study cohort. *J. Gerontol. A. Biol. Sci. Med. Sci.* **2006**, *61*, 72–77. [CrossRef] [PubMed]
6. Chen, L.K.; Liu, L.K.; Woo, J.; Assantachai, P.; Auyeung, T.W.; Bahyah, K.S.; Chou, M.Y.; Chen, L.Y.; Hsu, P.S.; Krairit, O.; et al. Sarcopenia in Asia: Consensus report of the asian working group for sarcopenia. *J. Am. Med. Dir. Assoc.* **2014**, *15*, 95–101. [CrossRef] [PubMed]
7. Studenski, S.A.; Peters, K.W.; Alley, D.E.; Cawthon, P.M.; McLean, R.R.; Harris, T.B.; Ferrucci, L.; Guralnik, J.M.; Fragala, M.S.; Kenny, A.M.; et al. The fnih sarcopenia project: Rationale, study description, conference recommendations, and final estimates. *J. Gerontol. A. Biol. Sci. Med. Sci.* **2014**, *69*, 547–558. [CrossRef] [PubMed]
8. Vellas, B.; Fielding, R.A.; Bens, C.; Bernabei, R.; Cawthon, P.M.; Cederholm, T.; Cruz-Jentoft, A.J.; Del Signore, S.; Donahue, S.; Morley, J.; et al. Implications of ICD-10 for sarcopenia clinical practice and clinical trials: Report by the international conference on frailty and sarcopenia research task force. *J. Frailty Aging* **2018**, *7*, 2–9. [PubMed]
9. Proctor, D.N.; O'Brien, P.C.; Atkinson, E.J.; Nair, K.S. Comparison of techniques to estimate total body skeletal muscle mass in people of different age groups. *Am. J. Physiol.* **1999**, *277*, 489–495. [CrossRef] [PubMed]
10. Lee, R.C.; Wang, Z.M.; Heymsfield, S.B. Skeletal muscle mass and aging: Regional and whole-body measurement methods. *Can. J. Appl. Physiol.* **2001**, *26*, 102–122. [CrossRef] [PubMed]
11. Chien, M.Y.; Huang, T.Y.; Wu, Y.T. Prevalence of sarcopenia estimated using a bioelectrical impedance analysis prediction equation in community-dwelling elderly people in taiwan. *J. Am. Geriatr. Soc.* **2008**, *56*, 1710–1715. [CrossRef] [PubMed]
12. Chumlea, W.C.; Guo, S.S. Bioelectrical impedance and body composition: Present status and future directions. *Nutr. Rev.* **1994**, *52*, 123–131. [CrossRef] [PubMed]
13. Batsis, J.A.; Mackenzie, T.A.; Barre, L.K.; Lopez-Jimenez, F.; Bartels, S.J. Sarcopenia, sarcopenic obesity and mortality in older adults: Results from the national health and nutrition examination survey III. *Eur. J. Clin. Nutr.* **2014**, *68*, 1001–1007. [CrossRef] [PubMed]
14. Janssen, I.; Heymsfield, S.B.; Baumgartner, R.N.; Ross, R. Estimation of skeletal muscle mass by bioelectrical impedance analysis. *J. Appl. Physiol.* **2000**, *89*, 465–471. [CrossRef] [PubMed]
15. Kyle, U.G.; Genton, L.; Hans, D.; Pichard, C. Validation of a bioelectrical impedance analysis equation to predict appendicular skeletal muscle mass (ASMM). *Clin. Nutr.* **2003**, *22*, 537–543. [CrossRef]
16. Sergi, G.; De Rui, M.; Veronese, N.; Bolzetta, F.; Berton, L.; Carraro, S.; Bano, G.; Coin, A.; Manzato, E.; Perissinotto, E. Assessing appendicular skeletal muscle mass with bioelectrical impedance analysis in free-living caucasian older adults. *Clin. Nutr.* **2015**, *34*, 667–673. [CrossRef] [PubMed]
17. Scafoglieri, A.; Clarys, J.P.; Bauer, J.M.; Verlaan, S.; Van Malderen, L.; Vantieghem, S.; Cederholm, T.; Sieber, C.C.; Mets, T.; Bautmans, I.; et al. Predicting appendicular lean and fat mass with bioelectrical impedance analysis in older adults with physical function decline - the provide study. *Clin. Nutr.* **2017**, *36*, 869–875. [CrossRef] [PubMed]
18. Karelis, A.D.; Chamberland, G.; Aubertin-Leheudre, M.; Duval, C.; Ecological mobility in Aging and Parkinson (EMAP) Group. Validation of a portable bioelectrical impedance analyzer for the assessment of body composition. *Appl. Physiol. Nutr. Metab.* **2013**, *38*, 27–32. [CrossRef] [PubMed]
19. Ling, C.H.; de Craen, A.J.; Slagboom, P.E.; Gunn, D.A.; Stokkel, M.P.; Westendorp, R.G.; Maier, A.B. Accuracy of direct segmental multi-frequency bioimpedance analysis in the assessment of total body and segmental body composition in middle-aged adult population. *Clin. Nutr.* **2011**, *30*, 610–615. [CrossRef] [PubMed]
20. Kim, H.; Kim, C.H.; Kim, D.W.; Park, M.; Park, H.S.; Min, S.S.; Han, S.H.; Yee, J.Y.; Chung, S.; Kim, C. External cross-validation of bioelectrical impedance analysis for the assessment of body composition in korean adults. *Nutr. Res. Pract.* **2011**, *5*, 246–252. [CrossRef] [PubMed]
21. Gibson, A.L.; Holmes, J.C.; Desautels, R.L.; Edmonds, L.B.; Nuudi, L. Ability of new octapolar bioimpedance spectroscopy analyzers to predict 4-component-model percentage body fat in hispanic, black, and white adults. *Am. J. Clin. Nutr.* **2008**, *87*, 332–338. [CrossRef] [PubMed]
22. Malavolti, M.; Mussi, C.; Poli, M.; Fantuzzi, A.L.; Salvioli, G.; Battistini, N.; Bedogni, G. Cross-calibration of eight-polar bioelectrical impedance analysis versus dual-energy x-ray absorptiometry for the assessment of total and appendicular body composition in healthy subjects aged 21–82 years. *Ann. Hum. Biol.* **2003**, *30*, 380–391. [CrossRef] [PubMed]

23. Ward, L.C. Segmental bioelectrical impedance analysis: An update. *Curr. Opin. Clin. Nutr. Metab. Care* **2012**, *15*, 424–429. [CrossRef] [PubMed]
24. Hanaoka, M.; Yasuno, M.; Ishiguro, M.; Yamauchi, S.; Kikuchi, A.; Tokura, M.; Ishikawa, T.; Nakatani, E.; Uetake, H. Morphologic change of the psoas muscle as a surrogate marker of sarcopenia and predictor of complications after colorectal cancer surgery. *Int. J. Colorectal Dis.* **2017**, *32*, 847–856. [CrossRef] [PubMed]
25. Mourtzakis, M.; Prado, C.M.; Lieffers, J.R.; Reiman, T.; McCargar, L.J.; Baracos, V.E. A practical and precise approach to quantification of body composition in cancer patients using computed tomography images acquired during routine care. *Appl. Physiol. Nutr. Metab.* **2008**, *33*, 997–1006. [CrossRef] [PubMed]
26. Vallet-Pichard, A.; Mallet, V.; Nalpas, B.; Verkarre, V.; Nalpas, A.; Dhalluin-Venier, V.; Fontaine, H.; Pol, S. Fib-4: An inexpensive and accurate marker of fibrosis in hcv infection. Comparison with liver biopsy and fibrotest. *Hepatology* **2007**, *46*, 32–36. [CrossRef] [PubMed]
27. Buckinx, F.; Reginster, J.Y.; Dardenne, N.; Croisiser, J.L.; Kaux, J.F.; Beaudart, C.; Slomian, J.; Bruyere, O. Concordance between muscle mass assessed by bioelectrical impedance analysis and by dual energy x-ray absorptiometry: A cross-sectional study. *BMC Musculoskelet Disord.* **2015**, *16*, 60. [CrossRef] [PubMed]
28. Forslund, A.H.; Johansson, A.G.; Sjodin, A.; Bryding, G.; Ljunghall, S.; Hambraeus, L. Evaluation of modified multicompartment models to calculate body composition in healthy males. *Am. J. Clin. Nutr.* **1996**, *63*, 856–862. [CrossRef] [PubMed]
29. Han, D.S.; Chang, K.V.; Li, C.M.; Lin, Y.H.; Kao, T.W.; Tsai, K.S.; Wang, T.G.; Yang, W.S. Skeletal muscle mass adjusted by height correlated better with muscular functions than that adjusted by body weight in defining sarcopenia. *Sci. Rep.* **2016**, *6*, 19457. [CrossRef] [PubMed]
30. Montano-Loza, A.J.; Meza-Junco, J.; Prado, C.M.; Lieffers, J.R.; Baracos, V.E.; Bain, V.G.; Sawyer, M.B. Muscle wasting is associated with mortality in patients with cirrhosis. *Clin. Gastroenterol. Hepatol.* **2012**, *10*, 166–173. [CrossRef] [PubMed]
31. Dasarathy, S.; Merli, M. Sarcopenia from mechanism to diagnosis and treatment in liver disease. *J. Hepatol.* **2016**, *65*, 1232–1244. [CrossRef] [PubMed]
32. Baumgartner, R.N.; Koehler, K.M.; Gallagher, D.; Romero, L.; Heymsfield, S.B.; Ross, R.R.; Garry, P.J.; Lindeman, R.D. Epidemiology of sarcopenia among the elderly in new mexico. *Am. J. Epidemiol.* **1998**, *147*, 755–763. [CrossRef] [PubMed]
33. Shafiee, G.; Keshtkar, A.; Soltani, A.; Ahadi, Z.; Larijani, B.; Heshmat, R. Prevalence of sarcopenia in the world: A systematic review and meta- analysis of general population studies. *J. Diabetes Metab. Disord* **2017**, *16*, 21. [CrossRef] [PubMed]
34. Dent, E.; Morley, J.E.; Cruz-Jentoft, A.J.; Arai, H.; Kritchevsky, S.B.; Guralnik, J.; Bauer, J.M.; Pahor, M.; Clark, B.C.; Cesari, M.; et al. International clinical practice guidelines for Sarcopenia (ICFSR): Screening, Diagnosis and Management. *J. Nutr. Health. Aging* **2018**, *22*, 1148–1161. [CrossRef] [PubMed]
35. Cruz-Jentoft, A.J.; Bahat, G.; Bauer, J.; Boirie, Y.; Bruyere, O.; Cederholm, T.; Cooper, C.; Landi, F.; Rolland, Y.; Sayer, A.A.; et al. Sarcopenia: Revised European consensus on definition and diagnosis. *Age Ageing* **2019**, *48*, 16–31. [CrossRef] [PubMed]
36. Kim, T.N.; Park, M.S.; Yang, S.J.; Yoo, H.J.; Kang, H.J.; Song, W.; Seo, J.A.; Kim, S.G.; Kim, N.H.; Baik, S.H.; et al. Prevalence and determinant factors of sarcopenia in patients with type 2 diabetes: The Korean Sarcopenic Obesity Study (KSOS). *Diabetes Care* **2010**, *33*, 1497–1499. [CrossRef] [PubMed]
37. Kim, T.N.; Yang, S.J.; Yoo, H.J.; Lim, K.I.; Kang, H.J.; Song, W.; Seo, J.A.; Kim, S.G.; Kim, N.H.; Baik, S.H.; et al. Prevalence of sarcopenia and sarcopenic obesity in Korean adults: The Korean sarcopenic obesity study. *Int. J. Obes.* **2009**, *33*, 885–892. [CrossRef] [PubMed]
38. Lau, E.M.; Lynn, H.S.; Woo, J.W.; Kwok, T.C.; Melton, L.J., 3rd. Prevalence of and risk factors for sarcopenia in elderly Chinese men and women. *J. Gerontol. A Biol. Sci. Med. Sci.* **2005**, *60*, 213–216. [CrossRef] [PubMed]
39. Lim, S.; Kim, J.H.; Yoon, J.W.; Kang, S.M.; Choi, S.H.; Park, Y.J.; Kim, K.W.; Lim, J.Y.; Park, K.S.; Jang, H.C. Sarcopenic obesity: Prevalence and association with metabolic syndrome in the Korean Longitudinal Study on Health and Aging (KLoSHA). *Diabetes Care* **2010**, *33*, 1652–1654. [CrossRef] [PubMed]
40. Baumgartner, R.N.; Stauber, P.M.; McHugh, D.; Koehler, K.M.; Garry, P.J. Cross-sectional age differences in body composition in persons 60+ years of age. *J. Gerontol. A Biol. Sci. Med. Sci.* **1995**, *50*, M307–M316. [CrossRef] [PubMed]

41. Lloret Linares, C.; Ciangura, C.; Bouillot, J.L.; Coupaye, M.; Decleves, X.; Poitou, C.; Basdevant, A.; Oppert, J.M. Validity of leg-to-leg bioelectrical impedance analysis to estimate body fat in obesity. *Obes. Surg.* **2011**, *21*, 917–923. [CrossRef] [PubMed]
42. Pomeroy, E.; Macintosh, A.; Wells, J.C.K.; Cole, T.J.; Stock, J.T. Relationship between body mass, lean mass, fat mass, and limb bone cross-sectional geometry: Implications for estimating body mass and physique from the skeleton. *Am. J. Phys. Anthropol.* **2018**, *166*, 56–69. [CrossRef] [PubMed]
43. Bosy-Westphal, A.; Muller, M.J. Identification of skeletal muscle mass depletion across age and BMI groups in health and disease—There is need for a unified definition. *Int. J. Obes.* **2015**, *39*, 379–386. [CrossRef] [PubMed]
44. Lee, Y.H.; Jung, K.S.; Kim, S.U.; Yoon, H.J.; Yun, Y.J.; Lee, B.W.; Kang, E.S.; Han, K.H.; Lee, H.C.; Cha, B.S. Sarcopaenia is associated with NAFLD independently of obesity and insulin resistance: Nationwide surveys (KNHANES 2008-2011). *J. Hepatol.* **2015**, *63*, 486–493. [CrossRef] [PubMed]
45. Lee, Y.H.; Kim, S.U.; Song, K.; Park, J.Y.; Kim, D.Y.; Ahn, S.H.; Lee, B.W.; Kang, E.S.; Cha, B.S.; Han, K.H. Sarcopenia is associated with significant liver fibrosis independently of obesity and insulin resistance in nonalcoholic fatty liver disease: Nationwide surveys (KNHANES 2008–2011). *Hepatology* **2016**, *63*, 776–786. [CrossRef] [PubMed]
46. Lee, S.Y.; Ahn, S.; Kim, Y.J.; Ji, M.J.; Kim, K.M.; Choi, S.H.; Jang, H.C.; Lim, S. Comparison between dual-energy X-ray absorptiometry and bioelectrical impedance analyses for accuracy in measuring whole body muscle mass and appendicular skeletal muscle mass. *Nutrients* **2018**, *10*. [CrossRef] [PubMed]
47. Vermeiren, S.; Beckwee, D.; Vella-Azzopardi, R.; Beyer, I.; Knoop, V.; Jansen, B.; Delaere, A.; Antoine, A.; Bautmans, I.; Scafoglieri, A.; et al. Evaluation of appendicular lean mass using bio impedance in persons aged 80+: A new equation based on the BUTTERFLY-study. *Clin. Nutr.* **2018**. [CrossRef] [PubMed]
48. National Health Statistics. Available online: http://www.index.go.kr/potal/main/EachDtlPageDetail.do?idx_cd=1438 (accessed on 10 December 2018).

Journal of
Clinical Medicine

MDPI

Review

Abuse of Licit and Illicit Psychoactive Substances in the Workplace: Medical, Toxicological, and Forensic Aspects

Ricardo Jorge Dinis-Oliveira [1,2,3,*] and Teresa Magalhães [1,2,4,*]

1 IINFACTS—Institute of Research and Advanced Training in Health Sciences and Technologies, Department of Sciences, University Institute of Health Sciences (IUCS), CESPU, CRL, 4585-116 Gandra, Portugal
2 Department of Public Health and Forensic Sciences, and Medical Education, Faculty of Medicine, University of Porto, 4200-319 Porto, Portugal
3 UCIBIO-REQUIMTE, Laboratory of Toxicology, Department of Biological Sciences, Faculty of Pharmacy, University of Porto, 4050-313 Porto, Portugal
4 Fidelidade—Companhia de Seguros, SA, R. Direita de Campinas 324, 4100-207 Porto, Portugal
* Correspondence: ricardinis@med.up.pt (R.J.D.-O.); tmaga@med.up.pt (T.M.)

Received: 27 January 2020; Accepted: 10 March 2020; Published: 12 March 2020

Abstract: About one-third of adult life is spent in the workplace. The use of psychoactive substances is a major preventable cause of morbidity and mortality. The consumption of psychoactive substances during or outside working hours greatly increases the frequency and severity of labor accidents, as well as the workers' poor general state of health and productivity, implying higher costs for enterprises. It is the responsibility of organizations to ensure the safety and health of their workers. These cannot be limited to traditional routine clinical exams, as other aspects also have an impact on health. Thus, prevention and intervention in the consumption of psychoactive substances (e.g., ethanol, opioids, central nervous system stimulants or depressants, hallucinogens, *Cannabis* derivatives, dissociative substances, and inhalants) in labor activity should be considered as an investment of organizations and not as a cost, in view of the professional, personal, and family advantages for workers and employers, with a potential impact on productivity, security, health, and quality of life at work. Despite the extensive literature on the subject, each article generally focuses on one or another aspect of a very specific nature, not tackling the problem in a holistic way by confronting clinical, safety, and legal issues. This article presents a reflection on the legal, laboratorial, clinical, ethical, forensic, and safety concerns related to the consumption of psychoactive substances in the workplace, and can be a cross-cutting contribution to occupational medicine, forensic medicine, and insurance medicine, as well as for entrepreneurs, lawyers, judges, workers, and technicians from the public and private sectors that develop projects in this area. This discussion is based on general principles established internationally and highlights the role of the occupational healthcare system and other decision-making actors in the prevention and supervision of workplace psychoactive consumption.

Keywords: occupational medicine; forensic medicine; insurance medicine; psychoactive substances; safety; clinical; forensic; law; ethics

1. Introduction

Workplaces reflect, to a certain extent, the widespread presence of ethanol and other psychoactive substances in society and the type of work, but sometimes both. There has been global growth in the use of workplace drug testing as a response to drug-related risks to safety and productivity [1,2]. Depending on the countries, the labor sectors, and the professions, statistics suggest high percentages of workers suffer from or are at risk of becoming dependent on ethanol [3]. Civil construction workers,

transport workers, hotel staff, barmen, catering workers, farmers, and workers in the primary sectors are particularly affected by this problem, especially in terms of ethanol [3]. Additionally, ethanol-related indicators are often inversely proportional to the level of literacy, meaning that the higher this level is, the less likely workers are to become drunk [4]. However, in some countries, excessive ethanol or other substance consumption is clearly observed among some higher professional groups such as doctors and managers [5,6]. Regarding the consumption of other psychoactive substances in the workplace, the information is scarcer, with *Cannabis* being more prevalent among younger workers, whereas cocaine is more prevalent among highly skilled workers such as managers [7]. Indeed, because workplace drug testing is often performed on-site by occupational physicians, a global statistic spectrum is hard to obtain [8].

The consumption of psychoactive substances at work depends on the combination of multiple factors, some of them linked to individual characteristics and lifestyles, and others being of a professional nature, related for example to the typology of work, rhythms and cadences, shift work, and stress, among others [9,10].

It is obvious that the prevention and deterrence of problems associated with the use of psychoactive substances should be a global intervention involving the participation of all decision-making actors in the organization bodies, namely, occupational medicine, occupational safety and health, human resources, social action services, intermediate and direct leadership, workers' representatives, and workers themselves. Detection may or may not be part of the organization's health and safety policy procedures. In order to create a program that contemplates psychoactive substance testing, the underlying policy, objectives, and rights and responsibilities of all parties involved should be made explicit. Moreover, the analysis of results and the relationship with the worker must respect confidentiality, cooperation, mutual commitment, and capacity building. The final goals are beneficial both for employers and employees and include the increase of workplace productivity by reducing absenteeism, presenteeism, and workplace accidents, as well as raising awareness of the toxic effects of psychoactive substances and their consequences in workers' performance by adopting a more healthily behavior during and outside working hours [9,11,12].

This article presents a reflection on the legal, laboratorial, clinical, ethical, forensic, and safety concerns related to the consumption of psychoactive substances in the workplace, and can be a cross-cutting contribution to occupational medicine for businessmen, lawyers, juries, workers, and laboratory technicians that act in this area. We also aimed to discuss the best practices to be followed by laboratories providing workplace drug testing services.

2. Methods

A narrative review was performed by searching articles in English, French, Spanish, and Portuguese in PubMed, Scopus, Web of Science, and PsycINFO concerning legal, laboratorial, clinical, ethical, forensic, and safety concerns of workplace drug testing, without time limit. Besides these inclusion criteria, additional reports were obtained from the references of the articles identified in the original search. International reports from the World Health Organization; European Union; International Commission on Occupational Health; and legislation on psychoactive substance workplace consumption, testing, and prevention were also reviewed. Specifically, Portuguese legislation, under the European Union regulation, was used as a starting point and, regardless the jurisdiction, broad and transversal aspects general and useful for occupational medicine were discussed. The effects in workplace performance of ethanol, opioids, central nervous system stimulants or depressants, hallucinogens, *Cannabis* derivatives, dissociative substances, inhalants, and examples of new psychoactive substances were also reviewed. From this search, 192 documents were obtained and 105 were ultimately considered for the final version of the manuscript. The remaining documents were excluded if they focused on specific points of certain enterprises or on a very specific point of a country not considered useful for a broad discussion, or when those aspects do not fit well in the scope of occupational medicine, such as technical and analytical details of toxicological methods.

3. Psychoactive Substances and Occupational Risks

Psychoactive substances are those that act mainly on the central nervous system, where they alter brain function, resulting in temporary changes in perception, mood, consciousness, and behavior. The effect depends not only on the specific substance but also on the dose, previous consumption, tolerance, comorbidities, mixture with other substances, and route of administration, among other reasons. In a general overview, all psychoactive substances have, to a higher or lesser extent, a dysfunctional effect on work capability. Considering the clinical usefulness and their most important effects at the level of the central nervous systems at usual doses, psychoactive substances are classified into several groups, which are discussed below regarding their repercussions in terms of workplace performance.

Opioids are mainly used as analgesics and antitussives [13–17]. Physical work and high psychosocial work demands, excessive repetition of tasks, awkward postures, and heavy lifting are all known workplace risk factors for musculoskeletal pain and consequently lead to administration of opioids, according to prospective studies [18,19]. Interestingly, a recent randomized trial demonstrated that treatment with opioids was not better to treatment with nonopioid medications for improving pain-related function over 12 months in moderate to severe chronic back pain or hip or knee osteoarthritis pain [20]. Part of this group are natural compounds extracted from opium such as morphine, codeine, and thebaine; semisynthetic compounds such as heroin, oxycodone, hydrocodone, oxymorphone, etorphine, and hydromorphone; and the synthetics tramadol, tapentadol, meperidine (or pethidine), methadone, fentanyl, and pentazocine. Work-related injuries have been identified as a factor in the rise of opioid dependence and opioid-related overdoses [21,22]. Opioid use has also been associated with the risk of motor vehicle accidents in commercial drivers [23]. Indeed, several adverse effects are possible risk factors for workplace performance, namely, those mediated by μ-receptors, such as sedation, respiratory depression, suppression of cough reflex, sweating, euphoria, dysphoria, confusion, insomnia, agitation, fear, hallucinations, drowsiness, motor decoding and mood swings, miosis, and dependence [17,24,25].

Cocaine, amphetamines (i.e., *d*-amphetamine, *d*-methamphetamine, methylphenidate, 3,4-methylenedioxymethamphetamine), caffeine, nicotine, and ephedrine are **stimulants of the central nervous system** [26–28]. Xenobiotics belonging to this group exhibit, in general, a pronounced stimulating effect in the central nervous system. They reduce the feeling of mental (i.e., increased alertness) and physical fatigue (i.e., increased motor activity) and cause dependence [29,30]. Of these, the most consumed psychoactive substance is caffeine, which is present in coffee, tea, and chocolate, as well as in numerous foods, drugs, and beverages, as it increases energy and concentration, with no major negative labor effects having been reported [31], and it moreover has been linked to a reduced suicide risk [32]. In addition, considerable research suggests that nicotine enhances cognitive control-related processes (e.g., attention, memory) among nicotine-deprived smokers, both in terms of behavior and neural indices [33]. Indeed, smokers deprived of nicotine (e.g., 12 h smoking abstinence) exhibit reduced cognitive-attentional functioning [34]. Employees who tested positive for cocaine were four times more likely to be categorized as absentee employees and two times more likely to be terminated from employment than those who tested negative [35]. Comparing the consumers of amphetamines with those of other illicit psychoactive substances, it was found that they had higher absenteeism due to disease and/or accidents, as well as a higher incidence of behavioral risks. If consumption occurs outside working hours, the period of depression and asthenia and cognitive alterations may arise when they resume work activity. Low-to-moderate doses of stimulants (e.g., amphetamine, caffeine, modafinil) have been reported to be effective countermeasures for mood and performance decrements caused by sleep deprivation and fatigue [36,37].

More than 61 **cannabinoids** have been identified among more than 400 compounds documented in *Cannabis sativa* [16,38–40]. The most important forms are delta-9-tetrahydrocannabinol (Δ^9-THC; main psychoactive constituent), Δ^8-THC (almost as active as the previous one but at lower concentrations), cannabinol (low activity, but at high concentrations), and cannabidiol (not psychoactive, but in

high concentrations). Δ^9-THC is the most widely abused illicit psychoactive substance and the potency of the derivative (e.g., marijuana, hashish, hashish oil) depends on the percentage of this xenobiotic [16,39–41]. Δ^9-THC produces euphoria followed by relaxation, distortion in space and time, hallucinations (in higher doses than those normally found clinically), changes in short-term memory, motor incoordination, behavioral disinhibition, concentration and learning issues, decreased appetite, and in high doses paranoid psychosis. It possesses a low risk of psychological dependence and its abstinence is typically characterized by insomnia [42]. Marijuana cigarettes containing low-to-moderate Δ^9-THC concentrations can decrease some night shift-related performance and mood disruptions [43,44]. Occupational medicine additional concerns come from the use of medicinal *Cannabis* [45–47]. Workers who have been authorized to use *Cannabis* should be required to report any change in product, dose, frequency, and timing of use or route of administration, and an occupational physician trained and knowledgeable on the impact and evaluation of potentially impairing substances in the workplace should be included [47]. Moreover, new psychoactive substances, such as synthetic cannabinoids, are now an additional concern to deal with [48]. Indeed, these compounds also bind to cannabinoid receptors and abuse can cause anxiety; confusion; hypertension; psychosis; hallucinations; tachycardia; seizures; and, in severe cases, it can lead to death [49–51].

Hallucinogens, such as the diethylamide of lysergic acid (LSD), mescaline, psilocybin, and psilocin, also have a pronounced effect on workplace performance [52,53]. Hallucinogens intensify sensory experiences and lead to behavioral alterations, delusions, as well as emotional lability at the time of and after consumption, confusional states, and *flashbacks* (i.e., reviving experiences). Chronic consumption can lead to depression, violent behavior, anxiety, and alteration in the perception of time, but the risk of dependence is absent.

The **central nervous system depressant** group mainly includes psychotropic drugs, notably the illicit use of anxiolytics, sedatives, and hypnotics, such as benzodiazepines; barbiturates; and non-benzodiazepine hypnotics such as zolpidem and ethanol [54,55]. Of all substances, ethanol is the one with the highest negative impact on work [9,56]. Globally, ethanol is the world's number one risk factor for ill-health and premature death amongst the 25- to 59-year-old age group, the core of the working-age population [57]. Moreover, in the USA, ethanol-induced impairment directly affects an estimated 15% of the workforce and causes more than 22% of the deaths as a result of injuries at work [58]. The main labor consequences of ethanol consumption are slow reaction time, motor incoordination, decreased visual acuity, emotional lability, reduced concentration, lower intellectual ability, behavioral changes, attendance/punctuality problems, lower productivity, absenteeism, presenteeism, workplace injuries, and higher employee turnover [59]. Several studies have shown that workers who drink the most are more often absent from work due to drinking [60], and they more often report alcohol-related presenteeism and inefficiency at work due to alcohol use [61]. Ethanol is anxiolytic/disinhibiting and has a high potential for abuse. Withdrawal leads to sweating, nausea, tremor, insomnia, decreased appetite, restlessness, aggression, anxiety, and eventually hallucinations. It is often used in conjunction with other substances to enhance the overall effect. Major international organizations such as the World Health Organization, the Council of the European Union, and the International Labor Office advertise the need as priority to actualize policies and programs focused on the issue of ethanol and work.

Among psychoactive drugs, benzodiazepines are the most prescribed drugs, especially as anxiolytic, sedative, or hypnotic drugs, and less often as antiepileptics and/or muscle relaxants [55]. Benzodiazepines have largely replaced barbiturates as they are safer drugs with fewer enzyme-inducing effects, and thus less severe interactions, less severe withdrawal symptoms, and a broad therapeutic margin [62]. However, persistent and longer than recommended use and self-medication is a reality. Benzodiazepines are distinguished between anxiolytic and hypnotic [24,55]. This distinction is somewhat artificial because all are anxiolytic, and all can change sleep as long as certain doses are reached [55,63]. What distinguishes them, however, is that so-called hypnotic benzodiazepines are potent drugs that can, therefore, modify sleep conditions at relatively low doses, whereas so-called anxiolytic benzodiazepines are less potent, allowing for a "therapeutic window" to exist where anxiolytic

action can be obtained without significantly interfering with sleep [55]. All benzodiazepines may induce tolerance, dependence, and addiction, but to a lesser extent than barbiturates [62]. Short-acting benzodiazepines have the highest potential to induce dependence, and withdrawal syndrome may even occur when discontinuation is abrupt. The abuse of these drugs is generally higher in females and older workers. The toxic effects of benzodiazepines on quality of work are mostly related to the central nervous system depressant effects, particularly sleepiness, motor incoordination, and impaired thinking [64]. They have a significant impact on motor vehicle driving, workability, and interpersonal relationships, and thus their use should be taken into account in these circumstances [65]. They can also cause short-term memory impairment, that is, diminished ability to learn new information, leaving already learned information intact.

Among **dissociative drugs**, phencyclidine and ketamine were originally developed as dissociative general anesthetics, capable of promoting sensory loss and analgesia, amnesia, and paralysis, generating an intense sensation of dissociation of the environment but without real loss of consciousness and protective reflexes [66,67]. Only ketamine is still available in therapy due to the high frequency of delusions and hallucinations observed postoperatively with phencyclidine [67–70]. The most classic adverse effects of ketamine are delirium, hallucinations, tachycardia, mild respiratory depression, confusion, irrationality, violent or aggressive behavior, dizziness, ataxia, slurred speech, delayed reaction time, euphoria, altered body image, analgesia, amnesia, and coma [66,67].

4. Psychoactive Substances and Occupational Consequences

Employers, for their part, have a broad range of responsibilities, and it is the employer's obligation to ensure the individual and collective health of workers for the proper functioning of all work activities. Psychoactive substance abuse is a major preventable cause of morbidity and mortality and has direct impacts on workability [71]. Staying in the workplace under the influence of psychoactive substances depends on the combination of multiple factors, some linked to individual characteristics and lifestyles, and others of professional nature, such as work typology, rhythms and cadences, irregular working hours as occurs in shift work, stress, and psychological harassment, among others. The following are some of the consequences of the occupational psychoactive substance consumption and thus justify why employers aim for a "drug-free" workplace [72–74]: (i) increase in the rate of presenteeism (i.e., being present at workplace in an impaired state) and absenteeism that affects professional performance promoting errors, and hence the competitiveness and productivity of the enterprise, as well as the country's own wealth; (ii) creation of a negative image, leading to discrediting and despising of the organization; (iii) negative effect on the equipment integrity and therefore a potential cause of financial losses; (iv) negative effect on the workers' physical, psychosocial, and behavioral integrity; (v) risk of neglect and reduced decision-making and motor coordination, consequently leading to a higher number of errors and accidents and therefore costs (e.g., insurance premiums); (vi) workers being more often involved in conflict, violent behavior, and theft, and being more frequently the subject of complaint by coworkers who may also see their physical integrity or even their own lives affected as a result of lack of care or discernment, decreased alertness, or altered perceptions or judgements from others being under the influence of psychoactive substances; and (vii) workers tending to be non-punctual (arriving at work later and leaving early) than the rest of the working population, putting a greater strain on coworkers by introducing additional tasks that still need to be done.

5. Legal Aspects

Aiming at a legal interpretation of this theme, it is important to make clear what is an occupational accident and workplace. Although, these concepts may vary with the different legislation of each country or state, regardless the jurisdiction, these are broad concepts in most legislations. We used Portuguese legislation as a starting point for this approach.

The Portuguese Law no. 98/2009 of 4 September regulates the regime of compensation for occupational accidents and diseases, including occupational rehabilitation and reintegration, pursuant

to article 284 of the Labor Code, approved by Law no. 7/2009 of 12 February. In accordance with Article 8, an ***occupational accident*** is one that occurs at the place and time of work and that directly or indirectly results in bodily injury, functional disturbance, or illness resulting in reduced working or earning capacity, or death. Therefore, it is an event that has a professional factor for its occurrence, including acts of violence that occurred within the scope of the workplace concept. Its major differences from the concept of occupational disease, in which occupational factors are also determinant, include (i) very short time (at most a few minutes), usually sudden and unexpected (acute); (ii) easy identification of the cause (professional); and (iii) easy identification of the lesion.

The concept of the ***workplace*** means any place where a worker is or should be by virtue of his/her work and that are directly or indirectly under the employer's control, including road accidents during work-related activities. The term "working time beyond the normal working period" is defined in Portuguese law as the time that (i) precedes its commencement, including preparatory acts; (ii) follows, including related acts; (iii) resulted from normal or forced interruptions of work. Article 9 extends the concept, and also considers a workplace accident as one that occurs (i) on the way to or from the workplace in the routes normally used and considering the period of time usually spent by the worker in those routes: between any of their workplaces if they have more than one job (the destination employer being the responsible for the accident), between his/her habitual or occasional residence and the workplace or the place of payment or the place where they will receive any kind of assistance or treatment by reason of a previous accident, or between the workplace and the place of meal and between the new place defined by the employer for a specific work and the usual workplace or his/her habitual or occasional residence. It is also considered a work accident when the normal route has been interrupted or changed in order to satisfy workers' needs, as well as due to force majeure or due to a fortuitous event (ii) while performing spontaneous services that may result in economic gain for the employer; (iii) at and outside the workplace while representing workers; (iv) at the workplace, when attending a training course, or outside the workplace, if the express permission from the employer for such attendance was obtained; (v) at the place of payment of remuneration and while staying there for that purpose; (vi) at the place where the worker should receive any form of assistance or treatment due to a previous accident and while staying there for that purpose; (vii) while job searching during the hours granted by law to workers with process of termination of the current work contract; and (viii) off-site or working time when verified in the performance of services determined by the employer or consented by him/her.

5.1. Workplace Drug Testing

The legislation applied to workplace drug testing also has some differences among countries [75]. Indeed, a multinational company may not be able to implement the same procedure in all its offices. In this work, we focused upon and discussed general consentaneous aspects that need to be followed in the application of a program for workplace drug testing; the Portuguese legislation was used as a platform for critical reflection. This approach is even more interesting because by adopting a radical step, in July 2001 Portugal became the first country in the world to decriminalize the possession for consumption of all illicit substances (Law no. 30/2000 of 20 November). Rather than being arrested, those caught with a personal supply are obliged to undergo rehabilitation treatment.

Although there are no specific propositions referring to psychoactive substance testing in the workplace, the Portuguese law appoints a set of very strict general rules regarding the worker's health. Indeed, the human life (Article 24); the moral and physical integrity of persons (Article 25); the right to preserve the intimacy of private life (Article 26); the right to the protection of personal data and the use of computers (Article 35); and the rights of all workers (regardless of age, sex, race, citizenship, territory of origin, religion, or political or ideological convictions) to have a job and work under conditions of hygiene, safety, and health, as well as the rights to have assistance and fair rehabilitation when they are victims of workplace accidents or occupational disease (Article 59) are rights constitutionally enshrined in the Constitution of the Portuguese Republic. Moreover, in Europe, any regulation must

conform to the European Convention for the Protection of Human Rights and Fundamental Freedoms. The latter warranties for a person's right to privacy, which states that everyone has the right to his/her private and family life, his/her home, and his/her correspondence, and that public authorities must not interfere with the exercise of this right, except if the interests of national security, public safety, or the economic well-being of the country are at risk. On an international level, the matter might be covered by Universal Declaration of Human Rights (Article 12), which states that no one shall be subjected to arbitrary interference with his/her privacy.

The use of psychoactive substances in the workplace (or during working hours) is usually governed by health and safety laws that address the potential health and safety risks for themselves or co-workers [75–77]. Therefore, these laws usually legitimate the fact that the tests should be restricted to categories of workers whose activity may endanger their physical or third party integrity and make employers responsible for preventing psychoactive substance use in the workplace and impose them to carry out risk assessments and preventive measures [78]. The real question is how we can fulfil the purposes of the law, namely, the following Portuguese legislation: (i) Articles 15 (employer's general obligations), 16 (simultaneous or successive activities in the same workplace), and 17 (worker's obligations) of the legal regime promoting occupational safety and health (Law no. 102/2009 of 10 of September); (ii) Articles 19 (medical tests and examinations), 99 (internal company rules), 281 (general principles on safety and health at workplace), 282 (information, consultation, and training of workers), 283 (accidents at work and occupational diseases), and 284 (regulation of prevention and reparation) of the Labor Code (Law no. 7/2009 of 12 February); (iii) Decree-Law no. 4/2015 of 7 of January (Code of Administrative Procedure); and (iv) the General Data Protection Regulation 2016/679 of the European Parliament and of the Council of 27 of April of 2016. Besides these legislations, specific regulations of each enterprise should be also taking into account. The Portuguese State, through the Labor General Inspection, the Directorate-General of Health, and the National Protection Centre against Professional Risks, proceeds to inspections to see if the rules are followed.

5.1.1. The Legal Regime for the Promotion of Health and Safety at Work (Law no. 102/2009 of 10 September)

The legal regime for the promotion of health and safety at work (Law no. 102/2009 of 10 September) states that occupational safety and health must be based on a correct and permanent risk assessment and be developed according to principles, policies, standards, and programs, focusing on the promotion and monitoring of occupational health and the enhancement of technical and scientific research in the field of safety and health at the workplace, particularly the emergence of new risk factors (Article 5). In Portugal, the existence of in-house occupational safety and health services is mandatory in organizations with more than 400 workers or in those with more than 30 workers if they are exposed to high-risk activities. In cases where the enterprise does not have an internal occupational health and safety service, this must be assumed by the external entities that provide occupational health and safety services (Article 78).

There are several general obligations that the employer must continuously and permanently fulfil (Articles 15 and 18): (i) ensuring safety and health in the workplace, (ii) identifying all foreseeable risks in all activities of the organization, (iii) mitigating monotonous and repetitive work, and (iv) reducing psychosocial risks. These concerns should be balanced accordingly to the risks to which the workers are potentially exposed to and are recognized as causes for psychoactive substance use, which in turn is a risk factor for accidents and enhancer of work-related diseases. Nevertheless, it should be borne in mind that employers cannot legitimately invoke the obligation to carry out psychoactive substance screening tests to accomplish their duties of ensuring the health of workers (Article 108). Indeed, the obligation to perform clinical exams (i.e., for admission, periodic, or occasional) is duly specified in the legislation and its purpose is aimed at attesting and evaluating the physical and mental fitness of the worker to perform the activity, as well as possible repercussions. In other words, compulsory psychoactive substance testing undermines the rights, freedoms and personal guarantees enshrined in

the Constitution of the Portuguese Republic, namely, the right to personal integrity (Article 25) and the right to privacy reserve (Article 26). This means that there will be no justification for drug testing in all (or random) workers in an organization, but only for those whose job requires high skills or involves considerable risk to themselves or other workers and in all who show manifested and serious signs of being influenced by psychoactive substances [79]. Moreover, it should be mentioned that the existence of an internal regulation that considers a program for workplace drug testing cannot be in itself a just cause for dismissal because it is not provided by law and violates the principle of job security and the fundamental right of workers accordingly the Article 53 of the Constitution of the Portuguese Republic.

On the other hand, it is also the worker's obligation to comply with the occupational safety and health requirements provided by legislation and collective labor regulation instruments, as well as for instructions determined for that purpose by the employer (Article 17).

5.1.2. The Labor Code (Law no. 7/2009 of 12 February)

Regarding the Portuguese Labor Code (Law no. 7/2009 of 12 February), the article 281 et seq. focuses on the prevention and reparation of accidents and occupational diseases in the workplace; workers have the right to provide work in safe and healthy conditions and should respect the occupational safety and health requirements laid down by law or collective labor regulation instruments, or those determined by the employer, and the employer should apply the necessary measures to provide such environments.

The carrying out of medical tests and examinations (Article 19) is within the scope of the organization of occupational safety and health services and must respect citizens' rights, freedoms, and guarantees. Regarding the detection of psychoactive substance use, this may or may not be part of the organization's health and safety policy. The most common procedure is to perform drug testing in those workers randomly nominated by the computer, as well as those appointed by the occupational physician, or to those who request according to the rules of procedure. As mentioned above, random drug testing has been prone to controversy because employers must ensure that every aspect of their policies is rooted in scientific evidence, linked rationally to the goal of workplace safety, and are ethically justifiable [79]. Some organizations also advocate testing following an accident of specific consequences (e.g., fatalities, injuries that require anyone to be removed from the scene for medical care, damage to vehicles or property above a specified monetary amount) in order to determine whether the abuse of psychoactive substances were a factor. Of course, a positive test for psychoactive substances cannot prove that this was the cause of the accident.

Considering the possibility of creating a program that includes toxicological analyses for detection of psychoactive substances, this must comply with the legal rules in force and be part of an internal regulation according to Article 99. This regulation should spell out the underlying policy, objectives, and rights and responsibilities of all parties involved, as well as issues regarding the protection of personal data, namely, assuring the right of the worker to privacy, the need for his/her consent to perform the tests, and the preventive and non-sanctioning character of the drug tests. These are important topics specially designed to increase the sensitivity of health and safety professionals to this problem and encourage them to raise awareness of them.

Screening tests for psychoactive substance use are restricted to the occupational physician or, under his/her guidance and control, to other health professionals obliged to professional secrecy (e.g., occupational nurses) and trained to use the kits for toxicological analysis. The same is true for the results of the tests because this represents health information. Therefore, clinical data can only be known to health professionals from the occupational medicine team, who are subject to confidentiality, and any hypothetic witnesses nominated by the worker.

In order to comply with professional secrecy and to guarantee the confidentiality of information resulting from medical examinations, the occupational physician must record results with "generic terms", namely, the worker (i) is fit (i.e., does not consume), (ii) is "fit with restrictions", or (iii) is temporarily unable (i.e., consumes) to perform his/her duties. According to no. 2 of Article 17 (on the

protection of personal data) of the Labor Code, at no time should the physician report the test results to the employer. Only in this way is the adequacy of the preventive and deterrent measures ensured and it represents a very serious offence the disrespect of these directives. Accordingly to Article 195 of the Portuguese Penal Code, the disclosure of other people's secrets, known to someone on account of his/her state, job, employment, profession (e.g., physician), or arts, is punished by a term imprisonment of 1 year or fine until 240 days. Moreover, the physician also follows the Deontological Code of the Medical Association, which compels him/her to professional secrecy (Article 71).

It is also made clear by the Labor Code (Article 19), in addition to the situations provided in occupational safety and health legislation, that the employer cannot, for admission purposes or job maintenance, require the applicant or worker to perform or present tests or medical examinations of any nature (such as the results of psychoactive substance use tests), or prove physical or mental conditions and therefore health. An exception occurs when these examinations aim for the protection and safety of the worker or third parties, or when the demands inherent to the activity justify it. In these cases, the reasons must be provided in writing to the job seeker or worker. Therefore, conducting screening tests for psychoactive substance use will only be legitimate in exceptional cases where health, welfare, worker, employer, or third-party concerns are at stake. In cases of job admission, even if screening tests for psychoactive substance can be justifiably performed in candidates, the possibility of the job applicant stopping drug use several days before testing should be born in mind. However, the screening will no longer be legally acceptable on the principles of proportionality, appropriateness, and reasonableness when there are no objective grounds to perform them in order to assure safety to the workers, service users, or the wider community, or when, from this point of view, the risks are minimal.

It is also clear that a job applicant or worker who has provided personal information has the right to control his/her personal data, to be aware of its content and the intended purpose, and to ask for its correction and update.

6. Diagnosis of the Influence of Psychoactive Substances and Toxicological Analysis

It is fundamental to have the substances that will be monitored clearly defined in the rules of procedure. Suspicion of psychoactive substance consumption can be made at various levels, of a subjective and objective nature. The loss of productivity and decreased quality of work, lack of punctuality and absenteeism, indiscipline and inappropriate behaviors, and the increase in workplace accidents are warning signs that cannot be neglected. However, these signs of suspicion should be part of a broader clinical and laboratory evaluation under the responsibility of occupational medicine and, in some cases, with the contribution of insurance medicine as well as forensic medicine.

Occupational medicine is a medical specialty concerned with the maintenance of health in the workplace, including prevention and treatment of diseases and injuries. In other words, the aim of occupational medicine focuses on workers' health. Accordingly, the International Code of Ethics for Occupational Health Professionals published by the International Commission on Occupational Health (ICOH) states, "the aim of occupational health practice is to protect and promote workers' health, to sustain and improve their working capacity and ability, to contribute to the establishment and maintenance of a safe and healthy working environment for all, as well as to promote the adaptation of work to the capabilities of workers, taking into account their state of health" [80].

Many enterprises establish a drug policy with little or no structure for drug testing, namely, (i) quality control, (ii) systematic confirmation of positives, and (iii) procedures to accomplish the chain of custody. Indeed, procedures that ensure the chain of custody compliance are of utmost importance in toxicological analyses. Moreover, drug testing is usually performed on-site by occupational physicians, who are usually not familiarized with analytical toxicological aspects. In Europe, the European Guidelines for Workplace Drug Testing have been prepared and updated by the European Workplace Drug Testing Society (EWDTS) for different samples. These guidelines are designed to [81] (i) establish best practice and standard procedures whilst allowing individual countries to operate within the

requirements of national customs and legislation, (ii) ensure that the entire drug testing process is conducted to give accurate and reliable information about a donor's drug use, (iii) maintain the legal defensibility and scrutiny either by an employment tribunal or a court of law, (iv) protect the dignity of the specimen donors and the validity of the specimen, (v) define common and critical quality control procedures for laboratories, and vi) help in the interpretation of the analytical results. Guidelines for collection of biological samples for toxicological analysis were recently published [82] and have largely conformed to the EWDTS guidelines, focusing on sample collection and testing in urine, hair, and oral fluid. Our work focused only on specific points relevant to the interpretation of workplace drug testing [75,81]. Trained personnel, who do not need to specifically be healthcare professionals, are required [81]. All samples must be kept for an agreed period or in respect of the national legislation of each country to allow rebuttal if any judicial sues are made regarding the obtained results. After the agreed time, the laboratory may discard the sample if the customer did not request the laboratory to retain the sample for an additional period. Samples must be retained within the secure laboratory area until the disposal date agreed with the customer [83].

6.1. Interpretation of Ethanol Results

The alcoholaemia is usually performed by quantitative breath analysis, using an alcohol meter (duly calibrated). In Portugal, there is a model certified by the Portuguese Institute of Quality that is based on the theoretical relationship defined in Law no. 18/2007 of 17 May, assuming that 1 mg/L in breath alcohol concentration (BrAC) is equal to 2.3 g/L of blood alcohol concentration (BAC) [65]. Counterproof should always be available to confirm results and should be provided by a referenced toxicology laboratory and by using a different technique and chemical principle from that of the screen test in order to ensure reliability and accuracy. Real BAC can be quantified by blood collection and further analysis by using gas chromatography with a flame ionization detector and a headspace system. The direct determination of ethanol itself in hair is not possible due to its volatility and its potential absorption from external sources [54]. Instead, the minor ethanol metabolites ethyl glucuronide (EtG) and/or fatty acid ethyl esters (FAEE) can be measured in hair by GC or LC coupled to MS/MS as a direct alcohol consumption marker [84].

There are no legally defined values for ethanol blood concentration in the workplace and it should be noted that the maximum limit established for the Portuguese Highway Code should not be generalized to all professions or tasks [65]. Nevertheless, several Portuguese organizations are governed by the values defined in the Highway Code. For instance, the Collective Bargaining Agreement for the Construction and Public Works Industry that has been in force since April 1, 2010, in clause no. 78 (on the prevention and control of alcoholaemia) considers that it is under the influence of alcohol the worker who, under examination by BrAC, reveals 0.5 g/L or more. For workers under the Highway Code, the BAC provided in that Code is applicable.

6.2. Interpretation of Other Psychoactive Substances Results

Regarding the analysis of other psychoactive substances, toxicological analysis can be performed in various biological samples, being more commonly analyzed in oral fluid [83–85] and urine [81,83] or in hair [86] and nails [87], given that, although analytically possible, the toxicological results are not relevant to allow for a quantitative interpretation [65]. Indeed, in most cases, only by using blood, serum, or plasma is it possible to truly document the impairment. Therefore, the most correct procedure will be to consider as positive the test that reveals the presence of illicit psychoactive substances above the defined concentrations cut-offs for screening and confirmation tests [88], regardless of the quantitative interpretation the test may provide [89]. Although urine is an important sample for toxicological screening, it is only useful if freshly voided and collection is witnessed and supervised to avoid adulteration. Indeed, the collection facilities should be arranged to prevent adulteration of the specimen as much as possible and adding coloring agents to toilet water has been recommended to reduce the risk [81]. Some authors suggest that all urine specimens taken for drug testing from

both the workplace and court settings need to be tested for validity [90]. Nevertheless, urinating is a personal act and most people feel inhibited about being observed in such circumstances by close family members, medical staff, or even sexual partners, and the situation can never be less than humiliating. Furthermore, individuals may not want to disclose pregnancy or venereal diseases. A wide variety of collection devices are available in the market to collect oral fluid for toxicological analysis. A minimum validity test should be performed for oral fluid, such as visual inspection of the sample(s), measurement of oral fluid volume and testing on matrix authenticity through measurement of endogenous biomarkers such as salivary amylase and cortisol [83–85].

Hair samples have been gaining popularity. This is by far the most expensive method, but it is claimed to be more secure in the event of legal challenges, to provide conclusive evidence of rebuttal, and is an excellent indicator of addiction, as opposed to occasional use, which may be particularly suitable for pre-employment testing programs. Analysis of nail clippings may be a useful back-up for hair analysis when hair is unavailable [87,91,92]. Useful applications of nail analysis may be where contamination is highly unlikely (i.e., if the fingers do not become contaminated) as is the case during the consumption of tablets or capsules as medication or where the question of time or quantity of use rather than the fact of use is being investigated [92].

The most frequently analyzed substances are amphetamines and derivatives, cannabinoids, cocaine, opioids, and benzodiazepines. Regarding cocaine, immunoassays detect the metabolite benzoylecgonine quite accurately without much concern for false positive or negatives. Nevertheless, the metabolite is inactive and may be present up to 3 days after use, whereas most clinical effects of cocaine occur within 6 to 12 hours of use [26,27]. As for amphetamines, there are many false positives, including antihistamines, decongestants, antidepressants, or acid-blockers, and these may be detected 1 to 3 days after use. Indeed, methamphetamine only differs from pseudoephedrine in a single atom of oxygen. In fact, many methods flowed by clandestine laboratories use *l*-ephedrine (present in stimulant supplements and used for weight loss) or *d*-pseudoephedrine (present in several decongestants drugs) as starting reagents to produce *d*-methamphetamine [93].

Regarding cannabinoids, common over-the-counter analgesics such as ibuprofen and naproxen, as well as the increasing interest in the use of medical *Cannabis*, can increase the probability of cross-reaction with this assay, creating false positives. Many opioids can be missed during routine screening, generating significant false-negatives, and it has been reported that screening is also prone to many false-positive results in the presence of poppy seeds and quinolone antibiotics [13,94]. Indeed, poppy seed paste used in foods could lead to positive urine tests for opioids, even if increased cut-off levels are used and different biomarkers are considered for the differential diagnosis, such as the presence of thebaine to ensure justice for each individual [13,95]. Moreover, if an individual is currently taking a lawful medication containing codeine, the test can suggest that he/she has used heroin.

Nevertheless, it should be mentioned that there are several unscreened substances in existence, such as ketamine, chloral hydrate, gamma-hydroxybutyrate (GHB), psilocybin, mescaline, and cathinones. Moreover, qualitative or semi-quantitative immunoassays may cross-react with other substances, many of which are licit, namely, pharmaceutical drugs [89]. In this case, tests showing the presence of prescription drugs should not be considered positive.

On the basis of the technical limitation of the immunoassay analysis, whenever positive results (i.e., above a predefined cut-off level) emerge at the screening stage, confirmatory tests on the sample must be carried out by a referenced toxicology laboratory, normally by gas and liquid chromatography coupled with mass spectrometry. If the screen results are all negative, no further analysis is necessary. Results of confirmatory analysis are usually not readily available, and even if positive results are obtained, interpretation is required by skilled toxicological professionals in conjunction with a qualified occupation medicine physician [96].

Additionally, of note is that there are many other clinical situations with signs and symptoms similar to psychoactive substance poisoning, and thus definitive diagnosis is only possible through clinical examination. Therefore, a differential diagnosis is needed to find out whether the reason for the

positive result can be explained by prescribed medication; some other acceptable reason; or whether it is, in fact, a question of drug abuse [97]. If an occupational physician is not present, the worker should be removed from activities that put his/her or others' safety at risk and should be properly referred to a health unit. It is underlined that in this case all labor rights, including remuneration, must be ensured for as long as the worker is away from work.

6.3. Worker's Refusal to Undergo Toxicological Analysis

Besides the health professional (or other) who is competent to carry out the sample collection for toxicological analysis, it should be made possible for the worker to request the assistance of a witness, having a defined time to do so, but the test cannot be stopped if the presentation of the witness was not feasible. Refusal to submit to screening tests for psychoactive substance use, although possible, does not imply that the worker is unsuitable for work. When unjustified, the worker may incur any disciplinary offence and, in some enterprises, the refusal leads to the assumption that the worker has a BAC of 0.5 g/L or higher and therefore the employee is unfit/inapt for work. Moreover, it should be highlighted that if a worker is declared unfit by the occupational medicine as a result of psychoactive substance use, it is not a just cause for dismissal or to affect career progression, as clinical results are confidential and as such cannot be made known to the hierarchical superiors or other entities outside the clinical sphere. Only the behavior that may result from the influenced state can be framed according to Article 351 (just cause for dismissal) of the Portuguese Labor Code and as such be subject to disciplinary sanction (and possibly dismissal) due to the unacceptable or breach of established rules.

As a precautionary measure, if a BAC of 0.5 g/L or greater is assumed or presumed, or if positive illicit psychoactive substances results are obtained, the worker will be immediately prevented from working during the remaining daily work period, with the consequent loss of remuneration for such period.

If the worker is found to abuse psychoactive substances in the workplace, it is possible and permissible to initiate disciplinary proceedings against the worker, especially in organizations with internal rules prohibiting the use of psychoactive substances in the workplace.

7. Conclusions and Future Perspectives

The use of psychoactive substances in the workplace is a public health problem that influences the safety and health of workers and extends beyond the workplace itself. It is a contentious issue that has moved up the human resources agenda in recent years, and different organizations take diverse approaches to psychoactive substance workplace testing, some of them following a zero-tolerance approach, whereas others have had to develop a more nuanced policy. Such cases may also have insurance and legal repercussions [98,99]. Indeed, workplace psychoactive substance testing is one of the latest components to be added to the discipline of forensic toxicology, which now comprises the triad fields of post-mortem, human performance, and workplace drug testing toxicology [88]. There are several studies that demonstrate that the psychic and motor disturbance due to the consumption of psychoactive substances is in the origin of several work accidents and related economic impacts [12,100,101]. Moreover, testing for psychoactive substances in the workplace, at random or by surprise, has a statistically significant preventive effect in overall professions [102]. Nevertheless, testing in the workplace is a complex topic because it is not often directly regulated by supranational or national law. Only a few countries report legislation that clearly and specifically address the issue of drug testing in the workplace [51].

The participation of workers and their representatives in the design of an occupational health promotion plan and in the definition of policies to be followed undoubtedly plays a decisive role in the implementation of prevention programs against the use of psychoactive substances in the workplace. Human resources policies aimed at promoting worker safety, health, and well-being that integrate worker assistance programs, information campaigns, and other interventions in this field reflect the level of organizational cultures that incorporate concepts and principles of corporate citizenship,

encouraging entrepreneurs and managers to good practices, the production of deontological ethics, and codes that value the image of the company or organization and its end products. Workers should also be aware that the employer is committed to creating a work environment that promotes safety and health and that policies and related measures should be applied to all elements of the organizational system so that the individual and collective rights of society will be fully accomplished. Indeed, if on the one hand the use of illicit drugs in the workplace raises issues pertaining to prevention and safety and the responsibility of the various members of staff, on the other hand it also brings into question the interface between work and private life [103].

As a prerequisite to testing, the company must have an antidrug program aiming to ensure health and safety, including a written drug testing internal regulation. In this regulation, problems related to the use of psychoactive substances should be considered as health problems and consequently should be treated like other health problems in the workplace regarding temporary disability, sick pay, and other social benefits, especially during periods of rehabilitation treatment.

It is important to clarify that under the law, the screening, counterproof, or confirmatory analysis should not be charged to the worker because they are preventive health and safety activities of companies and the costs must always be supported by the employer. All these measures are stimulating and represent facilitating factors for recovery, although under treatment, the employer must also ensure that the workstation is maintained or transferred to other duties that do not pose a risk to the safety of himself or others without loss of rights or other benefits. Nevertheless, an employee's failure to successfully complete treatment requirements can eventually result in the termination of their employment.

Nowadays, we are witnessing a paradigm shift from an evolution centered on the treatment and rehabilitation of proven psychoactive substance dependents to a cross-cutting approach through the implementation of a specific regulation that includes (i) prevention of consumption through information actions on health consequences of the use of psychoactive substances; (ii) early detection; (iii) treatment at the enterprise (if it has the means), or through referral to other specialist physician, or to local or regional specialist services (e.g., integrated response centers or alcohololology units) to recover workers in compliance with personal freedom, and these treatments cannot be imposed by coercion; and (iv) socio-professional rehabilitation of the worker, aiming at relapse prevention.

For the full implementation of these objectives, the availability of skilled human resources specifically in occupational medicine with toxicological background and skilled toxicologists in clinical, forensic, and analytical aspects of psychoactive substances, to correctly interpret toxicological results, are needed. Given this reality, fulfilling the purposes of the law, namely, Article 19 of the Portuguese Labor Code, is an almost impossible task. There should be a balance between prevention and repression.

Finally, policies and educational interventions to reduce psychoactive substance consumption in the workplace are needed [104]. These should address awareness to workplace hazards, physician education to promote best practices for treatment, and overdose prevention. In addition, the phenomenon of the new psychoactive substances are also motifs of concern in the workplace that should not be disregarded [51]. Indeed, the consumption of new psychoactive substances poses a significant risk to public health and a challenge to national and international drug policies for occupational medicine, and only very recently have laboratories begun to offer screening and confirmation analysis for new psychoactive substances at lower costs [105].

Author Contributions: R.J.D.-O. created and designed the study, wrote the first draft of the manuscript, and analyzed and interpreted the literature. T.M. also analyzed and interpreted literature and wrote and revised the final version of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

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

References

1. George, S. A snapshot of workplace drug testing in the UK. *Occup. Med. (Lond.)* **2005**, *55*, 69–71. [CrossRef] [PubMed]
2. Burke, R.J.; Richardsen, A.M. *Increasing Occupational Health and Safety in Workplaces: Individual, Workplaces*; Edward Elger Publishing: Cheltenham, UK, 2019.
3. World Health Organization. *Global Status Report on Alcohol and Health*; World Health Organization: Geneva, Switzerland, 2018.
4. Mackenbach, J.P.; Stirbu, I.; Roskam, A.-J.R.; Schaap, M.M.; Menvielle, G.; Leinsalu, M.; Kunst, A.E.; European Union Working Group on Socioeconomic Inequalities in Health. Socioeconomic inequalities in health in 22 European countries. *N. Engl. J. Med.* **2008**, *358*, 2468–2481. [CrossRef] [PubMed]
5. Cope, R.B. The potential for occupational exposure of veterinarians to ketamine resulting in positive drug tests. *Aust. Vet. J.* **2018**, *96*, 59–61. [CrossRef] [PubMed]
6. Airagnes, G.; Lemogne, C.; Renuy, A.; Goldberg, M.; Hoertel, N.; Roquelaure, Y.; Limosin, F.; Zins, M. Prevalence of prescribed benzodiazepine long-term use in the French general population according to sociodemographic and clinical factors: Findings from the CONSTANCES cohort. *BMC Public Health* **2019**, *19*, 566. [CrossRef] [PubMed]
7. Corral, A.; Durán, J.; Isusi, I. *Use of Alcohol and Drugs at the Workplace*; European Foundation for the Improvement of Living and Working Conditions: Dublin, Ireland, 2012.
8. Lillsunde, P.; Mukala, K.; Partinen, R.; Lamberg, M. Role of occupational health services in workplace drug testing. *Forensic Sci. Int.* **2008**, *174*, 103–106. [CrossRef] [PubMed]
9. Buvik, K.; Moan, I.S.; Halkjelsvik, T. Alcohol-related absence and presenteeism: Beyond productivity loss. *Int. J. Drug Policy* **2018**, *58*, 71–77. [CrossRef] [PubMed]
10. Ervasti, J.; Kivimäki, M.; Head, J.; Goldberg, M.; Airagnes, G.; Pentti, J.; Oksanen, T.; Salo, P.; Suominen, S.; Jokela, M.; et al. Sociodemographic differences between alcohol use and sickness absence: Pooled analysis of four cohort studies. *Alcohol Alcohol. (Oxf. Oxfs.)* **2018**, *53*, 95–103. [CrossRef]
11. Dams, J.; Buchholz, A.; Kraus, L.; Reimer, J.; Scherbaum, N.; Konnopka, A.; König, H.-H. Excess costs of alcohol-dependent patients in German psychiatric care compared with matched non-alcohol-dependent individuals from the general population: A secondary analysis of two datasets. *BMJ Open* **2018**, *8*, e020563. [CrossRef]
12. Sullivan, T.; Edgar, F.; McAndrew, I. The hidden costs of employee drinking: A quantitative analysis. *Drug Alcohol Rev.* **2019**, *38*, 543–553. [CrossRef]
13. Dinis-Oliveira, R.J. Metabolism and metabolomics of opiates: A long way of forensic implications to unravel. *J. Forensic Leg. Med.* **2019**, *61*, 128–140. [CrossRef]
14. Faria, J.; Barbosa, J.; Moreira, R.; Queirós, O.; Carvalho, F.; Dinis-Oliveira, R.J. Comparative pharmacology and toxicology of tramadol and tapentadol. *Eur. J. Pain (Lond. Engl.)* **2018**, *22*, 827–844. [CrossRef] [PubMed]
15. Ventura, L.; Carvalho, F.; Dinis-Oliveira, R.J. Opioids in the frame of new psychoactive substances network: A complex pharmacological and toxicological issue. *Curr. Mol. Pharmacol.* **2018**, *11*, 97–108. [CrossRef] [PubMed]
16. Dinis-Oliveira, R.J. Licit and ilicit uses of medicines. *Acta Med. Port.* **2014**, *27*, 755–766. [CrossRef] [PubMed]
17. Dinis-Oliveira, R.J.; Carvalho, F.; Moreira, R.; Duarte, J.A.; Proenca, J.B.; Santos, A.; Magalhaes, T. Clinical and forensic signs related to opioids abuse. *Curr. Drug Abus. Rev.* **2012**, *5*, 273–290. [CrossRef] [PubMed]
18. da Costa, B.R.; Vieira, E.R. Risk factors for work-related musculoskeletal disorders: A systematic review of recent longitudinal studies. *Am. J. Ind. Med.* **2010**, *53*, 285–323. [CrossRef]
19. Buttorff, C.; Trujillo, A.J.; Castillo, R.; Vecino-Ortiz, A.I.; Anderson, G.F. The impact of practice guidelines on opioid utilization for injured workers. *Am. J. Ind. Med.* **2017**, *60*, 1023–1030. [CrossRef]
20. Krebs, E.E.; Gravely, A.; Nugent, S.; Jensen, A.C.; DeRonne, B.; Goldsmith, E.S.; Kroenke, K.; Bair, M.J.; Noorbaloochi, S. Effect of opioid vs nonopioid medications on pain-related function in patients with chronic back pain or hip or knee osteoarthritis pain: The SPACE randomized clinical trial. *JAMA* **2018**, *319*, 872–882. [CrossRef]
21. Kowalski-McGraw, M.; Green-McKenzie, J.; Pandalai, S.P.; Schulte, P.A. Characterizing the interrelationships of prescription opioid and benzodiazepine drugs with worker health and workplace hazards. *J. Occup. Environ. Med.* **2017**, *59*, 1114–1126. [CrossRef]

22. Cheng, M.; Sauer, B.; Johnson, E.; Porucznik, C.; Hegmann, K. Comparison of opioid-related deaths by work-related injury. *Am. J. Ind. Med.* **2013**, *56*, 308–316. [CrossRef]
23. Howard, M.E.; Desai, A.V.; Grunstein, R.R.; Hukins, C.; Armstrong, J.G.; Joffe, D.; Swann, P.; Campbell, D.A.; Pierce, R.J. Sleepiness, sleep-disordered breathing, and accident risk factors in commercial vehicle drivers. *Am. J. Respir. Crit. Care Med.* **2004**, *170*, 1014–1021. [CrossRef]
24. Katzung, B.G. *Basic & Clinical Pharmacology*, 14th ed.; Katzung, B.G., Trevor, A.J., Eds.; McGraw-Hill: New York, NY, USA, 2018.
25. Negus, S.S.; Banks, M.L. Medications development for opioid abuse. *Cold Spring Harb. Perspect. Med.* **2013**, *3*, a012104. [CrossRef] [PubMed]
26. Dinis-Oliveira, R.J. Metabolomics of cocaine: Implications in toxicity. *Toxicol. Mech. Methods* **2015**, *25*, 494–500. [PubMed]
27. Oliveira, N.G.; Dinis-Oliveira, R.J. Drugs of abuse from a different toxicological perspective: An updated review of cocaine genotoxicity. *Arch. Toxicol.* **2018**, *92*, 2987–3006. [CrossRef] [PubMed]
28. Nóbrega, L.; Dinis-Oliveira, R.J. The synthetic cathinone α-pyrrolidinovalerophenone (α-PVP): Pharmacokinetic and pharmacodynamic clinical and forensic aspects. *Drug Metab. Rev.* **2018**, *50*, 125–139. [CrossRef] [PubMed]
29. Carvalho, M.; Carmo, H.; Costa, V.M.; Capela, J.P.; Pontes, H.; Remiao, F.; Carvalho, F.; Bastos Mde, L. Toxicity of amphetamines: An update. *Arch. Toxicol.* **2012**, *86*, 1167–1231. [CrossRef] [PubMed]
30. Lhermitte, M.; Frimat, P.; Labat, L.; Haguenoer, J.M. Use of illicit substances in the workplace. *Ann. Pharm. Fr.* **2012**, *70*, 3–14. [CrossRef]
31. McLellan, T.M.; Caldwell, J.A.; Lieberman, H.R. A review of caffeine's effects on cognitive, physical and occupational performance. *Neurosci. Biobehav. Rev.* **2016**, *71*, 294–312. [CrossRef]
32. Park, H.; Suh, B.S.; Lee, K. Relationship between daily coffee intake and suicidal ideation. *J. Affect. Disord.* **2019**, *256*, 468–472. [CrossRef]
33. Sutton, S.K.; Van Rensburg, K.J.; Jentink, K.G.; Drobes, D.J.; Evans, D.E. Nicotine-induced cortical activation among nonsmokers with moderation by trait cognitive control. *Psychopharmacology* **2016**, *233*, 2301–2308. [CrossRef]
34. Leventhal, A.M.; Waters, A.J.; Boyd, S.; Moolchan, E.T.; Lerman, C.; Pickworth, W.B. Gender differences in acute tobacco withdrawal: Effects on subjective, cognitive, and physiological measures. *Exp. Clin. Psychopharmacol.* **2007**, *15*, 21–36. [CrossRef]
35. Normand, J.; Salyards, S.D.; Mahoney, J.J. An evaluation of preemployment drug testing. *J. Appl. Psychol.* **1990**, *75*, 629–639. [CrossRef] [PubMed]
36. Hart, C.L.; Haney, M.; Vosburg, S.K.; Comer, S.D.; Gunderson, E.; Foltin, R.W. Modafinil attenuates disruptions in cognitive performance during simulated night-shift work. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* **2006**, *31*, 1526–1536. [CrossRef] [PubMed]
37. Hart, C.L.; Ward, A.S.; Haney, M.; Nasser, J.; Foltin, R.W. Methamphetamine attenuates disruptions in performance and mood during simulated night-shift work. *Psychopharmacology* **2003**, *169*, 42–51. [CrossRef] [PubMed]
38. Ross, S.A.; ElSohly, M.A.; Sultana, G.N.; Mehmedic, Z.; Hossain, C.F.; Chandra, S. Flavonoid glycosides and cannabinoids from the pollen of *Cannabis sativa* L. *Phytochem. Anal.* **2005**, *16*, 45–48. [CrossRef] [PubMed]
39. Dinis-Oliveira, R.J. Metabolomics of Delta9-tetrahydrocannabinol: Implications in toxicity. *Drug Metab. Rev.* **2016**, *48*, 80–87. [CrossRef] [PubMed]
40. Dinis-Oliveira, R.J. The Clinical Toxicology Perspective on the Therapeutic Use of Cannabis and Cannabinoids. *Acta Med. Port.* **2019**, *32*, 87–90. [CrossRef]
41. Sharma, P.; Murthy, P.; Bharath, M.M. Chemistry, metabolism, and toxicology of cannabis: Clinical implications. *Iran. J. Psychiatry* **2012**, *7*, 149–156.
42. Crane, N.A.; Schuster, R.M.; Fusar-Poli, P.; Gonzalez, R. Effects of cannabis on neurocognitive functioning: Recent advances, neurodevelopmental influences, and sex differences. *Neuropsychol. Rev.* **2013**, *23*, 117–137. [CrossRef]
43. Keith, D.R.; Gunderson, E.W.; Haney, M.; Foltin, R.W.; Hart, C.L. Smoked marijuana attenuates performance and mood disruptions during simulated night shift work. *Drug Alcohol Depend.* **2017**, *178*, 534–543. [CrossRef]
44. Dahlgren, M.K.; Sagar, K.A.; Smith, R.T.; Lambros, A.M.; Kuppe, M.K.; Gruber, S.A. Recreational cannabis use impairs driving performance in the absence of acute intoxication. *Drug Alcohol Depend.* **2020**. [CrossRef]

45. Anderson, D.M.; Rees, D.I.; Tekin, E. Medical marijuana laws and workplace fatalities in the United States. *Int. J. Drug Policy* **2018**, *60*, 33–39. [CrossRef] [PubMed]
46. Phillips, J.A.; Holland, M.G.; Baldwin, D.D.; Meuleveld, L.G.; Mueller, K.L.; Perkison, B.; Upfal, M.; Dreger, M. Marijuana in the workplace: Guidance for occupational health professionals and employers: Joint Guidance Statement of the American Association of Occupational Health Nurses and the American College of Occupational and Environmental Medicine. *J. Occup. Environ. Med.* **2015**, *57*, 459–475. [CrossRef] [PubMed]
47. Goldsmith, R.S.; Targino, M.C.; Fanciullo, G.J.; Martin, D.W.; Hartenbaum, N.P.; White, J.M.; Franklin, P. Medical marijuana in the workplace: Challenges and management options for occupational physicians. *J. Occup. Environ. Med.* **2015**, *57*, 518–525. [CrossRef] [PubMed]
48. Cho, B.; Cho, H.S.; Kim, J.; Sim, J.; Seol, I.; Baeck, S.K.; In, S.; Shin, D.H.; Kim, E. Simultaneous determination of synthetic cannabinoids and their metabolites in human hair using LC-MS/MS and application to human hair. *Forensic Sci. Int.* **2020**, *306*, 110058. [CrossRef]
49. Adamowicz, P.; Gieron, J.; Gil, D.; Lechowicz, W.; Skulska, A.; Tokarczyk, B. The effects of synthetic cannabinoid UR-144 on the human body-A review of 39 cases. *Forensic Sci. Int.* **2017**, *273*, e18–e21. [CrossRef]
50. Gounder, K.; Dunuwille, J.; Dunne, J.; Lee, J.; Silbert, P.; Lawn, N. The other side of the leaf: Seizures associated with synthetic cannabinoid use. *Epilepsy Behav.* **2020**, *104*, 106901. [CrossRef]
51. Salomone, A.; Palamar, J.J.; Vincenti, M. Should NPS be included in workplace drug testing? *Drug Test. Anal.* **2019**. [CrossRef]
52. Dinis-Oliveira, R.J.; Pereira, C.L.; da Silva, D.D. Pharmacokinetic and pharmacodynamic aspects of peyote and mescaline: Clinical and forensic repercussions. *Curr. Mol. Pharmacol.* **2019**, *12*, 184–194. [CrossRef]
53. Dinis-Oliveira, R.J. Metabolism of psilocybin and psilocin: Clinical and forensic toxicological relevance. *Drug Metab. Rev.* **2017**, *49*, 84–91. [CrossRef]
54. Dinis-Oliveira, R.J. Oxidative and Non-Oxidative Metabolomics of Ethanol. *Curr. Drug Metab.* **2016**, *17*, 327–335. [CrossRef]
55. Dinis-Oliveira, R.J. Metabolic profile of oxazepam and related benzodiazepines: Clinical and forensic aspects. *Drug Metab. Rev.* **2017**, *49*, 451–463. [CrossRef] [PubMed]
56. Gomez-Recasens, M.; Alfaro-Barrio, S.; Tarro, L.; Llaurado, E.; Sola, R. A workplace intervention to reduce alcohol and drug consumption: A nonrandomized single-group study. *BMC Public Health* **2018**, *18*, 1281. [CrossRef] [PubMed]
57. Griswold, M.; Fullman, N.; Hawley, C.; Arian, N.; Zimsen, S.R.M.; Tymeson, H.D.; Venkateswaran, V.; Tapp, A.D.; Forouzanfar, M.H.; Abate, K.H.; et al. Alcohol use and burden for 195 countries and territories, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *Lancet* **2018**, *392*, 1015–1035. [CrossRef]
58. Spicer, R.S.; Miller, T.R.; Smith, G.S. Worker substance use, workplace problems and the risk of occupational injury: A matched case-control study. *J. Stud. Alcohol* **2003**, *64*, 570–578. [CrossRef]
59. Bragazzi, N.L.; Dini, G.; Toletone, A.; Rahmani, A.; Montecucco, A.; Massa, E.; Manca, A.; Guglielmi, O.; Garbarino, S.; Debarbieri, N.; et al. Patterns of harmful alcohol consumption among truck drivers: Implications for occupational health and work safety from a systematic review and meta-analysis. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1121. [CrossRef]
60. Edvardsen, H.M.; Moan, I.S.; Christophersen, A.S.; Gjerde, H. Use of alcohol and drugs by employees in selected business areas in Norway: A study using oral fluid testing and questionnaires. *J. Occup. Med. Toxicol.* **2015**, *10*, 46. [CrossRef]
61. Aas, R.W.; Haveraaen, L.; Sagvaag, H.; Thorrisen, M.M. The influence of alcohol consumption on sickness presenteeism and impaired daily activities. The WIRUS screening study. *PLoS ONE* **2017**, *12*, e0186503. [CrossRef]
62. Brunton, L.L.; Knollman, B.C.; Hilal-Dandan, R. *Goodman & Gilman's—The Pharmacological Basis of Therapeutics*, 13th ed.; McGraw-Hill—Medical Publishing Division: New York, NY, USA, 2017.
63. Dinis-Oliveira, R.J. Metabolic profile of flunitrazepam: Clinical and forensic toxicological aspects. *Drug Metab. Lett.* **2017**, *11*, 14–20. [CrossRef]
64. Hunt, D.L.; Artuso, R.D.; Kalia, N.; Leung, N.; Bernacki, E.J.; Tao, X.G. Association of opioid, anti-depressant, and benzodiazepines with workers' compensation cost: A cohort study. *J. Occup. Environ. Med.* **2019**, *61*, e206–e211. [CrossRef]

65. Dinis-Oliveira, R.J.; Nunes, R.; Carvalho, F.; Santos, A.; Teixeira, H.; Vieira, D.N.; Magalhaes, T. Ethical, technical and legal procedures of the medical doctor responsibility to accomplish the road enforcement law about driving under the influence of alcohol and psychotropic substances. *Acta Med. Port.* **2010**, *23*, 1059–1082.
66. Morgan, C.J.; Curran, H.V. Ketamine use: A review. *Addiction* **2012**, *107*, 27–38. [CrossRef] [PubMed]
67. Dinis-Oliveira, R.J. Metabolism and metabolomics of ketamine: A toxicological approach. *Forensic Sci. Res.* **2017**, *2*, 2–10. [CrossRef] [PubMed]
68. Orhurhu, V.J.; Claus, L.E.; Cohen, S.P. Ketamine toxicity. In *StatPearls*; StatPearls Publishing: Treasure Island, FL, USA, 2019.
69. Corriger, A.; Pickering, G. Ketamine and depression: A narrative review. *Drug Des. Dev. Ther.* **2019**, *13*, 3051–3067. [CrossRef] [PubMed]
70. Kintz, P.; Villain, M.; Dumestre, V.; Cirimele, V. Evidence of addiction by anesthesiologists as documented by hair analysis. *Forensic Sci. Int.* **2005**, *153*, 81–84. [CrossRef] [PubMed]
71. Cook, R.F.; Back, A.; Trudeau, J. Substance abuse prevention in the workplace: Recent findings and an expanded conceptual model. *J. Prim. Prev.* **1996**, *16*, 319–339. [CrossRef]
72. White, T. Drug testing at work: Issues and perspectives. *Subst. Use Misuse* **2003**, *38*, 1891–1902. [CrossRef] [PubMed]
73. Frone, M.R.; Trinidad, J.R. Perceived physical availability of alcohol at work and workplace alcohol use and impairment: Testing a structural model. *Psychol. Addict. Behav.* **2014**, *28*, 1271–1277. [CrossRef]
74. Takeshita, T. Prevention of alcohol-related health harm in the workplace. *Ind. Health* **2017**, *55*, 403–405. [CrossRef]
75. Pierce, A. Regulatory aspects of workplace drug testing in Europe. *Drug Test. Anal.* **2012**, *4*, 62–65. [CrossRef]
76. Rosso, G.L.; Montomoli, C.; Morini, L.; Candura, S.M. Seven years of workplace drug testing in Italy: A systematic review and meta-analysis. *Drug Test. Anal.* **2017**, *9*, 844–852. [CrossRef]
77. Ruehle, L. How to maintain a drug-free workplace and ensure audit success. *Occup. Health Saf.* **2016**, *85*, 76, 78, 80. [PubMed]
78. Waehrer, G.M.; Miller, T.R.; Hendrie, D.; Galvin, D.M. Employee assistance programs, drug testing, and workplace injury. *J. Saf. Res.* **2016**, *57*, 53–60. [CrossRef] [PubMed]
79. Christie, T. A discussion of the ethical implications of random drug testing in the workplace. *Healthc. Manag. Forum* **2015**, *28*, 172–174. [CrossRef] [PubMed]
80. International Commission on Occupational Health. *International Code of Ethics for Occupational Health Professionals*, 3rd ed.; International Commission on Occupational Health: Rome, Italy, 2014.
81. Taskinen, S.; Beck, O.; Bosch, T.; Brcak, M.; Carmichael, D.; Fucci, N.; George, C.; Piper, M.; Salomone, A.; Schielen, W.; et al. European guidelines for workplace drug testing in urine. *Drug Test. Anal.* **2017**, *9*, 853–865. [CrossRef] [PubMed]
82. Dinis-Oliveira, R.J.; Vieira, D.N.; Magalhaes, T. Guidelines for collection of biological samples for clinical and forensic toxicological analysis. *Forensic Sci. Res.* **2016**, *1*, 42–51. [CrossRef]
83. Brcak, M.; Beck, O.; Bosch, T.; Carmichael, D.; Fucci, N.; George, C.; Piper, M.; Salomone, A.; Schielen, W.; Steinmeyer, S.; et al. European guidelines for workplace drug testing in oral fluid. *Drug Test. Anal.* **2018**, *10*, 402–415. [CrossRef]
84. Moore, C. Oral fluid for workplace drug testing: Laboratory implementation. *Drug Test. Anal.* **2012**, *4*, 89–93. [CrossRef]
85. Desrosiers, N.A.; Huestis, M.A. Oral fluid drug testing: Analytical approaches, issues and interpretation of results. *J. Anal. Toxicol.* **2019**, *43*, 415–443. [CrossRef]
86. Salomone, A.; Tsanaclis, L.; Agius, R.; Kintz, P.; Baumgartner, M.R. European guidelines for workplace drug and alcohol testing in hair. *Drug Test. Anal* **2016**, *8*, 996–1004. [CrossRef]
87. Stowe, G.N.; Paulsen, R.B.; Hill, V.A.; Schaffer, M.I. A retrospective analysis of selected opioids in hair of workplace drug testing subjects. *J. Anal. Toxicol.* **2019**, *43*, 553–563. [CrossRef]
88. Karch, S.B. *Workplace Drug Testing*; CRC Press: Boca Raton, FL, USA, 2008.
89. Dinis-Oliveira, R.J. Heterogeneous and homogeneous immunoassays for drug analysis. *Bioanalysis* **2014**, *6*, 2877–2896. [CrossRef] [PubMed]
90. Lin, S.Y.; Lee, H.H.; Lee, J.F.; Chen, B.H. Urine specimen validity test for drug abuse testing in workplace and court settings. *J. Food Drug Anal.* **2018**, *26*, 380–384. [CrossRef] [PubMed]

91. Solimini, R.; Minutillo, A.; Kyriakou, C.; Pichini, S.; Pacifici, R.; Busardo, F.P. Nails in forensic toxicology: An update. *Curr. Pharm. Des.* **2017**, *23*, 5468–5479. [CrossRef] [PubMed]
92. Hill, V.A.; Stowe, G.N.; Paulsen, R.B.; Schaffer, M. Nail analysis for drugs: A role in workplace testing? *J. Anal. Toxicol.* **2018**, *42*, 425–436. [CrossRef] [PubMed]
93. Brzeczko, A.W.; Leech, R.; Stark, J.G. The advent of a new pseudoephedrine product to combat methamphetamine abuse. *Am. J. Drug Alcohol Abus.* **2013**, *39*, 284–290. [CrossRef]
94. Phan, H.M.; Yoshizuka, K.; Murry, D.J.; Perry, P.J. Drug testing in the workplace. *Pharmacotherapy* **2012**, *32*, 649–656. [CrossRef]
95. Ozbunar, E.; Aydogdu, M.; Doger, R.; Bostanci, H.I.; Koruyucu, M.; Akgur, S.A. Morphine Concentrations in Human Urine Following Poppy Seed Paste Consumption. *Forensic Sci. Int.* **2019**, *295*, 121–127. [CrossRef]
96. Dinis-Oliveira, R.J.; Carvalho, F.; Duarte, J.A.; Remiao, F.; Marques, A.; Santos, A.; Magalhaes, T. Collection of biological samples in forensic toxicology. *Toxicol. Mech. Methods* **2010**, *20*, 363–414. [CrossRef]
97. Verstraete, A.G.; Pierce, A. Workplace drug testing in Europe. *Forensic Sci. Int.* **2001**, *121*, 2–6. [CrossRef]
98. Smith, D.E.; Glatt, W.; Tucker, D.E.; Deutsch, R.; Seymour, R.B. Drug testing in the workplace: Integrating medical review officer duties into occupational medicine. *Occup. Med.* **2002**, *17*, 79–90.
99. Clark, H.W.; Johnson, B. Legal issues facing physicians: Substance abuse in the workplace. *Occup. Med.* **2002**, *17*, 119–135. [PubMed]
100. Ompad, D.C.; Gershon, R.R.; Sandh, S.; Acosta, P.; Palamar, J.J. Construction trade and extraction workers: A population at high risk for drug use in the United States, 2005–2014. *Drug Alcohol Depend.* **2019**, *205*, 107640. [CrossRef]
101. Carter, T.; Heaton, K.; Shirey, M.R. Should nursing institute random drug screening? *Workplace Health Saf.* **2019**, *67*, 604. [CrossRef]
102. Marques, P.H.; Jesus, V.; Olea, S.A.; Vairinhos, V.; Jacinto, C. The effect of alcohol and drug testing at the workplace on individual's occupational accident risk. *Saf. Sci.* **2014**, *68*, 108–120. [CrossRef]
103. Fantoni-Quinton, S.; Bossu, B.; Morgenroth, T.; Frimat, P. Illicit drugs, testing, prevention and work in France: Ethical and legal issues. *Med. Law* **2010**, *29*, 353–372. [PubMed]
104. Hawkins, D.; Roelofs, C.; Laing, J.; Davis, L. Opioid-related overdose deaths by industry and occupation-Massachusetts, 2011–2015. *Am. J. Ind. Med.* **2019**, *62*, 815–825. [CrossRef]
105. Kimble, A.N.; DeCaprio, A.P. Systematic analysis of novel psychoactive substances. II. Development of a screening/confirmatory LC-QqQ-MS/MS method for 800+ compounds and metabolites in urine. *Forensic Chem.* **2019**, *16*, 100189. [CrossRef]

Journal of
Clinical Medicine

MDPI

Review

Role of Nutrition and Exercise Programs in Reducing Blood Pressure: A Systematic Review

Roman Jurik and Petr Stastny *

Department of Sport, Faculty of Physical Education and Sport, Charles University, 162 52 Prague, Czech Republic; rjurik@centrum.cz

* Correspondence: stastny@ftvs.cuni.cz; Tel.: +420-777-198764

Received: 30 July 2019; Accepted: 3 September 2019; Published: 5 September 2019

Abstract: The combined effect of diet and strength training (ST) on blood pressure (BP) seems to be very important for the treatment of prehypertension and hypertension (HT). Therefore, the aim of this study was to determine whether ST alone or combined with nutrition or supplementation has an impact on the arterial pressure reduction in normotensive and hypertensive populations. A systematic computerized literature search was performed according to the PRISMA guidelines using PubMed, Scopus and Google Scholar; only English language studies published from 1999 until 2018 were included. This systematic search identified the results of 303 individuals from nine studies. The ST program alone had a similar effectiveness as the nutrition program (NP) alone; however, their combination did not result in increased effectiveness in terms of a high BP reduction. The consumption of L-citrulline had a similar effect as ST on lowering BP; on the other hand, caffeine led to an increase in BP during the ST session. Our data suggest that a combination of ST 2–3 times a week at moderate intensity and a NP seems to be equally effective in terms of lowering BP (systolic and diastolic) as ST and NP alone.

Keywords: resistance training; hypertension; arterial pressure; disease prevention; caffeine

1. Introduction

The nonpharmacological approach to hypertension (HT) reduction is based on lifestyle changes using nutritional and exercise strategies; different training interventions [1–7] or nutritional plans [8–10] have been shown to decrease arterial pressure values. Previous reviews have summarized the approaches of aerobic training [11–13], anaerobic training [14–17] and nutrition [18–21], and those strategies were effective for HT reduction with expected decreases of 5 mmHg in systolic blood pressure (SBP) and 3 mmHg in diastolic pressure (DBP) [22] after three months. However, nonpharmacological strategies do not contraindicate each other, and their combination has been shown to be effective for other health improvements, such as bodyweight reduction [10,23–27]. Therefore, there is a question as to what kind of intervention or their combination has an improved effect on the HT decrease.

One of the well-documented interventions that has been shown to reduce arterial pressure is strength training (ST), which has already been reviewed to set optimal training loads, such as the number of sets [28–31], repetitions [32–35] and rest intervals during training sessions [32–34,36–38]. One of the ST effects is eliciting high muscle protein degradation followed by protein synthesis, which increases basal metabolism and is therefore usually accompanied by changes in nutritional requirements. In clinical practice, it is typical that ST is prescribed along with a low carb diet [39–41], specific protein intake [42–49] or another strategy that might support and magnify the arterial pressure decrease. However, there is currently no recommendation regarding whether the ST program should be accompanied by specific nutritional support that could result in a greater effect of HT reduction.

Numerous studies have shown that different nutritional programs (NP) and strategies might lead to nonpharmacological decreases in arterial pressure [50–55], while many strategies place a high

demand on patients to make changes in their eating habits. One way to easily change food intake might be accomplished by using food supplements, which have a synergistic effect with the application of ST. Some food supplements have been shown to decrease arterial pressure when used with aerobic training [56–59] or alone. On the other hand, some food supplements might have a positive effect alone but when used simultaneously with exercise might cause side effects, such as post-exercise hypotension [60–66].

In the current literature, there is no current overview of whether ST or nutritional strategies have a stronger effect on HT reduction. Therefore, the aim of this study was to determine whether ST alone or ST combined with nutrition or a supplement has the greatest impact on arterial pressure reduction in normotensive and hypertensive populations. Additionally, this review aimed to summarize what kind of combined ST and NP might be effective for HT reduction without side effects. The main hypothesis of this article is that the biggest effect on HT should be observed when strength training is combined with energy intake restriction. Based on this result, practitioners can establish nonpharmacological treatment for the individuals with increased blood pressure (BP).

2. Methods

This systematic review is reported in accordance with recommendations as presented in the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement [67]. The protocol for this systematic review was published on PROSPERO under registration number CRD42019130631.

2.1. Search Strategy

A systematic computerized literature search was performed using PubMed, and Scopus and included studies published in English from 1966 until November 2018. The search formula included the following terms: Blood pressure AND hypertension OR cardiovascular disease OR hypotension AND resistance training OR strength training OR weightlifting OR bodybuilding OR exercise. The search was limited by article types, species, subjects, language, age, and text availability. A manual search was performed using the identified reviews, reference lists of the selected articles and Google Scholar.

2.2. Types of Studies

The review considered cohort studies, analytical cross-sectional studies, randomized control trials, nonrandomized control trials, intervention studies, case-control studies and others that included BP and HR measurements as well as data on NP and ST in all adult populations. The review studies were used for manual searches of their reference list. Dissertations, theses, conference proceedings, conference monographies, and other reviews were not included. Retrospective studies were not included because the area of interest requires performing experiments. The qualitative component also considered the type of ST and NP and methodological designs. All titles and abstracts were screened according to the above-mentioned inclusion criteria after removing the duplicates. Full texts of eligible articles were retrieved and assessed by two reviewers (R. J. and P.S.). Any discrepancies between the two reviewers were managed by a consensus discussion.

2.3. Types of Outcomes

The review considered studies that included the following outcome measures: Acute SBP and DBP variability before, during, and after exercise or a nutritional treatment; the HR variability before, during, and after exercise or treatment; and a mean arterial pressure before, during, and after exercise. The exclusion criteria were as follows: Full text was not available in English, the study did not contain an appropriate description of measurement devices and procedures, the study did not include a proper exercise or nutrition task, or the study did not report how raw data were processed.

2.4. Data Extraction and Evaluation

Data extraction included aspects of the study population, such as the average age and sex, specific aspects of the NP and ST intervention (sample size, type of exercise performed, presence of supervision, frequency, and, duration of each session, type of diet, type of supplements), outcome measures and results presented, and the values of BP or HR (Tables S1–S3); however, the studies were not rejected if any part were missing. The Physiotherapy Evidence Database (PEDro) scale was used to assess the methodological quality of a study based on general criteria, such as concealed allocation, intention-to-treat analysis, and adequacy of follow-up. These characteristics make the PEDro scale a useful tool to assess methodological quality [68]. Extraction was performed by two reviewers (R. J. and P.S.). The lack of clarity during the extraction process was resolved by the reviewer's discussion. The PEDro scale based on a Delphi list [69] was used for all articles even if the trials had already been rated by trained evaluators of the PEDro database (http://www.pedro.fhs.usyd.edu.au/).

3. Results

3.1. Study Characteristics

The systematic literature search through database searching identified 15,302 records. However, after duplicates were removed, 11,558 records were screened based on the title and abstract. The title and abstract screening resulted in 144 records for full-text eligibility. From these records, nine studies satisfied the quality and exclusion criteria and were selected for this systematic review after full-text screening (Figure 1).

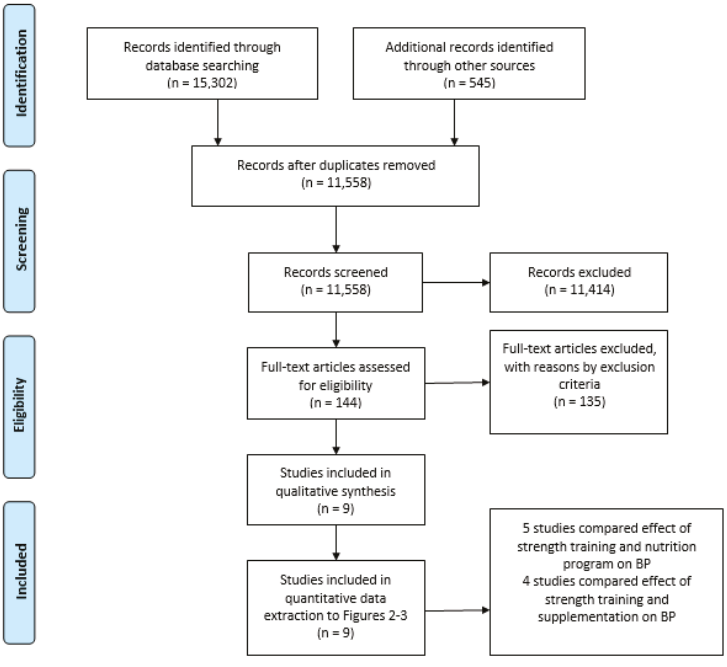


Figure 1. Flowchart of the selection process. BP = blood pressure.

Table 1 provides a general description of each study, sample, and intervention characteristics. Only two studies included seniors [70] and [71]. One study included only men [72], and three included a mixed sample of men and women [70,71,73]. Five studies included only women [74–78]. In total,

five studies compared the effect of ST with the NP intervention effect on BP [70,73–76]. Another three studies compared the effect of ST and supplementation intervention on BP [71,72,78,79] and one study analysed the effect of ST and supplementation sesion in cross sectional design [71].

Based on the study results compilation, it can be stated that ST alone has a positive effect on the BP values (Figure 2A,B). However, the effect did not depend on the type of ST (resistance training, bodyweight training or training on vibration platforms, etc. Table 2). An important prerequisite for effective ST is the selection of suitable parameters and methods; otherwise, there is a risk of injury or side effects. All studies used training parameters (Table 2) in accordance with recommendations established by The American College of Sport Medicine (ACSM) and the Canadian Hypertension Education Program, according to Pescatello et al. [80]. However, each study chose different training parameters based on the performance and health state of their subjects.

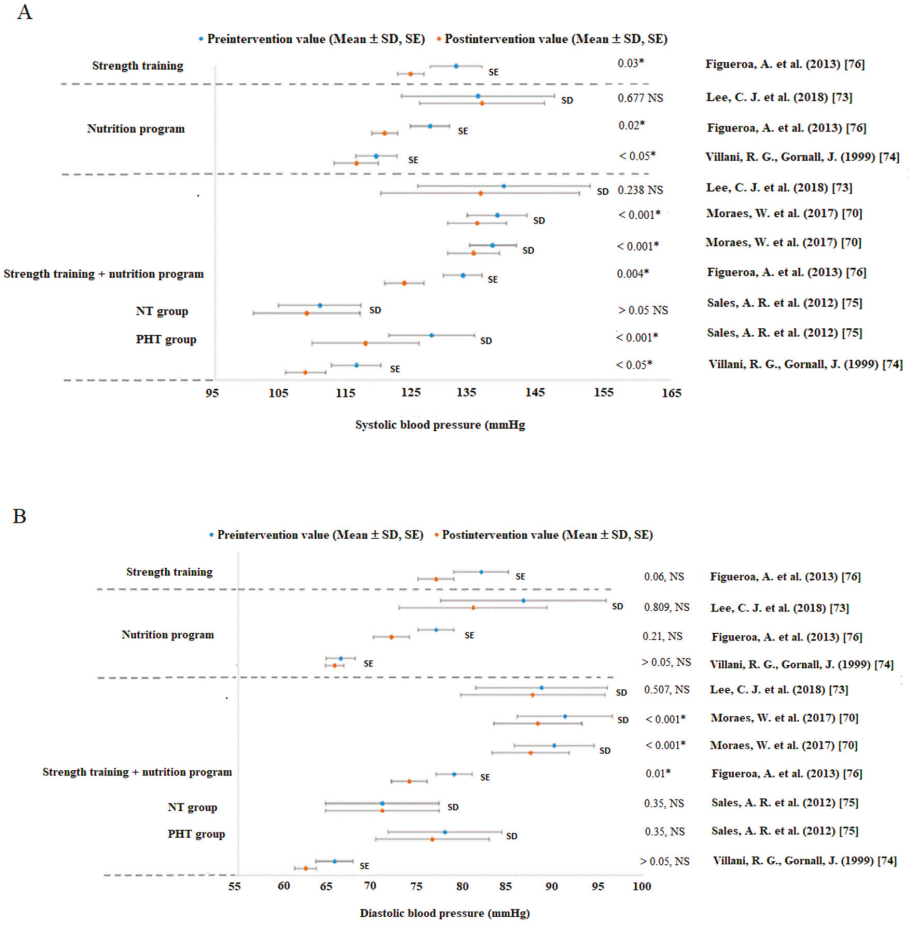


Figure 2. Changes in systolic (**A**) and diastolic (**B**) blood pressure (mmHg) in nutrition and strength training program studies. Abbreviations: NP = nutrition program, NT = normotension group, PHT = prehypertension group, SD = standard deviation, SE = standard error, ST = strength training, NS = not significant change—no change; * significant difference on reported *p* value.

Table 1. General description of each study, sample, and intervention characteristics.

Authors	Subjects	Aim	Results
Villani and Gornall (1999) [74]	Premenopausal women NP-only group: n = 10, mean age (y) = 33, weight (kg) = 75.78 ± 3.2 NP + Strength training group: n = 10, mean age (y) = 33, weight (kg) = 79.50 ± 2.86	The aim of the study was to investigate the combined influences of very-low-kilojoule diets and strength training on BP.	Resistance exercise did not significantly alter the BP reduction observed with short-term severe dieting.
Sales et al. (2012) [75]	Women with prehypertension: n = 10, age (y) = 39 ± 6, weight (kg) = 71.5 ± 10.7 Women with normotension n = 10, age (y) = 35 ± 11, weight (kg) = 66.5 ± 6.8	The aim of the study was to investigate the effect of diet and exercise training on BP and autonomic modulation in women with prehypertension.	Diet and exercise training reduced SBP in women with prehypertension, and this was associated with parasympathetic augmentation and probably reduction in sympathetic cardiac modulation.
Astorino and Martin (2013) [72]	Hypertensive men: n = 7, age (y) = 23.9 ± 4.6, height (m) = 1.8 ± 0.1, mass (kg) = 89 ± 16.2 Normotensive men: n = 7, age (y) = 22.4 ± 4.0, height (m) = 1.8 ± 0.1, weight (kg) = 77.9 ± 6.4	The primary aim of the study was to compare changes in BP in normotensive and prehypertensive men completing resistance exercise following caffeine ingestion.	Post-exercise hypotension did not occur in either treatment, suggesting that intense resistance training with or without caffeine intake may mitigate the BP-lowering effect of resistance exercise.
Figueroa et al. (2013) [76]	Postmenopausal women LIRET: n = 14, age (y) = 54 ± 1, height (m) = 1.66 ± 0.02, weight (kg) = 88.4 ± 4.6 Postmenopausal women NP: n = 13, age (y) = 54 ± 1, height (m) = 1.62 ± 0.02, weight (kg) = 89.0 ± 4.4 Postmenopausal women NP + LIRET: n = 14, age (y) = 54 ± 1, height (m) = 1.63 ± 0.02, weight (kg) = 86.7 ± 2.7	The aim of the study was to evaluate the independent and combined effects of a hypocaloric diet and LIRET with slow movement on PWV and body composition.	A hypocaloric diet decreases baPWV mainly by reducing legPWV, and this reduction was related to the loss of truncal fat. Although LIRET alone does not affect PWV or body composition, LIRET combined with diet improves baPWV and muscle strength while preventing loss of lean body mass in obese postmenopausal women.
Arazi et al. (2014) [77]	Middle-aged women: n = 24, age (y) = 46.4 ± 6.3, height (m) = 1.66 ± 4.2, weight (kg) = 66.6 ± 9.2 kg	The aim of the study was to investigate the effects of green tea extract on BP, HR, and RPP responses to low-intensity resistance exercise in hypertensive women.	Three weeks of green tea extract ingestion did not influence SBP, DBP or HR but may be have a favorable effect on MAP and RPP responses to an acute resistance exercise during a 1-h exercise recovery.

Table 1. *Cont.*

Authors	Subjects	Aim	Results
Wong et al. (2016) [78]	Postmenopausal women whole-body vibration training + Placebo: n = 14, age (y) = 58 ± 4.0, height (m) = 1.6 ± 0.06, weight (kg) = 89.5 ± 10.6 Postmenopausal women L-citrulline: n = 14, age (y) = 58 ± 4.0, height (m) = 1.6 ± 0.05, weight (kg) = 83.8 ± 8.4 Postmenopausal women WBVT + L-citrulline: n = 13, age (y) = 58 ± 3.0, height (m) = 1.62 ± 0.05, weight (kg) = 88.3 ± 3.9	The aim of the study was to examine the combined and independent effects of whole-body vibration training and L-citrulline supplementation on aortic hemodynamics and plasma nitric oxide metabolites in postmenopausal women.	This study supports the effectiveness of whole-body vibration training + L-citrulline as a potential intervention for the prevention of hypertension-related cardiac diseases in obese postmenopausal women.
Moraes et al. (2017) [70]	Low milk intake group: n = 16, age (y) = 70.2 ± 4.9 and weight (kg) = 70.1 ± 7.6 High milk intake group: n = 12, age (y) = 70.3 ± 5.0, weight (kg) = 68.6 ± 7.7	The aim of the study was to investigate whether the maintenance of exercise training benefits are associated with adequate milk and dairy product intake in elderly hypertensive subjects after detraining.	Maintenance of exercise training benefits related to pressure levels, lower extremity strength and aerobic capacity is associated with adequate milk and dairy product intake in hypertensive elderly subjects following six weeks of detraining.
Romero et al. (2017) [71]	Adults (men and women): n = 9, age (y) = 68 ± 5, height (m) = 1.65 ± 5, weight (kg) = 70 ± 8	The purpose of this study was to test the hypothesis that folic acid ingestion improves skeletal muscle blood flow in aged adults performing graded handgrip and plantar flexion exercise via increased vascular conductance.	Folic acid ingestion increases blood flow to active skeletal muscle primarily via improved local vasodilation in aged adults.
Lee et al. (2018) [73]	Adults (men and women) Advice-only comparison group: n = 28, age (y) = 43.4 ± 14.5, weight (kg) = 69.9 ± 9.2 Diet education group: n = 30, age (y) = 43.0 ± 13.5, weight (kg) = 72.8 ± 15.2 Diet and exercise education group: n = 27, age (y) 49.1 ± 10.1, weight (kg) = 76.6 ± 10.7	The aim of this study was to evaluate the effectiveness of a home-based lifestyle modification intervention on BP management.	Lifestyle modification emphasizing both diet and exercise was effective for lowering BP and should be favored over diet-only modifications.

baPWV, brachial–ankle pulse-wave velocity; BP, blood pressure; DBP, diastolic blood pressure; HR, heart rate; LIRET, low-intensity resistance exercise training, PWV, pulse-wave velocity; MAP, mean arterial pressure; RPP, rate pressure product; SBP, systolic blood pressure; y, year.

Table 2. Strength training parameters and methods in each study.

Authors	Sets	Rest Between Sets	Repetitions	Intensity	Frequency	Strength Training Methods	Number of Exercises
Villani and Gornall (1999) [74]	3	1–2 min	10	10 RM	3x per week	Resistance training	6
Sales et al. (2012) [75]	3	-	10	45–85% 1RM	3x per week	Resistance + aerobic training	10
Astorino and Martin (2013) [72]	4	2 min	As much as possible	70–80% 1RM	2 measurements	Resistance free weight training	4
Figueroa et al. (2013) [76]	2–3	-	18–22x	-	3x per week	Resistance machine training	4
Arazi et al. (2014) [77]	2	2 min	6–10	50% 1 RM	1 measurement	Resistance training	6
Wong et al. (2016) [78]	1–5	0.5–1 min	0.5–1 min	25–40 Hz	3x per week	Bodyweight training	8
Moraes et al. (2017) [70]	-	-	-	-	2x per week	Resistance + endurance + flexibility + stability training	-
Romero et al. (2017) [71]	2	20 min	1	3, 6, 9 kg	2 measurements	Isometric training	2
Lee et al. (2018) [73]	3	-	-	Mild–moderate intensity	Every day	Circuit training	5–7

RM = repetition maximum.

3.2. Strength Training Intervention

In studies by Figueroa et al. [76], Arazi et al. [77], Wong et al. [78] and Astorino et al. [72], a reduction in SBP after ST alone was detected in prehypertensive individuals. Only Astorino et al. [72] observed a negligible improvement in the BP values in normotensive individuals. Figueroa et al. [76] also detected a reduction in DBP after ST. It seems that ST is one of the initiators of post-exercise hypotension in adults, regardless of the type of ST (see in Figure 2A,B, Figure 3 and Tables S1–S3).

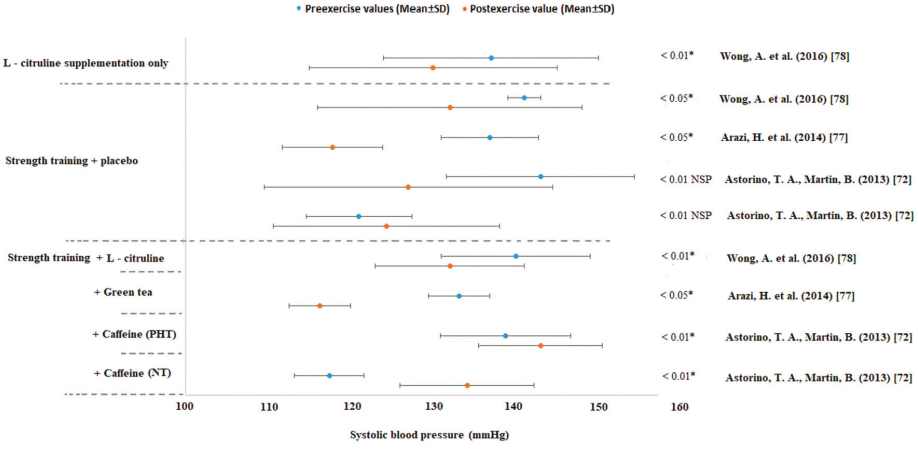


Figure 3. Changes in systolic blood pressure (mmHg) in supplement studies. Abbreviations: SD = standard deviation, * significant difference on reported "*p*" value. NSP = not significant according to the post hoc test.

Body composition and strength parameters have been improved by ST alone in the study by Figueroa et al. [76]. This confirmed the assumption that ST has a generally positive effect on health status because it might reduce body fat and increase muscle mass and other conditioning values. Strength abilities were improved in participants of studies combining ST and nutrition, and these findings have been reported by Sales et al. [75], Figueroa et al. [76], Moraes et al. [70], and Lee et al. [73]. Strength abilities worsened in the nontraining group (only NP) in the study by Figueroa et al. [76]. In contrast, the Dietary Approaches to Stop Hypertension (DASH) diet group had no negative effect on strength according to Lee et al. [73]. A comparison of the effect of ST alone or ST and NP on absolute and relative strength bring very similar results as those shown by Figueroa et al. [76]. The aerobic parameters were improved in the ST group with NP in the studies by Moraes et al. [70] and Lee et al. [73], and the VO2max was improved in the study by Sales et al. [75].

3.3. Nutrition Program

Figure 2A,B show that SBP and DBP decreased after a nutrition program [74,76]. Only a study by Lee et al. [73] reported a negligible increase in SBP after eight weeks of a NP in individuals with prehypertension (PHT) and HT. Moraes et al. [70] showed that a higher intake of dairy products together with combined aerobic and anaerobic training can lead to a slightly larger drop in BP than that associated with a lower intake of dairy products. Moreover, the higher intake of dairy products resulted in a smaller increase in BP (i.e., return of BP to the initial values) after six weeks of nontraining [70]. It appears that the inclusion of more dairy products in the NP together with combined aerobic and anaerobic training has a positive effect on the reductions in the SBP and DBP values.

A study by Villani, Gornall [74], Figueroa et al. [76], and Lee et al. [73] revealed that a training program combined with a NP leads to a significant drop in BP. Moreover, one study showed that ST alone had similar effectiveness as that of the NP [76].

3.4. The Effectivity of Food Supplements

Caffeine is a very popular pre-workout stimulant among athletes. Astorino et al. [72] found that caffeine intake immediately before training increases SBP ($p < 0.05$). However, the effect on HR ($p = 0.16$) and DBP ($p = 0.10$) was similar for the caffeine and placebo. Values of a HR and BP were significantly higher in men with PHT than in normotensive men ($p < 0.05$).

A study by Arazi et al. [77] performed on middle-aged women with HT revealed practically the same reduction in BP after resistance training with or without green tea supplementation (Figure 3). There was no significant difference between the placebo and green tea intake groups (500 mg daily = 245 mg polyphenol, 75 mg epigallocatechin gallate, 25 mg caffeine) or placebo (490 mg maltodextrin). The participants performed circuit training consisting of two sets with a resistance of 50% one repetition maximum (1RM). BP was measured at zero, 15, 30, 45, and 60 min after training.

Wong et al. [78] observed the BP changes in postmenopausal obese women with a BMI ≥ 25 kg/m^2. There was a similar decrease in the brachial and arterial BP in all groups ($p < 0.05$). Vibration training in combination with the placebo, vibration training with L-citrulline or L-citrulline alone can be used as effective ways to lower BP (Figure 3). However, there was no improvement in body composition between the study groups.

Romero et al. [71] observed the effects of folic acid supplementation for six weeks among seniors. The authors found that the folic acid intake immediately before training reduces HR but not mean arterial pressure. At the end of the six-week experiment, the HR values and mean arterial pressure were higher than the baseline values ($p = 0.05$ compared to acute folic acid ingestion). A significant positive feature is the fact that folic acid increases the blood flow to active skeletal muscles, mainly due to better local vasodilation.

4. Discussion

The present study is the first systematic review to analyze the evidence for the effectiveness of ST combined with the NP or supplementation on the BP values. This systematic research analyzed the results of 303 individuals from nine studies. Studies by Moraes et al. [70], Villani, Gornall [74], Sales et al. [75], Figueroa et al. [76], and Lee et al. [73] compared the effects of ST combined with a NP on BP. All of these studies used different methods and parameters for the ST protocol and the NP. A common feature of these five studies was the reductions in SBP and DBP or mean arterial pressure in all experimental groups, regardless of the type of training or diet. Only four studies with supplements met the criteria for inclusion in the systematic review: Astorino et al. [72], Arazi et al. [77], Wong et al. [78] and Romero et al. [71]. The effects of caffeine, L-citrulline, folic acid, and green tea were investigated. Only two supplements had a positive impact on BP: L-citrulline and green tea. On the other hand, caffeine and folic acid did not decrease BP, and caffeine has even been reported to increase BP.

In general, the issue of ST and diet or supplementation has been considered as completely different disciplines, which should be examined separately. However, in the case of treatment of PHT and HT, their combination may appear to be the most effective method. Although this study was not able to conclude whether ST alone or ST combined with the NP has the greatest impact on arterial pressure reduction, it successfully summarized current studies comparing different ST and NP interventions.

4.1. Blood Pressure Reduced by Strength Training Alone

It is generally known that ST increases strength, muscle mass, and bone mass and simultaneously helps reduce the symptoms of various chronic diseases, such as heart disease [81–85]. In 2005, the AHA (American Heart Association) started to recommended ST for lowering BP [2] because ST induces a post-exercise BP decrease, as supported by many reviews [14–16,86]. One of the first long-term ongoing studies conducted on cardiovascular disease was the Framingham Heart Study [24], where one of the findings reported was that a 2 mmHg reduction in DBP was associated with an estimated 17%

decrease in the prevalence of HT. However, unsuitable training parameters such as the work load, reps per set, rest interval, etc., can increase BP [87] above the recommended values, i.e., 220/105 mmHg [1]. This study summarize that ST alone can decrease SBP from 132 ± 4 mmHg to 125 ± 2 mmHg [76] or from 141 ± 2 mmHg to 132 ± 16 mmHg [78], and can decrease DBP from 82 ± 3 mmHg to 77 ± 2 mmHg [76], which is much more than the smallest significant values of 2 mmHg. The one-time effect of ST induced a SBP decrease (from 136.88 ± 5.9 mmHg to 117.82 ± 6.09 mmHg) in the study by Arazi et al. [77] and in the prehypertensive group of the study by Astorino et al. [72] (from 143 ± 11.4 mmHg to 131.7 ± 16.6 mmHg). However, in a study by Arazi et al. [77], the biggest BP reduction was not observed immediately after training but after 1 h. There is great diversity between the training sessions in this systematic review. None of the included studies was identical in terms of the training parameters, methods, or exercises (Table 2), but all included studies set up the ST intervention according to the recommended parameters and methods published by Pescatello et al. [80]. Moreover, no study reported a dangerous increase of BP over 220/105 mmHg (according to the ACSM) or substantial post-exercise hypotension. According to the results of this review, ST alone can be recommended as an effective nonpharmacological intervention and prevention method for people with HT or PHT.

4.2. Blood Pressure was Reduced by the Nutrition Program

Nutrition-based approaches are recommended as a first-line therapy for the prevention of HT [72], where the AHA recommends a specific program called the DASH diet to treat and prevent HT. However, some food components such as alcohol, sodium, simple sugar, and saturated fat, have been shown to increase BP [9,51,88]. It has also been found that for HT reduction, weight loss is essential, which has been shown to reduce BP in overweight hypertensive and prehypertensive individuals [10,89–92]. This systematic review included only one study that used the DASH diet. The Korean variation of the DASH diet in a study by Lee et al. [73] improved SBP and DBP only in conjunction with ST (from 139.3 ± 13.2 mmHg to 135.7 ± 15.3 mmHg). Diet alone increased SBP (from 135.3 ± 11.8 mmHg to 135 ± 9.6 mmHg), but decreased DBP (from 86.7 ± 9.2 mmHg to 81.1 ± 8.2 mmHg [73]. Conversely, Villani, Gornall [74], Sales et al. [75] and Figueroa et al. [76] included hypocaloric diets, which led to a reduction in BP in both groups (ST group and NP alone). Moreover, it emerged that ST alone had similar effectiveness as the NP alone [76]. Based on this review result, we recommend a NP that takes into account the individual components of the food and that also leads to a drop in BP. Moraes et al. [70] showed that the inclusion of more dairy products in the NP, together with training, can have a positive effect on the lowering of the SBP (from 138.3 ± 4.6 mmHg to 135.2 ± 4.5 mmHg) and DBP (from 91.3 ± 5.3 mmHg to 88.3 ± 4.9 mmHg) values. The same higher intake of dairy products resulted in a smaller increase in BP (i.e., return of BP to the initial values) after six weeks of nontraining. It seems that a higher intake of dairy products prolongs training hypotension. Therefore, multiple effects of ST and a NP should not be expected for the BP lowering. Although their combination is not more effective, their combination might result in more health benefits (such as weight loss) than ST and NP alone.

4.3. Effect of Supplements on Blood Pressure

Caffeine is a central nervous stimulant whose physiological effects for increasing sport performance are extensive, but sometimes conflicting or contradicting [93,94]. Caffeine supplementation has been shown to decrease feelings of fatigue and promote mood and perceptual responses during any exercise, including ST [94–96]. It has been found that SBP and DBP increase after caffeine ingestion due to the vasoconstrictive effects of caffeine [93]. However, some studies have reported mixed results regarding caffeine intake and BP. In randomized controlled trials (RCTs), short-term caffeine intake caused an acute increase in SBP and DBP by 2/1 mmHg, respectively, compared with the effects of decaffeinated coffee or abstinence [8,97]. Caffeine supplementation, according to the results of Astorino et al. [72], cannot be recommended in individuals with HT or PHT because it not only increases the resting BP but also maintains the BP at a high level after the end of the workout. Therefore, the hypotensive effect

of ST is completely lost [72]. For that reason, we cannot recommend caffeine to individuals with HT or PHT although caffeine reduces body fat, increases sport performance, and delays fatigue [94–96].

L-citrulline is a precursor of L-arginine, a substrate for nitric oxide synthase, in the production of nitric oxide. Deficiencies in the L-arginine supply have been strongly implicated in cardiovascular diseases, including HT, atherosclerosis, diabetic vascular disease, hyperhomocysteinemia, heart failure, etc. [98–100]. In a study by Wong et al. [78], L-citrulline decreased BP in individuals performing ST or not performing ST. Both combinations led to a reduction, but slightly better results were achieved in the group that performed the ST alone. Here, again, this study confirmed that ST plays a primary role in the entire process of hypotension. The main and indisputable advantage of L-citrulline is that it reduces BP alone but also in combination with NP or ST. However, more studies are needed to explore this supplement.

A link between green tea and BP reduction has been explored for decades in Chinese populations [101]. However, there are few studies regarding the long-term effects of tea drinking on the risk of HT, and the results of the few studies investigating the relationship between tea consumption and BP were opposing. In epidemiological studies, a higher consumption of black tea in Norwegian individuals was associated with a lower SBP [102], while the green tea intake in Japanese self-defense officials was unrelated to BP [103]. Long-term effects of green tea consumption do not reduce the values of SBP or DBP after training in comparison with the effects of training alone. In contrast, a significant difference between the two groups was observed in the increase in HR after the end of the training. The green tea group performing RT showed a smaller increase in HR than the training group alone. These positive effects can be attributed, at least in part, to the antioxidant properties and vasodilating effects of the catechins in green tea [104]. Unlike caffeine, green tea is not dangerous for individuals with a high BP because it does not increase BP. For this reason, it can be included for its stimulating effects as a preworkout drink. A very significant drop in the systolic blood pressure was measured in the study by Arazi et al. [77], where blood pressure dropped from preintervention values from 133.12 mmHg (± 3.7) to 116.25 mmHg (± 3.71) in the group that combine ST and green tea. According to this information, green tea can be recommended as a suitable food supplement for the treatment of HT. However, supplementation with green tea together with ST does not lead to a bigger reduction in BP in comparison with the effects of the ST alone.

The primary risk factor for a stroke, is HT as stated by Meschia et al. [105]. For this reason, the effect of folic acid on the frequency of infarcts was investigated. Most of the relevant randomized trials were designed for secondary prevention and have not shown a beneficial effect of folic acid for cardiovascular disease prevention [106–109]. Based on this study result, folic acid had an effect on the HR and mean arterial pressure [71]. The acute folate intake before exercise induced a reduction in HR but not a long-term decrease. Even the long-term intake did not lead to reduction. For all measurements, a similar increase in HR and mean arterial pressure during isometric exercises was observed with the acute and long-term intake of folic acid [71]. Based on this study, folic acid supplementation in elderly individuals does not lead to a significant decrease in the mean arterial BP or HR during exercise compared to the measurements of the control subjects [71]. Similar results were observed in studies by Huo et al. [63] and Hernández-Díaz et al. [110]. According to these findings, we cannot recommend folic acid alone for lowering BP.

4.4. Combined Effect of Strength Training, Nutrition Program and Supplementation on Blood Pressure

Although BP lowering is effective using ST alone [76–78], NP alone [74,76] and some food supplements alone [77,78], the combination of ST, NP, and supplementation did not have additional systolic BP lowering effectivity (Supplemetary Table S1). On the other hand, the combination of ST and NP seems to be necessary to decrease the diastolic BP [70,78] (Supplemetary Table S2). Moreover, the combination of ST with green tea or L-citrulline decreased the systolic BP on a higher significance or effect size (Supplemetary Table S3) than the ST alone, although the significant difference between the ST and ST with supplementation has not been directly reported. This findings supports the notion

that total behavioral modification such as lifestyle modifications, rather than focusing on modifying a single behavioral target, is more important. Therefore, a combination of ST with the energy intake restriction NP, which include food supplement stimulants green tea and L-citrulline should be used for BP lowering and additional health benefits such as body composition changes.

4.5. Changes in Body Composition

Body composition improvements have been shown as a result of ST and NPs in studies by Villani et al. [74], Sales et al. [75], Figueroa et al. [76], Moraes et al. [70], and Lee et al. [73]. Slightly better results, in terms of improved body composition, were achieved by a NP alone in the studies by Villani et al. [74] and Figueroa et al. [76]. On the other hand, better results were not achieved by the DASH alone compared with the DASH diet and a training program in the latest study by Lee et al. [73]. Only a small improvement in the body composition was recorded by Figueroa et al. [76] in the ST group. Furthermore, Figueroa et al. [76] found that a hypocaloric diet decreased the brachial–ankle pulse-wave velocity mainly by reducing the leg pulse-wave velocity, and this reduction was related to fat loss. Although ST alone does not affect the pulse-wave velocity or body composition, ST combined with NP improves the brachial–ankle pulse-wave velocity and muscle strength while preventing the loss of lean body mass in obese postmenopausal women. The results of the pulse-wave velocity are considered an independent predictor of systolic hypertension [111].

In this case, a NP appears to be important for body composition transformation. However, the advantages of ST cannot be overlooked, and their combination seems to be the optimal variant. This systematic review reported some important findings concerning the different effects of various training, nutrition, and supplementary strategies on HT. Individuals with a high BP can improve their BP values, not only by using medication but also by utilizing a nonpharmacological method, which will not only have a positive impact on BP but also on other health components, such as body composition, muscle strength, and bone density. Future studies should more closely examine the effect of specific training, nutrition, and supplementation programs on individuals with HT, normotension, and PHT. Additionally, future studies should explore mechanisms of how different nutrition programs and supplements lower BP.

5. Conclusions

Presented data suggest that a combination of ST and NP seems to be an effective solution for lowering BP (systolic and diastolic). This method can be recommended for individuals with PHT or HT. However, an essential role can be attributed to the strength training program alone. In both aerobic and anaerobic forms, exercise causes post-exercise hypotension, but the effect depends on the selected training parameters. A NP based on the restriction of energy intake positively affects BP and body fat. However, this study cannot recommend caffeine supplementation because it increases BP during ST, while a more suitable stimulant is green tea or L-citrulline, which lower BP. For BP lowering in clinical practice, it is recommended to prescribe a combination of ST with the energy intake restriction NP, which include food supplements stimulants green tea and L-citrulline.

Supplementary Materials: The following are available online at http://www.mdpi.com/2077-0383/8/9/1393/s1. Table S1: Changes in the systolic blood pressure in strength training and nutrition program groups; Table S2: Changes in the diastolic blood pressure in strength training and nutrition program groups; Table S3: Changes in the systolic blood pressure in strength training and supplementary groups.

Author Contributions: Conceptualization, R.J. and P.S.; Methodology, R.J. and P.S.; Writing—original draft preparation, R.J. and P.S.; Writing—review and editing, R.J. and P.S.

Funding: This research was funded by Charles University grant no. UNCE/HUM/032, SVV 2017-2019 260466 and GACR GA19-12150S.

Acknowledgments: This study was supported by Charles University grant no. UNCE/HUM/032, SVV 2017-2019 260466 and GACR GA19-12150S.

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

References

1. American College of Sports Medicine. American College of Sports Medicine position stand. Progression models in resistance training for healthy adults. *Med. Sci. Sports Exerc.* **2009**, *41*, 687. [CrossRef] [PubMed]
2. Pickering, T.G.; Hall, J.E.; Appel, L.J.; Falkner, B.E.; Graves, J.W.; Hill, M.N.; Jones, D.W.; Kurtz, T.; Sheps, S.G.; Roccella, E.J. Recommendations for blood pressure measurement in humans: An AHA scientific statement from the council on high blood pressure research professional and public education subcommittee. *J. Clin. Hypertens.* **2005**, *7*, 102–109. [CrossRef]
3. Chobanian, A.V.; Bakris, G.L.; Black, H.R.; Cushman, W.C.; Green, L.A.; Izzo, J.L., Jr.; Jones, D.W.; Materson, B.J.; Oparil, S.; Wright, J.T., Jr.; et al. Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. *Hypertension* **2003**, *42*, 1206–1252. [CrossRef] [PubMed]
4. Lamotte, M.; Niset, G.; Van De Borne, P. The effect of different intensity modalities of resistance training on beat-to-beat blood pressure in cardiac patients. *Eur. J. Cardiovasc. Prev. Rehabil.* **2005**, *12*, 12–17. [CrossRef]
5. Meyer, K. Resistance exercise in chronic heart failure—Landmark studies and implications for practice. *Clin. Investig. Med.* **2006**, *29*, 166–169.
6. Pescatello, L.S.; Franklin, B.A.; Fagard, R.; Farquhar, W.B.; Kelley, G.A.; Ray, C.A. American College of Sports Medicine position stand. Exercise and hypertension. *Med. Sci. Sports Exerc.* **2004**, *36*, 533–553. [CrossRef] [PubMed]
7. Thompson, P.D.; Arena, R.; Riebe, D.; Pescatello, L.S.; American College of Sports Medicine. ACSM's new preparticipation health screening recommendations from ACSM's guidelines for exercise testing and prescription, ninth edition. *Curr. Sports Med. Rep.* **2013**, *12*, 215–217. [CrossRef]
8. Hamer, M. The anti-hypertensive effects of exercise: Integrating acute and chronic mechanisms. *Sports Med.* **2006**, *36*, 109–116. [CrossRef] [PubMed]
9. Sacks, F.M.; Svetkey, L.P.; Vollmer, W.M.; Appel, L.J.; Bray, G.A.; Harsha, D.; Obarzanek, E.; Conlin, P.R.; Miller, E.R.; Simons-Morton, D.G.; et al. Effects on Blood Pressure of Reduced Dietary Sodium and the Dietary Approaches to Stop Hypertension (DASH) Diet. *N. Engl. J. Med.* **2001**, *344*, 3–10. [CrossRef]
10. Stevens, V.J.; Obarzanek, E.; Cook, N.R.; Lee, I.M.; Appel, L.J.; West, D.S.; Milas, N.C.; Mattfeldt-Beman, M.; Belden, L.; Bragg, C.; et al. Long-term weight loss and changes in blood pressure: Results of the trials of hypertension prevention, phase II. *Ann. Intern. Med.* **2001**, *134*, 1–11. [CrossRef]
11. LaMonte, M.J.; Yanowitz, F.G. Aerobic exercise for lowering blood pressure: A metaanalysis. *Clin. J. Sport Med.* **2002**, *12*, 407. [PubMed]
12. Pescatello, L.S.; Kulikowich, J.M. The aftereffects of dynamic exercise on ambulatory blood pressure. *Med. Sci. Sports Exerc.* **2001**, *33*, 1855–1861. [CrossRef] [PubMed]
13. Whelton, S.P.; Chin, A.; Xin, X.; He, J. Effect of aerobic exercise on blood pressure: A meta-analysis of randomized, controlled trials. *Ann. Intern. Med.* **2002**, *136*, 493–503. [CrossRef] [PubMed]
14. da Rocha, P.E.C.P.; da Silva, V.S.; Camacho, L.A.B.; Vasconcelos, A.G.G. Effects of long-term resistance training on blood pressure: A systematic review. *Revista Brasileira de Cineantropometria Desempenho Humano* **2017**, *19*, 730–742.
15. Inder, J.D.; Carlson, D.J.; Dieberg, G.; McFarlane, J.R.; Hess, N.C.; Smart, N.A. Isometric exercise training for blood pressure management: A systematic review and meta-analysis to optimize benefit. *Hypertens. Res.* **2016**, *39*, 88–94. [CrossRef]
16. Lemes, I.R.; Ferreira, P.H.; Linares, S.N.; MacHado, A.F.; Pastre, C.M.; Netto, J. Resistance training reduces systolic blood pressure in metabolic syndrome: A systematic review and meta-analysis of randomised controlled trials. *Br. J. Sports Med.* **2016**, *50*, 1438–1442. [CrossRef] [PubMed]
17. Wewege, M.A.; Thom, J.M.; Rye, K.A.; Parmenter, B.J. Aerobic, resistance or combined training: A systematic review and meta-analysis of exercise to reduce cardiovascular risk in adults with metabolic syndrome. *Atherosclerosis* **2018**, *274*, 162–171. [CrossRef]
18. Hovell, M.F. The experimental evidence for weight-loss treatment of essential hypertension: A critical review. *Am. J. Public Health* **1982**, *72*, 359–368. [CrossRef]
19. Ndanuko, R.N.; Tapsell, L.C.; Charlton, K.E.; Neale, E.P.; Batterham, M.J. Dietary Patterns and Blood Pressure in Adults: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Adv. Nutr.* **2016**, *7*, 76–89. [CrossRef]

20. Saneei, P.; Salehi-Abargouei, A.; Esmaillzadeh, A.; Azadbakht, L. Influence of Dietary Approaches to Stop Hypertension (DASH) diet on blood pressure: A systematic review and meta-analysis on randomized controlled trials. *Nutr. Metab. Cardiovasc. Dis.* **2014**, *24*, 1253–1261. [CrossRef]
21. Schwingshackl, L.; Chaimani, A.; Schwedhelm, C.; Toledo, E.; Pünsch, M.; Hoffmann, G.; Boeing, H. Comparative effects of different dietary approaches on blood pressure in hypertensive and pre-hypertensive patients: A systematic review and network meta-analysis. *Crit. Rev. Food Sci Nutr.* **2018**. [CrossRef] [PubMed]
22. Herrod, P.J.J.; Doleman, B.; Blackwell, J.E.M.; O'Boyle, F.; Williams, J.P.; Lund, J.N.; Phillips, B.E. Exercise and other nonpharmacological strategies to reduce blood pressure in older adults: A systematic review and meta-analysis. *J. Am. Soc. Hypertens.* **2018**, *12*, 248–267. [CrossRef] [PubMed]
23. Aucott, L.; Poobalan, A.; Smith, W.C.S.; Avenell, A.; Jung, R.; Broom, J. Effects of weight loss in overweight/obese individuals and long-term hypertension outcomes: A systematic review. *Hypertension* **2005**, *45*, 1035–1041. [CrossRef] [PubMed]
24. Moore, L.L.; Visioni, A.J.; Qureshi, M.M.; Bradlee, M.L.; Ellison, R.C.; D'Agostino, R. Weight loss in overweight adults and the long-term risk of hypertension the framingham study. *Arch. Intern. Med.* **2005**, *165*, 1298–1303. [CrossRef] [PubMed]
25. Mulrow, C.D.; Chiquette, E.; Angel, L.; Cornell, J.; Summerbell, C.; Anagnostelis, B.; Grimm, R., Jr.; Brand, M.B. Dieting to reduce body weight for controlling hypertension in adults. *Cochrane Database Syst. Rev.* **2000**, *2*, CD000484.
26. Sjostrom, M.; Karlsson, A.B.; Kaati, G.; Yngve, A.; Green, L.W.; Bygren, L.O. A four week residential program for primary health care patients to control obesity and related heart risk factors: Effective application of principles of learning and lifestyle change. *Eur J. Clin. Nutr* **1999**, *53* (Suppl. 2), S72–S77. [CrossRef]
27. Stevens, V.J.; Corrigan, S.A.; Obarzanek, E.; Bernauer, E.; Cook, N.R.; Hebert, P.; Mattfeldt-Beman, M.; Oberman, A.; Sugars, C.; Dalcin, A.T.; et al. Weight loss intervention in phase 1 of the Trials of Hypertension Prevention. The TOHP Collaborative Research Group. *Arch. Intern. Med.* **1993**, *153*, 849–858. [CrossRef]
28. Figueiredo, T.; Rhea, M.R.; Peterson, M.; Miranda, H.; Bentes, C.M.; dos Reis, V.M.; Simao, R. Influence of number of sets on blood pressure and heart rate variability after a strength training session. *J. Strength Cond. Res.* **2015**, *29*, 1556–1563. [CrossRef]
29. Moraes, M.R.; Bacurau, R.F.; Simoes, H.G.; Campbell, C.S.; Pudo, M.A.; Wasinski, F.; Pesquero, J.B.; Wurtele, M.; Araujo, R.C. Effect of 12 weeks of resistance exercise on post-exercise hypotension in stage 1 hypertensive individuals. *J. Hum. Hypertens.* **2012**, *26*, 533–539. [CrossRef]
30. Scher, L.M.; Ferriolli, E.; Moriguti, J.C.; Scher, R.; Lima, N.K. The effect of different volumes of acute resistance exercise on elderly individuals with treated hypertension. *J. Strength Cond. Res.* **2011**, *25*, 1016–1023. [CrossRef]
31. Tokmakidis, S.P.; Kalapotharakos, V.I.; Smilios, I.; Parlavantzas, A. Effects of detraining on muscle strength and mass after high or moderate intensity of resistance training in older adults. *Clin. Physiol. Funct. Imaging* **2009**, *29*, 316–319. [CrossRef] [PubMed]
32. Baum, K.; Rüther, T.; Essfeld, D. Reduction of blood pressure response during strength training through intermittent muscle relaxations. *Int. J. Sports Med.* **2003**, *24*, 441–445. [PubMed]
33. Gjovaag, T.; Hjelmeland, A.K.; Øygard, J.B.; Vikne, H.; Mirtaheri, P. Acute hemodynamic and cardiovascular responses following resistance exercise to voluntary exhaustion. Effects of different loadings and exercise durations. *J. Sports Med. Phys. Fit.* **2016**, *56*, 616–623.
34. Gjøvaag, T.; Hjelmeland, A.K.; Øygard, J.B.; Vikne, H.; Mirtaheri, P. Resistance exercise and acute blood pressure responses. *J. Sports Med. Phys. Fit.* **2016**, *56*, 616–623.
35. Nascimento Dda, C.; Tibana, R.A.; Benik, F.M.; Fontana, K.E.; Ribeiro Neto, F.; Santana, F.S.; Santos-Neto, L.; Silva, R.A.; Silva, A.O.; Farias, D.L.; et al. Sustained effect of resistance training on blood pressure and hand grip strength following a detraining period in elderly hypertensive women: A pilot study. *Clin. Interv. Aging* **2014**, *9*, 219–225. [CrossRef] [PubMed]
36. Lamotte, M.; Fleury, F.; Pirard, M.; Jamon, A.; van de Borne, P. Acute cardiovascular response to resistance training during cardiac rehabilitation: Effect of repetition speed and rest periods. *Eur. J. Prev. Cardiol.* **2010**, *17*, 329–336. [CrossRef] [PubMed]
37. Paulo, A.C.; Tricoli, V.; Queiroz, A.C.C.; Laurentino, G.; Forjaz, C.L.M. Blood Pressure Response During Resistance Training of Different Work-to-Rest Ratio. *J. Strength Cond. Res.* **2019**, *33*, 399–407. [CrossRef]

38. Veloso, J.; Polito, M.D.; Riera, T.; Celes, R.; Vidal, J.C.; Bottaro, M. Effects of rest interval between exercise sets on blood pressure after resistance exercises. *Arquivos Brasileiros de Cardiologia* **2010**, *94*, 482–487.
39. Boden, G.; Sargrad, K.; Homko, C.; Mozzoli, M.; Stein, T.P. Effect of a low-carbohydrate diet on appetite, blood glucose levels, and insulin resistance in obese patients with type 2 diabetes. *Ann. Intern. Med.* **2005**, *142*, 403–411. [CrossRef]
40. Yancy, W.S., Jr.; Foy, M.; Chalecki, A.M.; Vernon, M.C.; Westman, E.C. A low-carbohydrate, ketogenic diet to treat type 2 diabetes. *Nutr. Metab.* **2005**, 2. [CrossRef]
41. Yancy, W.S., Jr.; Westman, E.C.; McDuffie, J.R.; Grambow, S.C.; Jeffreys, A.S.; Bolton, J.; Chalecki, A.; Oddone, E.Z. A randomized trial of a low-carbohydrate diet vs orlistat plus a low-fat diet for weight loss. *Arch. Intern. Med.* **2010**, *170*, 136–145. [CrossRef]
42. Burke, V.; Hodgson, J.M.; Beilin, L.J.; Giangiulioi, N.; Rogers, P.; Puddey, I.B. Dietary protein and soluble fiber reduce ambulatory blood pressure in treated hypertensives. *Hypertension* **2001**, *38*, 821–826. [CrossRef]
43. Attina, D.A.; Cupelli, V.; Milicia, R.; Catalano, G.; Scavelli, A. Effects of physical exercise of different duration in hypertensive elderly subjects. *Int. J. Sports Cardiol.* **1996**, *5*, 67–70.
44. Fekete, A.A.; Giromini, C.; Chatzidiakou, Y.; Givens, D.I.; Lovegrove, J.A. Whey protein lowers blood pressure and improves endothelial function and lipid biomarkers in adults with prehypertension and mild hypertension: Results from the chronic Whey2Go randomized controlled trial. *Am. J. Clin. Nutr.* **2016**, *104*, 1534–1544. [CrossRef]
45. Fekete, Á.A.; Giromini, C.; Chatzidiakou, Y.; Givens, D.I.; Lovegrove, J.A. Whey protein lowers systolic blood pressure and Ca-caseinate reduces serum TAG after a high-fat meal in mildly hypertensive adults. *Sci. Rep.* **2018**, 8. [CrossRef]
46. Fekete, Á.A.; Ian Givens, D.; Lovegrove, J.A. Casein-derived lactotripeptides reduce systolic and diastolic blood pressure in a meta-analysis of randomised clinical trials. *Nutrients* **2015**, *7*, 659–681. [CrossRef]
47. He, J.; Gu, D.; Wu, X.; Chen, J.; Duan, X.; Chen, J.; Whelton, P.K. Effect of soybean protein on blood pressure: A randomized, controlled trial. *Ann. Intern. Med.* **2005**, *143*, 1–9. [CrossRef]
48. Pal, S.; Ellis, V. The chronic effects of whey proteins on blood pressure, vascular function, and inflammatory markers in overweight individuals. *Obesity* **2010**, *18*, 1354–1359. [CrossRef]
49. Rebholz, C.M.; Friedman, E.E.; Powers, L.J.; Arroyave, W.D.; He, J.; Kelly, T.N. Dietary protein intake and blood pressure: A meta-analysis of randomized controlled trials. *Am. J. Epidemiol.* **2012**, *176*, S27–S43. [CrossRef]
50. Al-Solaiman, Y.; Jesri, A.; Mountford, W.K.; Lackland, D.T.; Zhao, Y.; Egan, B.M. DASH lowers blood pressure in obese hypertensives beyond potassium, magnesium and fibre. *J. Hum. Hypertens.* **2010**, *24*, 237–246. [CrossRef]
51. Appel, L.J.; Moore, T.J.; Obarzanek, E.; Vollmer, W.M.; Svetkey, L.P.; Sacks, F.M.; Bray, G.A.; Vogt, T.M.; Cutler, J.A.; Windhauser, M.M.; et al. A clinical trial of the effects of dietary patterns on blood pressure. *N. Engl. J. Med.* **1997**, *336*, 1117–1124. [CrossRef]
52. Bazzano, L.A.; Green, T.; Harrison, T.N.; Reynolds, K. Dietary approaches to prevent hypertension. *Curr. Hypertens. Rep.* **2013**, *15*, 694–702. [CrossRef]
53. Estruch, R.; Martínez-González, M.A.; Corella, D.; Salas-Salvadó, J.; Ruiz-Gutiérrez, V.; Covas, M.I.; Fiol, M.; Gómez-Gracia, E.; López-Sabater, M.C.; Vinyoles, E.; et al. Effects of a Mediterranean-Style Diet on Cardiovascular Risk Factors a Randomized Trial. *Ann. Intern. Med.* **2006**, *145*, 1–11. [CrossRef]
54. Rees, K.; Hartley, L.; Flowers, N.; Clarke, A.; Hooper, L.; Thorogood, M.; Stranges, S. 'Mediterranean' dietary pattern for the primary prevention of cardiovascular disease. *Cochrane Database Syst. Rev.* **2013**, *2013*. [CrossRef]
55. Svetkey, L.P.; Simons-Morton, D.; Vollmer, W.M.; Appel, L.J.; Conlin, P.R.; Ryan, D.H.; Ard, J.; Kennedy, B.M. Effects of dietary patterns on blood pressure: Subgroup analysis of the Dietary Approaches to Stop Hypertension (DASH) randomized clinical trial. *Arch. Intern. Med.* **1999**, *159*, 285–293. [CrossRef]
56. Barbeau, P.A.; Holloway, T.M.; Whitfield, J.; Baechler, B.L.; Quadrilatero, J.; van Loon, L.J.C.; Chabowski, A.; Holloway, G.P. α-Linolenic acid and exercise training independently, and additively, decrease blood pressure and prevent diastolic dysfunction in obese Zucker rats. *J. Physiol.* **2017**, *595*, 4351–4364. [CrossRef]
57. Moradi, H.; Nikbakht, H.; Ebrahim, P.R.; Abed-Natanzi, H. The effects of 8 weeks of aerobic training and grape seed extract on some cardiovascular functional indexes in non-active men. *J. Isfahan Med. Sch.* **2017**, *34*, 1439–1444.

58. Neto, M.M.; da Silva, T.F.; de Lima, F.F.; Siqueira, T.M.Q.; Toscano, L.T.; de Moura, S.; Silva, A.S. Whole Red Grape Juice Reduces Blood Pressure at Rest and Increases Post-exercise Hypotension. *J. Am. Coll. Nutr.* **2017**, *36*, 533–540. [CrossRef] [PubMed]
59. Puga, G.M.; De, P.; Novais, I.; Katsanos, C.S.; Zanesco, A. Combined effects of aerobic exercise and l-arginine ingestion on blood pressure in normotensive postmenopausal women: A crossover study. *Life Sci.* **2016**, *151*, 323–329. [CrossRef]
60. Digiesi, V.; Cantini, F.; Brodbeck, B. Effect of coenzyme Q10 on essential arterial hypertension. *Curr. Ther. Res.* **1990**, *47*, 841–845.
61. Digiesi, V.; Cantini, F.; Bisi, G.; Guarino, G.; Oradei, A.; Littarru, G. Mechanism of action of coenzyme Q10 in essential hypertension. *Curr. Ther. Res.* **1992**, *51*, 668–672.
62. Dyckner, T.; Wester, P.O. Effect of magnesium on blood pressure. *Br. Med. J.* **1983**, *286*, 1847–1849. [CrossRef]
63. Huo, Y.; Li, J.; Qin, X.; Huang, Y.; Wang, X.; Gottesman, R.F.; Tang, G.; Wang, B.; Chen, D.; He, M.; et al. Efficacy of folic acid therapy in primary prevention of stroke among adults with hypertension in China: The CSPPT randomized clinical trial. *JAMA* **2015**, *313*, 1325–1335. [CrossRef] [PubMed]
64. Mente, A.; O'Donnell, M.J.; Rangarajan, S.; McQueen, M.J.; Poirier, P.; Wielgosz, A.; Morrison, H.; Li, W.; Wang, X.; Di, C.; et al. Association of urinary sodium and potassium excretion with blood pressure. *N. Engl. J. Med.* **2014**, *371*, 601–611. [PubMed]
65. Simopoulos, A.P. Omega-3 fatty acids in health and disease and in growth and development. *Am. J. Clin. Nutr.* **1991**, *54*, 438–463. [CrossRef] [PubMed]
66. Xiong, X.J.; Wang, P.Q.; Li, S.J.; Li, X.K.; Zhang, Y.Q.; Wang, J. Garlic for hypertension: A systematic review and meta-analysis of randomized controlled trials. *Phytomedicine* **2015**, *22*, 352–361. [CrossRef] [PubMed]
67. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; The, P.G. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLOS Med.* **2009**, *6*, e1000097. [CrossRef] [PubMed]
68. Maher, C.G.; Sherrington, C.; Herbert, R.D.; Moseley, A.M.; Elkins, M. Reliability of the PEDro scale for rating quality of randomized controlled trials. *Phys. Ther.* **2003**, *83*, 713–721. [PubMed]
69. Verhagen, A.P.; de Vet, H.C.; de Bie, R.A.; Kessels, A.G.; Boers, M.; Bouter, L.M.; Knipschild, P.G. The Delphi List: A criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by delphi consensus. *J. Clin. Epidemiol.* **1999**, *51*, 1235–1241. [CrossRef]
70. Moraes, W.; Santos, N.S.D.; Aguiar, L.P.; Sousa, L.G.O. Maintenance of exercise training benefits is associated with adequate milk and dairy products intake in elderly hypertensive subjects following detraining. *Einstein* **2017**, *15*, 289–294. [CrossRef] [PubMed]
71. Romero, S.A.; Gagnon, D.; Adams, A.N.; Moralez, G.; Kouda, K.; Jaffery, M.F.; Cramer, M.N.; Crandall, C.G. Folic acid ingestion improves skeletal muscle blood flow during graded handgrip and plantar flexion exercise in aged humans. *Am. J. Physiol.-Heart Circ. Physiol.* **2017**, *313*, H658–H666. [CrossRef] [PubMed]
72. Astorino, T.A.; Martin, B.J.; Schachtsiek, L.; Wong, K. Caffeine ingestion and intense resistance training minimize postexercise hypotension in normotensive and prehypertensive men. *Res. Sports Med.* **2013**, *21*, 52–65. [CrossRef] [PubMed]
73. Lee, C.J.; Kim, J.Y.; Shim, E.; Hong, S.H.; Lee, M.; Jeon, J.Y.; Park, S. The effects of diet alone or in combination with exercise in patients with prehypertension and hypertension: A randomized controlled trial. *Korean Circ. J.* **2018**, *48*, 637–651. [CrossRef] [PubMed]
74. Villani, R.G.; Gornall, J. Short-Term resistance exercise does not alter the hypotensive effect of low-energy dieting. *J. Strength Cond. Res.* **1999**, *13*, 286–288.
75. Sales, A.R.; Silva, B.M.; Neves, F.J.; Rocha, N.G.; Medeiros, R.F.; Castro, R.R.; Nobrega, A.C. Diet and exercise training reduce blood pressure and improve autonomic modulation in women with prehypertension. *Eur. J. Appl. Physiol.* **2012**, *112*, 3369–3378. [CrossRef] [PubMed]
76. Figueroa, A.; Vicil, F.; Sanchez-Gonzalez, M.A.; Wong, A.; Ormsbee, M.J.; Hooshmand, S.; Daggy, B. Effects of diet and/or low-intensity resistance exercise training on arterial stiffness, adiposity, and lean mass in obese postmenopausal women. *Am. J. Hypertens.* **2013**, *26*, 416–423. [CrossRef] [PubMed]
77. Arazi, H.; Samami, N.; Kheirkhah, J.; Taati, B. The effect of three weeks green tea extract consumption on blood pressure, heart rate responses to a single bout resistance exercise in hypertensive women. *High Blood Press. Cardiovasc. Prev.* **2014**, *21*, 213–219. [CrossRef] [PubMed]

78. Wong, A.; Alvarez-Alvarado, S.; Jaime, S.J.; Kinsey, A.W.; Spicer, M.T.; Madzima, T.A.; Figueroa, A. Combined whole-body vibration training and l-citrulline supplementation improves pressure wave reflection in obese postmenopausal women. *Appl. Physiol. Nutr. Metab.* **2016**, *41*, 292–297. [CrossRef] [PubMed]
79. Arazi, H.; Ghiasi, A.; Asgharpoor, S. A comparative study of cardiovascular responses to two rest intervals between circuit resistance exercises in normotensive woman. *Revista Brasileira de Medicina do Esporte* **2013**, *19*, 176–180. [CrossRef]
80. Pescatello, L.S.; MacDonald, H.V.; Lamberti, L.; Johnson, B.T. Exercise for Hypertension: A Prescription Update Integrating Existing Recommendations with Emerging Research. *Curr. Hypertens. Rep.* **2015**, *17*, 87. [CrossRef]
81. Castaneda, C.; Gordon, P.L.; Uhlin, K.L.; Levey, A.S.; Kehayias, J.J.; Dwyer, J.T.; Fielding, R.A.; Roubenoff, R.; Singh, M.F. Resistance Training To Counteract the Catabolism of a Low-Protein Diet in Patients with Chronic Renal Insufficiency: A Randomized, Controlled Trial. *Ann. Intern. Med.* **2001**, *135*, 965–976. [CrossRef] [PubMed]
82. Kannel, W.B. Blood Pressure as a Cardiovascular Risk Factor: Prevention and Treatment. *JAMA* **1996**, *275*, 1571–1576. [CrossRef] [PubMed]
83. Nelson, M.E.; Fiatarone, M.A.; Morganti, C.M.; Trice, I.; Greenberg, R.A.; Evans, W.J. Effects of high-intensity strength training on multiple risk factors for osteoporotic fractures: A randomized controlled trial. *JAMA* **1994**, *272*, 1909–1914. [CrossRef] [PubMed]
84. Skelton, D.A.; Young, A.; Greig, C.A.; Malbut, K.E. Effects of Resistance Training on Strength, Power, and Selected Functional Abilities of Women Aged 75 and Older. *J. Am. Geriatr. Soc.* **1995**, *43*, 1081–1087. [CrossRef] [PubMed]
85. Singh, N.A.; Clements, K.M.; Fiatarone, M.A. A randomized controlled trial of progressive resistance training in depressed elders. *J. Gerontol. Ser. A Biol. Sci. Med. Sci.* **1997**, *52*, M27–M35. [CrossRef] [PubMed]
86. Cornelissen, V.A.; Fagard, R.H.; Coeckelberghs, E.; Vanhees, L. Impact of resistance training on blood pressure and other cardiovascular risk factors: A meta-analysis of randomized, controlled trials. *Hypertension* **2011**, *58*, 950–958. [CrossRef]
87. MacDougall, J.D.; Tuxen, D.; Sale, D.G.; Moroz, J.R.; Sutton, J.R. Arterial blood pressure response to heavy resistance exercise. *J. Appl. Physiol.* **1985**, *58*, 785–790. [CrossRef]
88. Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. The Sixth Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. *Arch. Intern. Med.* **1997**, *157*, 2413–2446. [CrossRef]
89. Reisin, E.; Frohlich, E.D.; Messerli, F.H.; Dreslinski, G.R.; Dunn, F.G.; Jones, M.M.; Batson, H.M. Cardiovascular changes after weight reduction in obesity hypertension. *Ann. Intern. Med.* **1983**, *98*, 315–319. [CrossRef]
90. Langford, H.G.; Blaufox, M.D.; Oberman, A.; Hawkins, C.M.; Curb, J.D.; Cutter, G.R.; Wassertheil-Smoller, S.; Pressel, S.; Babcock, C.; Abernethy, J.D. Dietary therapy slows the return of hypertension after stopping prolonged medication. *JAMA* **1985**, *253*, 657–664. [CrossRef]
91. Stamler, R.; Stamler, J.; Gosch, F.C.; Civinelli, J.; Fishman, J.; McKeever, P.; McDonald, A.; Dyer, A.R. Primary prevention of hypertension by nutritional-hygienic means: Final report of a randomized, controlled trial. *JAMA* **1989**, *262*, 1801–1807. [CrossRef] [PubMed]
92. Stamler, R.; Stamler, J.; Grimm, R.; Gosch, F.C.; Elmer, P.; Dyer, A.; Berman, R.; Fishman, J.; Van Heel, N.; Civinelli, J. Nutritional therapy for high blood pressure: Final report of a four-year randomized controlled trial—the Hypertension Control Program. *JAMA* **1987**, *257*, 1484–1491. [CrossRef] [PubMed]
93. Happonen, P.; Voutilainen, S.; Salonen, J.T. Coffee drinking is dose-dependently related to the risk of acute coronary events in middle-aged men. *J. Nutr.* **2004**, *134*, 2381–2386. [CrossRef] [PubMed]
94. Grgic, J.; Grgic, I.; Pickering, C.; Schoenfeld, B.J.; Bishop, D.J.; Pedisic, Z. Wake up and smell the coffee: Caffeine supplementation and exercise performance—an umbrella review of 21 published meta-analyses. *Br. J. Sports Med.* **2019**. [CrossRef] [PubMed]
95. Glade, M.J. Caffeine—not just a stimulant. *Nutrition* **2010**, *26*, 932–938. [CrossRef]
96. Ali, A.; O'Donnell, J.; Foskett, A.; Rutherfurd-Markwick, K. The influence of caffeine ingestion on strength and power performance in female team-sport players. *J. Int. Soc. Sports Nutr.* **2016**, *13*, 46. [CrossRef] [PubMed]
97. Jee, S.H.; He, J.; Whelton, P.K.; Suh, I.; Klag, M.J. The Effect of Chronic Coffee Drinking on Blood Pressure. *Hypertension* **1999**, *33*, 647–652. [CrossRef]

98. Creager, M.A.; Cooke, J.P.; Mendelsohn, M.E.; Gallagher, S.J.; Coleman, S.M.; Loscalzo, J.; Dzau, V.J. Impaired vasodilation of forearm resistance vessels in hypercholesterolemic humans. *J. Clin. Investig.* **1990**, *86*, 228–234. [CrossRef]
99. Mayhan, W.G. Role of nitric oxide in disruption of the blood-brain barrier during acute hypertension. *Brain Res.* **1995**, *686*, 99–103. [CrossRef]
100. Rector, T.S.; Bank, A.J.; Mullen, K.A.; Tschumperlin, L.K.; Sih, R.; Pillai, K.; Kubo, S.H. Randomized, double-blind, placebo-controlled study of supplemental oral L-arginine in patients with heart failure. *Circulation* **1996**, *93*, 2135–2141. [CrossRef]
101. Yang, Y.-C.; Lu, F.-H.; Wu, J.-S.; Wu, C.-H.; Chang, C.-J. The protective effect of habitual tea consumption on hypertension. *Arch. Intern. Med.* **2004**, *164*, 1534–1540. [CrossRef]
102. Stensvold, I.; Tverdal, A.; Solvoll, K.; Foss, O.P. Tea consumption. Relationship to cholesterol, blood pressure, and coronary and total mortality. *Prev. Med.* **1992**, *21*, 546–553. [CrossRef]
103. Wakabayashi, K.; Kono, S.; Shinchi, K.; Honjo, S.; Todoroki, I.; Sakurai, Y.; Umeda, T.; Imanishi, K.; Yoshizawa, N. Habitual coffee consumption and blood pressure: A study of self-defense officials in Japan. *Eur. J. Epidemiol.* **1998**, *14*, 669–673. [CrossRef] [PubMed]
104. Álvarez, E.; Campos-Toimil, M.; Justiniano-Basaran, H.; Lugnier, C.; Orallo, F. Study of the mechanisms involved in the vasorelaxation induced by (−)-epigallocatechin-3-gallate in rat aorta. *Br. J. Pharmacol.* **2006**, *147*, 269–280. [CrossRef]
105. Meschia, J.F.; Bushnell, C.; Boden-Albala, B.; Braun, L.T.; Bravata, D.M.; Chaturvedi, S.; Creager, M.A.; Eckel, R.H.; Elkind, M.S.V.; Fornage, M.; et al. Guidelines for the primary prevention of stroke: A statement for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* **2014**, *45*, 3754–3832. [CrossRef] [PubMed]
106. Armitage, J.M.; Bowman, L.; Clarke, R.J.; Wallendszus, K.; Bulbulia, R.; Rahimi, K.; Haynes, R.; Parish, S.; Sleight, P.; Peto, R. Effects of homocysteine-lowering with folic acid plus vitamin B12 vs placebo on mortality and major morbidity in myocardial infarction survivors: A randomized trial. *JAMA* **2010**, *303*, 2486–2494.
107. Lonn, E. Heart Outcomes Prevention Evaluation (HOPE) 2 Investigators. Homocysteine lowering with folic acid and B vitamins in vascular disease. *Nat. Clin. Pract. Cardiovasc. Med.* **2006**, *3*, 414–415.
108. Toole, J.F.; Malinow, M.R.; Chambless, L.E.; Spence, J.D.; Pettigrew, L.C.; Howard, V.J.; Sides, E.G.; Wang, C.-H.; Stampfer, M. Lowering homocysteine in patients with ischemic stroke to prevent recurrent stroke, myocardial infarction, and death: The Vitamin Intervention for Stroke Prevention (VISP) randomized controlled trial. *JAMA* **2004**, *291*, 565–575. [CrossRef]
109. Bønaa, K.H.; Njølstad, I.; Ueland, P.M.; Schirmer, H.; Tverdal, A.; Steigen, T.; Wang, H.; Nordrehaug, J.E.; Arnesen, E.; Rasmussen, K. Homocysteine lowering and cardiovascular events after acute myocardial infarction. *N. Engl. J. Med.* **2006**, *354*, 1578–1588. [CrossRef]
110. Hernández-Díaz, S.; Werler, M.M.; Louik, C.; Mitchell, A.A. Risk of gestational hypertension in relation to folic acid supplementation during pregnancy. *Am. J. Epidemiol.* **2002**, *156*, 806–812. [CrossRef]
111. Najjar, S.S.; Scuteri, A.; Shetty, V.; Wright, J.G.; Muller, D.C.; Fleg, J.L.; Spurgeon, H.P.; Ferrucci, L.; Lakatta, E.G. Pulse wave velocity is an independent predictor of the longitudinal increase in systolic blood pressure and of incident hypertension in the Baltimore Longitudinal Study of Aging. *J. Am. Coll. Cardiol.* **2008**, *51*, 1377–1383. [CrossRef] [PubMed]

Journal of
Clinical Medicine

MDPI

Review

The Efficacy and Safety of Doripenem in the Treatment of Acute Bacterial Infections—A Systemic Review and Meta-Analysis of Randomized Controlled Trials

Chih-Cheng Lai [1], I-Ling Cheng [2], Yu-Hung Chen [2] and Hung-Jen Tang [3,*]

[1] Department of Intensive Care Medicine, Chi Mei Medical Center, Liouying 73657, Taiwan
[2] Department of Pharmacy, Chi Mei Medical Center, Liouying 73657, Taiwan
[3] Department of Medicine, Chi Mei Medical Center, Tainan 71004, Taiwan
* Correspondence: 8409d1@gmail.com; Tel.: +886-6-2812811

Received: 9 May 2019; Accepted: 1 July 2019; Published: 2 July 2019

Abstract: This study aims to assess the efficacy and safety of doripenem on treating patients with acute bacterial infections. The Pubmed, Embase, and Cochrane databases were searched up to April 2019. Only randomized clinical trials comparing doripenem and other comparators for the treatment of acute bacterial infection were included. The primary outcome was the clinical success rate and the secondary outcomes were microbiological eradication rate and risk of adverse events. Eight randomized controlled trials (RCTs) were included. Overall, doripenem had a similar clinical success rate with comparators (odds ratio [OR], 1.15; 95% CI, 0.79–1.66, I2 = 58%). Similar clinical success rates were noted between doripenem and comparators for pneumonia (OR, 0.84; 95% CI, 0.46–1.53, I^2 = 72%) and for intra-abdominal infections (OR, 1.00; 95% CI, 0.57–1.72). For complicated urinary tract infection, doripenem was associated with higher success rate than comparators (OR, 1.89, 95% CI, 1.13–3.17, I^2 = 0%). The pool analysis comparing doripenem and other carbapenems showed no significant differences between each other (OR, 0.96, 95% CI, 0.59–1.58, I^2 = 63%). Doripenem also had a similar microbiological eradication rate with comparators (OR, 1.08; 95% CI, 0.86–1.36, I^2 = 0%). Finally, doripenem had a similar risk of treatment-emergent adverse events as comparators (OR, 0.98; 95% CI, 0.83–1.17, I^2 = 33%). In conclusion, the clinical efficacy of doripenem is as high as that of the comparator drugs in the treatment of acute bacterial infection; furthermore, this antibiotic is as well tolerated as the comparators.

Keywords: doripenem; acute bacterial infection; pneumonia; intra-abdominal infection; complicated urinary tract infection

1. Introduction

Carbapenems, including imipenem and meropenem, remain the mainstay of treatment for hospital-acquired infections, especially for the multidrug-resistant organism associated infections [1]. Doripenem is another important carbapenem, and has excellent bactericidal activity against most nosocomial pathogens according to several in vitro studies [2–5]. A global surveillance showed that doripenem was at least two-fold more potent in in vitro activities than imipenem and meropenem against *Pseudomonas aeruginosa*—an important nosocomial pathogen [3]. For another notorious pathogen—*Acinetobacter baumannii*, doripenem displayed comparable in vitro activities to imipenem and meropenem [4]. Clinically, doripenem is approved for the treatment of patients with complicated intra-abdominal infection (cIAI), complicated urinary tract infection (cUTI) and pyelonephritis, and healthcare-associated pneumonia (HAP) including ventilator-associated pneumonia (VAP) in Europe and in other countries, other than United States. Although Qu et al. [6] conducted a meta-analysis of

doripenem for treating bacterial infections in 2015, only six clinical trials were enrolled and the number of patients was limited. Since then, two more studies investigating the efficacy of doripenem in comparison with other comparators were reported [7,8]. In Wagenlehner et al.'s study [7], 1033 randomized patients were enrolled, and they did the comparison between doripenem and ceftazidime-avibactam for the treatment of cUTI. In Oyake et al.'s study [8], they compared the empirical use of doripenem versus meropenem for febrile neutropenia in patients with acute leukemia. These two studies provided more patients and different types of infections compared to previse meta-analysis [6]. Therefore, we could conduct a comprehensive review and updated meta-analysis to assess the efficacy and safety of doripenem on treating patients with acute bacterial infections in comparison with other antibiotics, especially imipenem and meropenem.

2. Methods

2.1. Study Search and Selection

Studies were identified by a systematic review of the literature in the PubMed, Embase, and Cochrane databases until April 2019 using the following search terms—"doripenem," "infection," and "randomized" (Appendix A). Studies were considered eligible for inclusion if they directly compared the clinical effectiveness of doripenem with other antimicrobial agents in the treatment of adult patients with acute bacterial infections. Studies were excluded if they focused on in vitro activity, or pharmacokinetic-pharmacodynamic assessment. The articles of all languages of publication could be included. Two reviewers (I.-L.C. and Y.-H.C.) searched and examined publications independently. When they had disagreement, the third author (C.-C.L.) resolved the issue in time. The following data including year of publication, study design, type of infections, patients' demographic features, antimicrobial regimens, clinical and microbiological outcomes, and adverse events were extracted from every included study.

2.2. Definitions and Outcomes

The primary outcome was overall clinical success with resolution of clinical signs and symptoms of acute bacterial infection, or recovery to the pretreatment state at the end of treatment. Secondary outcomes included the microbiological eradication rate and adverse events. A microbiological eradication was defined as eradication (the baseline pathogen was absent) and presumed eradication (if an adequate source specimen was not available to culture, but the patient was assessed as clinically cured). Treatment-emergent adverse events were recorded, irrespective of causality. In addition, the risk of discontinuing due to adverse event and the incidence of serious adverse events, and some common events, including diarrhea, nausea, headache, constipation, and seizure were recorded.

2.3. Data Analysis

This study used the Cochrane risk of bias assessment tool to assess the quality of enrolled randomized controlled trials (RCTs) and the risk of bias [9]. The software Review Manager, version 5.3, was used to conduct the statistical analyses. The degree of heterogeneity was evaluated with the Q statistic generated from the χ^2 test. The proportion of statistical heterogeneity was assessed by the I^2 measure. Heterogeneity was considered significant when the p-value was less than 0.10 or the I^2 was more than 50%. The random-effects model was used when the data were significantly heterogeneous, and the fixed-effect model was used when the data were homogenous. Pooled odds ratios (OR) and 95% confidence intervals (CI) were calculated for outcome analyses. Sensitivity analysis was performed to ensure that the findings were not significantly affected by any individual study

3. Results

3.1. Study Selection and Characteristics

The search program yielded 499 references, including 263 from Pubmed, 170 from Embase, and 66 from Cochrane database. Then, 258 articles were screened after excluding 241 duplicated articles. Finally, a total of eight RCTs [7,8,10–15] fulfilling the inclusion criteria were included in this meta-analysis (Figure 1). All of studies were designed to compare the clinical efficacy and safety of doripenem with other antibiotics for patients with acute bacterial infection (Table 1) [7,8,10–15]. During the initial enrollment, doripenem and comparators were applied to 1736 and 1763 patients, respectively. Six studies [7,10–13,15] of them were multicenter studies. Three studies [10–12] focused on pneumonia, including two [12,16] on ventilator-associated pneumonia and one [10] on nosocomial pneumonia. Two studies focused on complicated urinary tract infections (cUTI) [7,13] and intra-abdominal infections (IAI) [14,15]. Only one study investigated febrile neutropenia [8]. Five studies [8,11,12,14,15] compared doripenem with other carbapenems including imipenem in three studies [11,12,14] and meropenem in two studies [8,15]. The regimen of doripenem was 1 g every eight hours in two studies [8,11] and 500 mg every eight hours in the other six studies [7,10,12–15]. For the two studies using double dose of doripenem (1 g every eight hour), the study drug (doripenem or meropenem) was used for at least five days in one study [8] and another one [11] compared seven-day doripenem versus 10-day imipenem-cilastatin. Figure 2 shows the analyses of risk of bias.

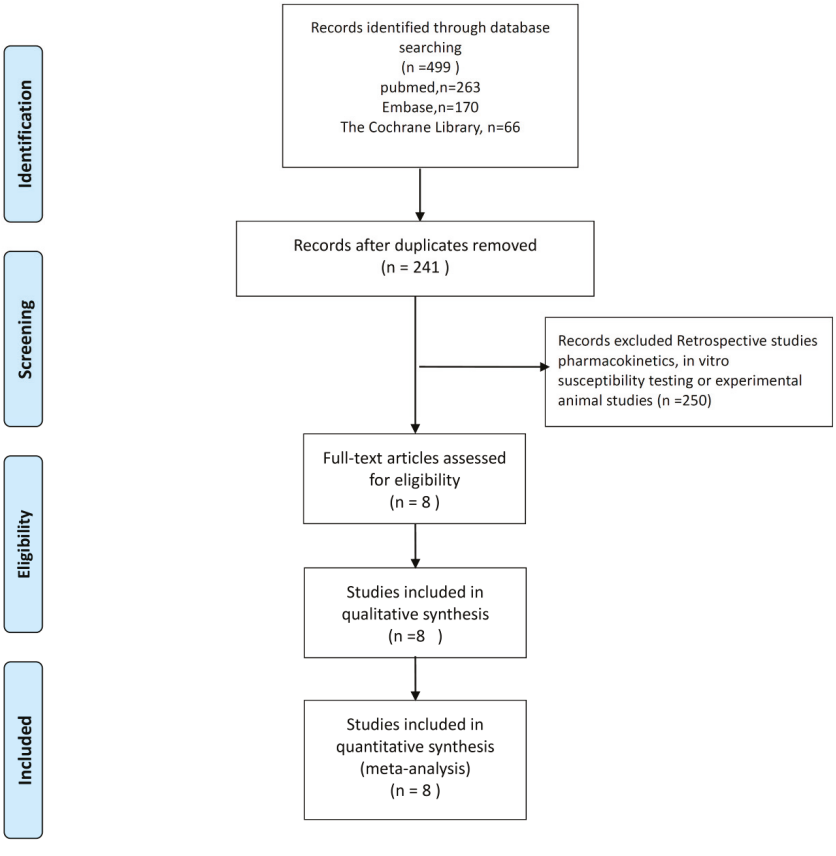


Figure 1. Flowchart of the study selection process.

Table 1. Characteristics of included studies.

Reference	RCT Study Design	Duration	Study Population	No. of Patients		Age of the Patients		Dose Regimen	
				Doripenem	Comparator	Doripenem	Comparator	Doripenem	Comparator
[10]	Randomized, open-label, multicenter trial	2004–2006	Nosocomial pneumonia	225	223	57.5	59.3	Doripenem 500 mg every 8 h	Piperacillin/tazobactam 4.5 g every 6 h
[15]	Prospective, multicenter, randomized, double-blind	2004–2006	Complicated intra-abdominal infection	237	239	46.9	46.4	Doripenem 500 mg every 8 h	Meropenem 1.0 g every 8 h
[12]	Prospective, multicenter, randomized, open-label trial	2004–2006	Ventilator-associated pneumonia	262	263	50.7	50.3	Doripenem 500 mg every 8 h	Imipenem/cilastatin 1 g every 8 h or 500 mg every 6 h
[13]	Prospective, multicenter, double-blind trial	2003–2006	Complicated UTI	377	376	51.2	51.1	Doripenem 500 mg every 8 h	Levofloxacin 250 mg everyday
[11]	Randomized, double-blind, multicenter trial	2008–2011	Ventilator-associated pneumonia	115	112	57.5	54.6	Doripenem 1 g every 8 h	Imipenem/cilastatin 1 g every 8 h
[14]	Randomized, open-label trial	2010–2013	Moderate or severe acute cholangitis or cholecystitis	62	65	74	73	Doripenem 500 mg every 8 h	Imipenem/cilastatin 500 mg every 8 h
[7]	Randomized, multicenter, double-blind, trials	2012–2014	Complicated UTI	393	417	53.3	51.4	Doripenem 500 mg every 8 h	Ceftazidime-avibactam 2000 mg/500 mg every 8 h
[8]	Randomized, open-label prospective trial	2011–2013	Febrile neutropenia in patients with acute leukemia or MDS-refractory anemia with excess blasts	65	68	57	56	Doripenem 1 g every 8 h	Meropenem 1.0 g every 8 h

MDS, myelodysplastic syndrome; UTI, urinary tract infection; RCT, randomized controlled trial.

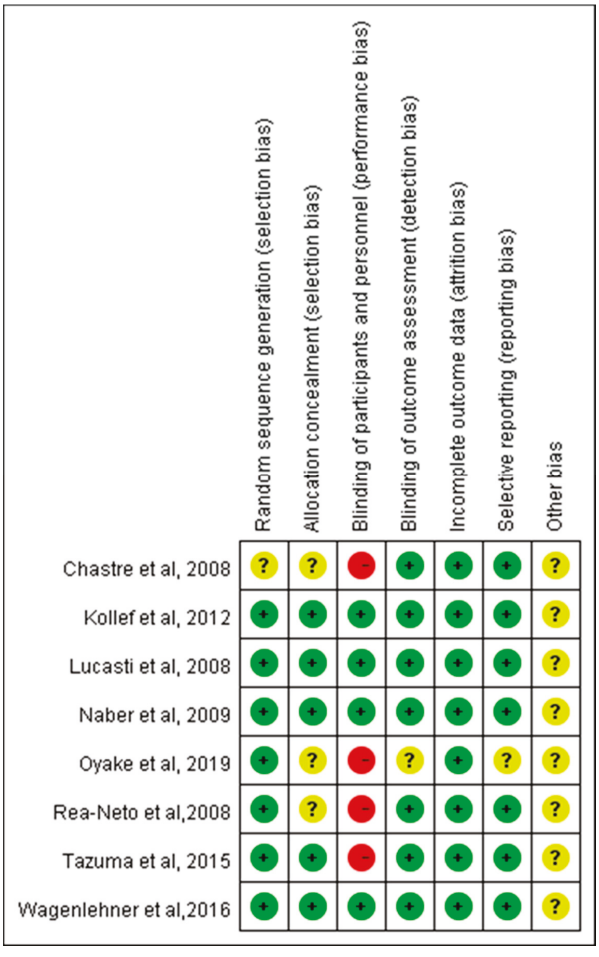


Figure 2. Risk of bias per study and domain.

3.2. Clinical Success

Overall, doripenem had a similar clinical success rate with comparators (OR, 1.15; 95% CI, 0.79–1.66, I^2 = 58%, Figure 3). Sensitivity analysis after randomly deleting an individual study each time to reflect the influence of the single data set to the pooled OR showed similar findings in most occasions. There was only one exception, when we deleted Kollef et al.'s study [11], doripenem showed better clinical success rate than other comparators in the pool analysis of the remaining seven studies [7,8,10,12–15] (OR, 1.33; 95% CI, 1.03–1.72, I^2 = 0%). In the different subgroup of patients with pneumonia, cUTI, and intra-abdominal infection, similar clinical success rates were noted between two different regimens for pneumonia (OR, 0.84; 95% CI, 0.46–1.53, I^2 = 72%) and for IAI (OR, 1.00; 95% CI, 0.57–1.72). For cUTI, doripenem was associated with a higher success rate than comparators (OR, 1.89, 95% CI, 1.13–3.17, I^2 = 0%). Three studies [11,12,14] compared the effect of doripenem and imipenem, and there was no difference in terms of clinical success rate between these two regimens (OR, 0.76; 95% CI, 0.38–1.55, I^2 = 66%). Two studies [8,15] compared doripenem and meropenem, their clinical success rates were similar (OR, 1.31, 95% CI, 0.75–2.28, I^2 = 34%). The pool analysis of these

five studies comparing doripenem and other carbapenems showed no significant differences between each other (OR, 0.96, 95% CI, 0.59–1.58, I^2 = 63%).

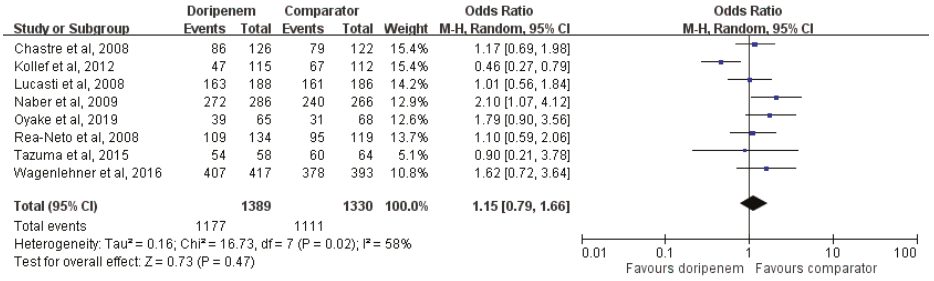


Figure 3. Overall clinical success rates of doripenem and comparators in the treatment of acute bacterial infections.

3.3. Microbiological Eradication

Only six studies [7,10,12–15] reported the data of microbiological eradication rate, and the pool analysis showed that doripenem had a similar microbiological eradication rate with comparators (OR, 1.08; 95% CI, 0.86–1.36, I^2 = 0%, Figure 4). Sensitivity analysis showed similar results. In the different subgroup of patients with pneumonia and IAI, similar microbiological eradication rates were found for both regimens (for pneumonia, OR, 1.25; 95% CI, 0.79–1.97, I^2 = 0%; for IAI, OR, 1.04; 95% CI, 0.49–2.17, I^2 = 54%). While comparing doripenem and other carbapenems in the pool analysis of four studies [7,12,14,15], the microbiological eradication rates were similar between these two regimens (OR, 1.13; 95% CI, 0.85–1.51, I^2 = 0%).

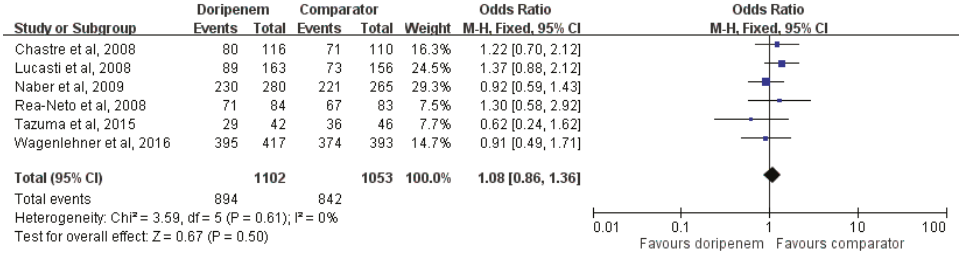


Figure 4. Overall microbiological eradication rates of doripenem and comparators in the treatment of acute bacterial infections.

3.4. Adverse Events

Six studies [7,8,11,13–15] reported the incidence of treatment-emergent adverse events, the doripenem had a similar risk with other antibiotics (OR, 0.98; 95% CI, 0.83–1.17, I^2 = 33%, Figure 5). Serious adverse events were reported in six studies [7,10,12–15], the overall incidence was similar between doripenem and other antibiotics (OR, 1.06; 95% CI, 0.85–1.31, I^2 = 43%). Six studies [7,10,12–15] reported the risk of discontinuing drug due to adverse event, the risk was similar between doripenem and comparators (OR, 0.75, 95% CI, 0.35–1.61, I^2 = 61%). Regarding some common adverse events, doripenem was associated with the similar risk as comparators in terms of diarrhea (OR, 0.91, 95% CI, 0.64–1.28, I^2 = 0%) in the pool analysis of eight studies [7,8,10–15], nausea (OR, 0.93, 95% CI, 0.45–1.93, I^2 = 62%) among five studies [11–15], headache (OR, 1.10, 95% CI, 0.82–1.48, I^2 = 0%) among three studies [13–15], and constipation (OR, 0.96, 95% CI, 0.61–1.52, I^2 = 0%) among three studies [11,13,14]. In the pooled analysis of four studies [10,12,13,15] that reported the risk of seizure, doripenem was

associated with a similar lower risk as comparators (OR, 0.37, 95% CI, 0.15–0.92, I^2 = 0%). Moreover, no seizure attack was reported to be related to doripenem in these four studies [10,12,13,15].

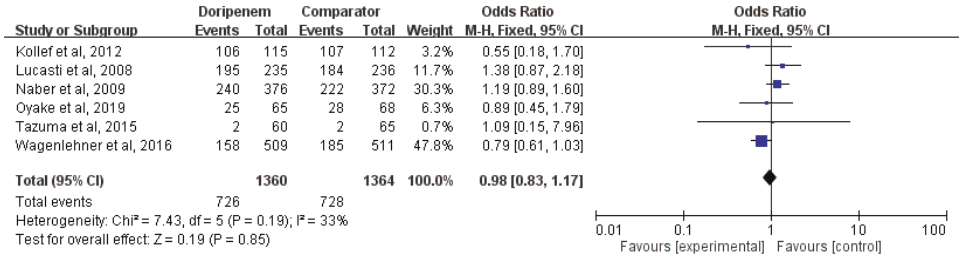


Figure 5. Risk of treatment-emergent adverse events of doripenem and comparators in the treatment of acute bacterial infections.

4. Discussion

This meta-analysis based on eight RCTs found that doripenem had a similar clinical success rate of treating acute bacterial infections with other comparators. The similar efficacy in terms of clinical response and microbiological eradication was found between doripenem and other carbapenems, including meropenem and imipenem. In addition, this result was not affected by the different types of infections—pneumonia, cUTI, or IAIs. Even for several specific types of infection—cholangitis, cholecystitis, appendicitis, lower urinary tract infection, and acute pyelonephritis—no statistical differences in terms of clinical efficacy was found between doripenem and comparators in the included studies [13–15]. In fact, in addition to Kollef et al.'s study [11], that showed doripenem was found to have non-significant higher rates of clinical failure and mortality compared to imipenem [7,10,12–15]. The difference between Kollef et al.'s study [11] and the other seven studies [7,10,12–15] may be explained by the fact that Kollef et al. compared a fixed seven-day course of doripenem with a fixed 10-day course of imipenem-cilastatin for treating VAP. Seven days of antibiotic treatment may have been too short for the patients with VAP, so the clinical outcome of VAP treated with a seven-day course of doripenem was worse than with a 10-day course of imipenem-cilastatin. In this meta-analysis, while we did sensitivity analysis after deleting this negative study [11] for doripenem, we found that the pooled analysis of the other seven studies [7,8,10,12–15] showed that doripenem was associated with better clinical outcome than comparators. Although this finding hints that the effect of doripenem may be better, or at least as good as, other antimicrobial agents in the treatment of acute bacterial infections, if doripenem can be used as long as the comparators, we still need further study to confirm this issue. Before that, the findings of this meta-analysis indicate that the clinical efficacy of doripenem is not inferior to other antimicrobial agents in the treatment of acute bacterial infections. Finally, several studies [16–18] demonstrated that doripenem was associated with lower medical resource utilization and hospital cost in the treatment of HAP and VAP versus comparators, including imipenem. Overall, doripenem could be both a life- and cost-saving antibiotic and could be recommended as the appropriate antibiotic in the treatment of acute bacterial infections, including pneumonia, cUTI, and cIAI.

In this meta-analysis, we also compared the microbiological response of doripenem with other antibiotics for acute bacterial infection. Overall, we found the microbiological eradication rates were similar between doripenem and comparators. Moreover, a similar trend was noted in the sensitivity analysis and subgroup analysis of pneumonia and IAIs. Finally, doripenem was comparable to other carbapenems, including imipenem and meropenem, in terms of microbiological eradication rate in the subgroup analysis. All these findings may be well explained by previous in vitro studies [3,4,19–21] that showed doripenem had a greater or similar in vitro activity against bacteria, including multi-drug resistant organisms. In this meta-analysis, we did not assess we did not evaluate the association between

in vitro activity and the in vivo response of different organisms, especially for antibiotic-resistant pathogens, because the associated information was limited. However, this meta-analysis demonstrates that doripenem is comparable to other antimicrobial agents in both the clinical and microbiological responses of treating acute bacterial infections.

In addition to the assessment of clinical efficacy and microbiological eradication, the safety issue is another important concern in the treatment of acute bacterial infection by doripenem. In this analysis, the risks of overall treatment-emergent adverse effects, common adverse effects (diarrhea, nausea, headache and constipation), serious adverse effects, and the risk of discontinuing the drug due to adverse effects were similar between doripenem and comparators. Seizure is another important concern for patients using carbapenems. In this meta-analysis, four studies [10,12,13,15] reported the incidence of seizure, and the doripenem group had a lower risk of seizure than comparators. Moreover, although six seizure events were reported in this meta-analysis, all these events occurred in patients with underlying risk factors and were not clearly related to doripenem. Therefore, all these findings indicate that doripenem may be as safe as conventional regimens in the treatment of acute bacterial infections.

This meta-analysis has several limitations. First, we did not evaluate the effect of doripenem and comparators against specific organisms in each type of bacterial infection and the confounding effect of the antibiotic resistance of these pathogens. Besides, the immune status and the age effect were not assessed in this meta-analysis due to limited information. Second, the use of doripenem for treating pneumonia remains a serious concern due to the negative findings of Kolleff et al.'s study [11] that showed a shorter course (seven days) of doripenem was associated with a worse outcome than a longer course (10 days) of imipenem for patients with VAP. However, doripenem was commonly used for treating pneumonia in many countries [22], and several studies [10,12,23–25] showed the clinical outcomes of pneumonia treated by doripenem were favorable. In the subgroup analysis of this meta-analysis, we found the clinical and microbiological responses of doripenem for treating pneumonia were as good as comparators. But, as only three RCTs [10–12] focusing on pneumonia were enrolled in this meta-analysis, the number of studies is limited, thus further study is warranted to clarify this issue.

In conclusion, based on the analysis of eight RCTs, no differences in terms of clinical success and microbiological eradication rates were found between doripenem and comparators in the treatment of acute bacterial infections. Moreover, doripenem was well tolerated and had comparable safety profiles to other antimicrobial agents.

Author Contributions: Conceptualization, I.-L.C., Y.-H.C. and C.-C.L.; formal analysis, I.-L.C. and Y.-H.C.; writing—original draft, C.-C.L.; writing—review and editing, H.-J.T.

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

Appendix A : List of Terms of the Search Strategy

Pubmed
1. "doripenem" [MeSH Terms]
2. "doripenem" [All Fields]
3. 1 OR 2
4. "infection" [MeSH Terms]
5. "infection" [All Fields]
6. 4 OR 5
7. "randomized" [All Fields]
8. "randomised" [All Fields]
9. 7 OR 8
10. 3 AND 6 AND 9

Embase

1. "doripenem"/exp
2. "doripenem"
3. 1 OR 2
4. "infection"
5. "randomized"
6. "randomised"
7. 5 OR 6
8. 3 AND 4 AND 7

Cochrane

1. doripenem
2. infection
3. #1 AND #2

References

1. Kalil, A.C.; Metersky, M.L.; Klompas, M.; Muscedere, J.; Sweeney, D.A.; Palmer, L.B.; Napolitano, L.M.; O'Grady, N.P.; Bartlett, J.G.; Carratalà, J.; et al. Management of adults with hospital-acquired and ventilator-associated pneumonia: 2016 Clinical practice guidelines by the Infectious Diseases Society of America and the American Thoracic Society. *Clin. Infect. Dis.* **2016**, *63*, e61–e111. [CrossRef] [PubMed]
2. Mazzei, T. The pharmacokinetics and pharmacodynamics of the carbapanemes: Focus on doripenem. *J. Chemother.* **2010**, *22*, 219–225. [CrossRef] [PubMed]
3. Castanheira, M.; Jones, R.N.; Livermore, D.M. Antimicrobial activities of doripenem and other carbapenems against *Pseudomonas aeruginosa*, other nonfermentative bacilli, and *Aeromonas* spp. *Diagn. Microbiol. Infect. Dis.* **2009**, *63*, 426–433. [CrossRef] [PubMed]
4. Douraghi, M.; Ghalavand, Z.; Nateghi Rostami, M.; Zeraati, H.; Aliramezani, A.; Rahbar, M.; Mohammadzadeh, M.; Ghourchian, S.; Boroumand, M.A.; Abdollahi, A. Comparative *in vitro* activity of carbapenems against clinical isolates of *Acinetobacter baumannii*. *J. Appl. Microbiol.* **2016**, *121*, 401–407. [CrossRef] [PubMed]
5. Li, Y.; Lv, Y.; Xue, F.; Zheng, B.; Liu, J.; Zhang, J. Antimicrobial resistance surveillance of doripenem in China. *J. Antibiot. Tokyo* **2015**, *68*, 496–500. [CrossRef] [PubMed]
6. Qu, X.Y.; Hu, T.T.; Zhou, W. A meta-analysis of efficacy and safety of doripenem for treating bacterial infections. *Braz. J. Infect. Dis.* **2015**, *19*, 156–162. [CrossRef] [PubMed]
7. Wagenlehner, F.M.; Sobel, J.D.; Newell, P.; Armstrong, J.; Huang, X.; Stone, G.G.; Yates, K.; Gasink, L.B. Ceftazidime-avibactam versus doripenem for the treatment of complicated urinary tract infections, including acute pyelonephritis: RECAPTURE, a phase 3 randomized trial program. *Clin. Infect. Dis.* **2016**, *63*, 754–762. [CrossRef] [PubMed]
8. Oyake, T.; Takemasa-Fujisawa, Y.; Sugawara, N.; Mine, T.; Tsukushi, Y.; Hanamura, I.; Fujishima, Y.; Aoki, Y.; Kowata, S.; Ito, S.; et al. Doripenem versus meropenem as first-line empiric therapy of febrile neutropenia in patients with acute leukemia: A prospective, randomized study. *Ann. Hematol.* **2019**, *98*, 1209–1216. [CrossRef] [PubMed]
9. Higgins, J.P.; Altman, D.G.; Gotzsche, P.C.; Juni, P.; Moher, D.; Oxman, A.D.; Savovic, J.; Schulz, K.F.; Weeks, L.; Sterne, J.A.; et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* **2011**, *343*, d5928. [CrossRef]
10. Rea-Neto, A.; Niederman, M.; Lobo, S.M.; Schroeder, E.; Lee, M.; Kaniga, K.; Ketter, N.; Prokocimer, P.; Friedland, I. Efficacy and safety of doripenem versus piperacillin/tazobactam in nosocomial pneumonia: A randomized, open-label, multicenter study. *Curr. Med. Res. Opin.* **2008**, *24*, 2113–2126. [CrossRef] [PubMed]
11. Kollef, M.H.; Chastre, J.; Clavel, M.; Restrepo, M.I.; Michiels, B.; Kaniga, K.; Cirillo, I.; Kimko, H.; Redman, R. A randomized trial of 7-day doripenem versus 10-day imipenem-cilastatin for ventilator-associated pneumonia. *Crit. Care* **2012**, *16*, R218. [CrossRef] [PubMed]
12. Chastre, J.; Wunderink, R.; Prokocimer, P.; Lee, M.; Kaniga, K.; Friedland, I. Efficacy and safety of intravenous infusion of doripenem versus imipenem in ventilator-associated pneumonia: A multicenter, randomized study. *Crit. Care Med.* **2008**, *36*, 1089–1096. [CrossRef] [PubMed]

13. Naber, K.G.; Llorens, L.; Kaniga, K.; Kotey, P.; Hedrich, D.; Redman, R. Intravenous doripenem at 500 milligrams versus levofloxacin at 250 milligrams, with an option to switch to oral therapy, for treatment of complicated lower urinary tract infection and pyelonephritis. *Antimicrob. Agents Chemother.* **2009**, *53*, 3782–3792. [CrossRef] [PubMed]
14. Tazuma, S.; Igarashi, Y.; Inui, K.; Ohara, H.; Tsuyuguchi, T.; Ryozawa, S. Clinical efficacy of intravenous doripenem in patients with acute biliary tract infection: A multicenter, randomized, controlled trial with imipenem/cilastatin as comparator. *J. Gastroenterol.* **2015**, *50*, 221–229. [CrossRef]
15. Lucasti, C.; Jasovich, A.; Umeh, O.; Jiang, J.; Kaniga, K.; Friedland, I. Efficacy and tolerability of IV doripenem versus meropenem in adults with complicated intra-abdominal infection: A phase III, prospective, multicenter, randomized, double-blind, noninferiority study. *Clin. Ther.* **2008**, *30*, 868–883. [CrossRef] [PubMed]
16. Kollef, M.H.; Nathwani, D.; Merchant, S.; Gast, C.; Quintana, A.; Ketter, N. Medical resource utilization among patients with ventilator-associated pneumonia: Pooled analysis of randomized studies of doripenem versus comparators. *Crit. Care* **2010**, *14*, R84. [CrossRef]
17. Kongnakorn, T.; Mwamburi, M.; Merchant, S.; Akhras, K.; Caro, J.J.; Nathwani, D. Economic evaluation of doripenem for the treatment of nosocomial pneumonia in the US: Discrete event simulation. *Curr. Med. Res. Opin.* **2010**, *26*, 17–24. [CrossRef] [PubMed]
18. Zilberberg, M.D.; Mody, S.H.; Chen, J.; Shorr, A.F. Cost-effectiveness model of empiric doripenem compared with imipenem-cilastatin in ventilator-associated pneumonia. *Surg. Infect. Larchmt* **2010**, *11*, 409–417. [CrossRef]
19. Drzewiecki, A.; Bulanda, M.; Talaga, K.; Sodo, A.; Adamski, P. Comparison of *in vitro* activity of doripenem, imipenem and meropenem against clinical isolates of *Enterobacteriaceae*, *Pseudomonas* and *Acinetobacter* in Southern Poland. *Pol. Przegl. Chir.* **2012**, *84*, 449–453.
20. Firsov, A.A.; Gilbert, D.; Greer, K.; Portnoy, Y.A.; Zinner, S.H. Comparative pharmacodynamics and antimutant potentials of doripenem and imipenem with ciprofloxacin-resistant *Pseudomonas aeruginosa* in an *in vitro* model. *Antimicrob. Agents Chemother.* **2012**, *56*, 1223–1228. [CrossRef]
21. Wali, N.; Mirza, I.A. Comparative In Vitro Efficacy of doripenem and imipenem against multi-drug resistant *Pseudomonas aeruginosa*. *J. Coll. Physicians Surg. Pak.* **2016**, *26*, 297–301. [PubMed]
22. Mustafa, M.; Chan, W.M.; Lee, C.; Harijanto, E.; Loo, C.M.; Van Kinh, N.; Anh, N.D.; Garcia, J. A PROspective study on the Usage patterns of Doripenem in the Asia-Pacific region (PROUD study). *Int. J. Antimicrob. Agents* **2014**, *43*, 353–360. [CrossRef] [PubMed]
23. Chao, C.M.; Chen, C.C.; Huang, H.L.; Chuang, Y.C.; Lai, C.C.; Tang, H.J. Clinical experience of patients receiving doripenem-containing regimens for the treatment of healthcare-associated infections. *PLoS ONE* **2016**, *11*, e0167522. [CrossRef] [PubMed]
24. Luyt, C.E.; Aubry, A.; Lu, Q.; Micaelo, M.; Brechot, N.; Brossier, F.; Brisson, H.; Rouby, J.J.; Trouillet, J.L.; Combes, A.; et al. Imipenem, meropenem, or doripenem to treat patients with *Pseudomonas aeruginosa* ventilator-associated pneumonia. *Antimicrob. Agents Chemother.* **2014**, *58*, 1372–1380. [CrossRef] [PubMed]
25. Muscedere, J.G.; Day, A.; Heyland, D.K. Mortality, attributable mortality, and clinical events as end points for clinical trials of ventilator-associated pneumonia and hospital-acquired pneumonia. *Clin. Infect. Dis.* **2010**, *51* (Suppl. 1), S120–S125. [CrossRef] [PubMed]

Journal of
Clinical Medicine

MDPI

Review

A Scoping Review of The Efficacy of Virtual Reality and Exergaming on Patients of Musculoskeletal System Disorder

Hui-Ting Lin [1], Yen-I Li [1], Wen-Pin Hu [2], Chun-Cheng Huang [3] and Yi-Chun Du [4,*]

[1] Department of Physical Therapy, I-Shou University No. 8, Yida Road, Yan-chao District, Kaohsiung 82445, Taiwan; huitinglin@isu.edu.tw (H.-T.L.); ssttaarrtt7616274@gmail.com (Y.-I.L.)
[2] Department of Bioinformatics and Medical Engineering, Asia University. 500, Lioufeng Road, Wufeng, Taichung 41354, Taiwan; wenpinhu@asia.edu.tw
[3] Department of Rehabilitation, E-DA Hospital, No.1, Yida Road, Yan-chao District, Kaohsiung 82445, Taiwan; ed109049@edah.org.tw
[4] Department of Electrical Engineering, Southern Taiwan University of Science and Technology, No. 1, Nan-Tai Street, Yungkang District, Tainan 71005, Taiwan
* Correspondence: terrydu@stust.edu.tw

Received: 3 April 2019; Accepted: 24 May 2019; Published: 4 June 2019

Abstract: To assess the effects of virtual reality on patients with musculoskeletal disorders by means of a scoping review of randomized controlled trials (RCTs). The databases included PubMed, IEEE, and the MEDLINE database. Articles involving RCTs with higher than five points on the Physiotherapy Evidence Database (PEDro) scale were reviewed for suitability and inclusion. The methodological quality of the included RCT was evaluated using the PEDro scale. The three reviewers extracted relevant information from the included studies. Fourteen RCT articles were included. When compared with simple usual care or other forms of treatment, there was significant pain relief, increased functional capacity, reduced symptoms of the disorder, and increased joint angles for the virtual reality treatment of chronic musculoskeletal disorders. Furthermore, burn patients with acute pain were able to experience a significant therapeutic effect on pain relief. However, virtual reality treatment of patients with non-chronic pain such as total knee replacement, ankle sprains, as well as those who went through very short virtual reality treatments, did not show a significant difference in parameters, as compared with simple usual care and other forms of treatment. Current evidence supports VR treatment as having a significant effect on pain relief, increased joint mobility, or motor function of patients with chronic musculoskeletal disorders. VR seems quite effective in relieving the pain of patients with acute burns as well.

Keywords: virtual reality; musculoskeletal disorders; randomized controlled tria

1. Introduction

Virtual reality (VR) of players using body movement to interact with a computer is a new form of treatment in rehabilitation settings. It generates a virtual world in three-dimensional space through a computer simulation that stimulates user senses, such as sight and hearing, making users feel as if they are immersed in it. VR has three elements: Interaction, Immersion, and Imagination [1]. It can be used in the teaching of human anatomy, online navigation of museums, 3D game teaching, flight training, and rehabilitation [2]. VR has become a therapeutic tool in many medical and rehabilitation fields. Due to the cost decline and ease of use of this technology, it has become an effective tool and trend in various fields.

However, its greatest obstacles lie in the lack of space, time, support staff, appropriate customer and customer incentives, therapist knowledge, and management support. The clinical use of VR often depends on the motivation and attitude of the therapist [3,4].

In the clinical investigations on the VR experience and perception of physical therapists (PTs) and occupational therapists (OTs) in Canada conducted by Levac et al., it was found that VR treatment is most commonly used for stroke (25.8%), brain injury (15.3%), musculoskeletal disorder (14.9%), cerebral palsy (10.5%), and neurodevelopmental disorders (6.3%) [3]. Most of the clinical applications of VR are for neurological problems. Moreover, numerous literature shows that VR is used to treat patients with stroke, cerebral palsy, Parkinson's disease, etc. [5–10]. Most researches in VR medical applications are used to the upper limb movement rehabilitation for stroke patients. The upper limb virtual reality rehabilitation systems were developed for the stroke group. The patient grasped and released the characteristic objects in the virtual environment, and finger movement control of the stroke patients after 4–6 weeks of VR intervention was improved [6,10]. Some scholars used Kinect and customized games to train the children with cerebral palsy (CP). The evidence appears to support the use of VR as a promising tool to be incorporated into the rehabilitation process of CP [7,11,12].

According to the World Heath Organization (WHO), musculoskeletal conditions affect muscles, bones, joints and associated tissues such as tendons and ligaments. To patients, musculoskeletal conditions are typically characterized by pain and limitations in mobility or functional ability.... Pain and restricted mobility are the consistent features of the range of musculoskeletal conditions. Musculoskeletal conditions are the second largest contributor to disability worldwide [13]. However, at present, there is less evidence on the therapeutic effect of VR on patients with musculoskeletal system disorder [14–18]. In addition, studies have shown that VR is beneficial in pain management, for example, in pain relief during dressing changes of burn patients [19]. VR can also reduce anxiety, distract from the fear of pain, and alleviate stress [20]. It can divert the attention of patients who are afraid of moving because of pain.

So far, there are no integrated and first-rate studies that explore which musculoskeletal disorders are suitable for VR treatment. The comparison of the effects of VR games and other treatments (e.g., traditional treatment, instrumental therapy, exercise) on patients with musculoskeletal disorder is inconclusive. Therefore, this article integrates the results of studies made in recent years into a scoping review to: (1) Compare the effectiveness of VR and other treatment interventions for patients with musculoskeletal disorder; (2) further explore whether there is any consistency in the VR treatment of patients with musculoskeletal system disorder, so as to give recommendations based on the highest level of evidence. This review only contains RCT articles with a PEDro Scale score ≥5 points.

2. Materials and Methods

2.1. Determination and Selection of Articles

The methodology of this scoping review was based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines because the main aim of this work is mapping all the available literature in the musculoskeletal field [21]. The use of the checklists based on PRISMA statement improve the quality and transparency of the scoping reviews [17]. Search was made in the PubMed, IEEE, and the MEDLINE library for reference literature using keywords and synonyms of "virtual reality", "pain", and "musculoskeletal". After performing a journal search, RCT (randomized controlled trial) journals that were written in English within the last 10 years (January 2008 to August 2018) were selected, and non-musculoskeletal diseases such as "stroke", "neurological", and "cognitive" were excluded using the Physiotherapy Evidence Database (PEDro) scale (http://www.pedro.org.au/). When reference materials could not be found on the PEDro website, scores were independently made by two authors who have completed the PEDro Scale training tutorial on the Physiotherapy Evidence Database.

When the scores were different, the clinical physiotherapist with more than five years of experience, and who completed the PEDro assessment training, was asked to conduct another assessment. When issues such as disagreement or ambiguity arose, they were resolved through discussions. Finally, literature with very low PEDro scores (<5/10) was excluded. The search process is shown in Figure 1. Since there are few studies on VR for musculoskeletal disorders, we do not explore the virtual reality (VR) outcomes for any specific pathology in our study, but explore the VR treatment effects, such as pain relief, joint mobility, function, range of motion (ROM), muscle strength, angular velocity and self-satisfaction for all musculoskeletal disorders.

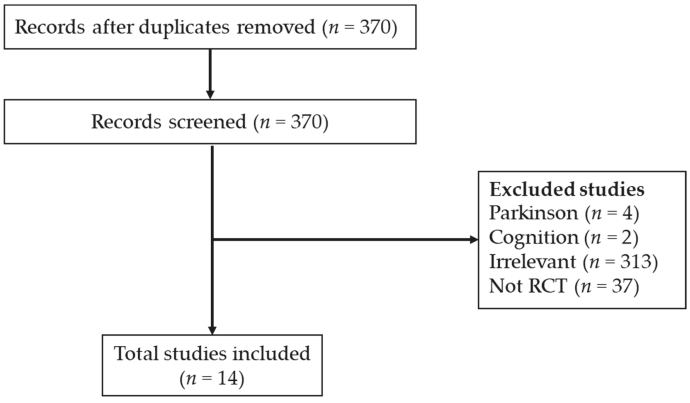


Figure 1. Flow chart displaying the screening process for studies included in this systematic review.

2.2. Data Extraction and Quality Assessment

Initially, the two authors completed the abstract review independently. When it was not possible to know whether an article could be included in the scoping review from its abstract, an assessment of the full article was made. All of the articles that had been included were reviewed in full. After sorting, the following were investigated: (1) Whether VR treatment improved the musculoskeletal system as compared with other treatments; (2) whether there was any consistency in the musculoskeletal disorder of patients that received VR treatment. The selected articles were summarized and analyzed with descriptive statistics. The author, publication year, subject, intervention, outcome measures, and mean between-group differences (95% confidence interval) were extracted from the references by the two authors of this study. A consensus was reached through discussion when the authors had different opinions.

3. Result

A database search was made to exclude articles with a PEDro score of less than 5 and non-English publications. A total of 14 articles were included. These 14 articles were included in this scoping review (Figure 1).

3.1. Quality of the Included Studies

The quality of included studies was presented in Table 1. The mean PEDro score of the included articles was 6.14 (range, 5–7). All studies were randomized (100%). 8 studies carry out concealed allocation (57.14%), and all studies baseline comparability (100%). All studies were analyzed between-group comparison (100%) and 13 studies reported point estimates and variability (92.86%). All studies didn't carry out blind therapist. One study carried out blind subjects (7.14%) and 7 studies carried out blind assessors (50%). 10 studies have adequate outcome measurement (71.43%), and 5 studies have an intension-to-treat analysis (35.71%).

3.2. Description of Included Studies

Each article abstract (including author, musculoskeletal disorder, design, participants, intervention, comparison, and outcome measure) is organized in Table 2. In terms of age, a study about frozen shoulder investigated subjects older than 20 years of age [15]; a research about subacromial impingement syndrome (SAIS) studied subjects between 18–60 years old [22]; subjects of two articles discussing chronic cervical pain were older than 18 years old [17,18]; three studies that explored low back pain (LBP) had subjects between 18–50 years old [23], and those between 40–55 years old [24,25]; an investigation on pelvic floor muscle had subjects older than 50 years of age [26]; two researches discussed the treatment of acute burn wounds in adolescents aged 10–18 years [27,28]; three studies discussed the treatment for patients with TKR aged in the sixties [14,16,29]; an article discussing ankle sprains had subjects aged 18–64, belonging to the working-age group [30]. In terms of experimental intervention, most of the study regarding VR intervention lasted 15 to 30 min, 2 to 4 times per week for 2 to 6 weeks. One research conducted VR intervention for 3 weeks [16]; 3 articles discussed 4 weeks of VR intervention [15,17,24]; 2 studies described 5 weeks of VR intervention [18,26]; and another 2 articles discussed 6 weeks of VR intervention [22,30]. One study conducted VR treatment beginning the second day after TKA until 6 months [29]. There were 5 studies that compared VR intervention and no intervention at all [17,23,27,28,30]. The rest made comparisons between VR and other treatments.

3.3. Virtual Reality Resources Choosing

Virtual Reality was applied using several resources. In the 14 studies, one study used Kinect [15]; 5 studies used Wii [14,22,24,26,30]; 5 studies used VR glasses (one of the studies used headphones and joysticks) [17,18,25,27,28]; two studies used 3-D TV and 3-D shutter glasses [23,29]; and one study used enhanced reality with VR and mirror therapy [16].

3.4. Heterogeneity of Included RCT

The outcome could not be pooled into meta-analysis due to the following reasons. Clinical heterogeneity (Table 2) can be clearly observed from the participant, intervention, exercise mode, and outcome measures of the included studies. Diversity is seen in patient conditions, frequency and duration of VR intervention, whether or not the patient does home exercise, received patient health education, whether the experiment conducted was pure VR (only VR) or VR mixed with traditional physical therapy or with exercise therapy, whether the outcome measure contains follow-up, and whether different estimate measures were inconsistent at different times.

3.5. Effect of Virtual Reality versus Other Interventions

In the articles included, a total of twelve studies compared the effects of VR treatment and other intervention on orthopedic conditions (Table 3). The research on patients suffering from frozen shoulder for more than three months shows that four weeks of VR plus modalities (hot pack and ultrasound) produced a significant 8% increase in their shoulder range of motion (ROM) when compared to traditional exercise training, plus modalities [15]. Another research showed that after 6 months of short-term training and one-month of follow-up, the subacromial impingement syndrome (SAIS) patients without a rotator cuff problem on the VR group and home exercise group (scapular muscles training), were able to significantly reduce their disability and improve their quality of life. Furthermore, the VR group showed significant improvement of SAIS and scapular dyskinesis symptoms when compared with the home exercise group [22]. Another article showed that patients with chronic cervical pain who went through 5 weeks of VR and cervical kinematic training (KT) had a big difference in the global perceived change (variations in different areas of patient self-assessment, such as satisfaction, self-reported pain differences), which could last for three months when compared to those in the KT group [18]. A study also showed that after four weeks of training, the VR group of patients with chronic cervical pain displayed a significant difference in terms of pain, physical condition, fear of moving the

neck, as well as in the mean and peak velocity from those in the laser beam projected group. However, there is no significant difference in cervical ROM during follow-up between the VR treatment group and the laser beam projected group [17]. Patients with chronic low back pain in another study were able to significantly improve pain, pressure algometry, disability, and the fear of low back pain after four weeks of VR training [24]. Another research proposed that VR with the supplementary traditional physical therapy can significantly reduce pain, fear, and increase functions for patients with subacute or chronic non-specific lower back pain [25]. Although a study comparing five weeks of pelvic floor muscle training via VR and traditional gym ball training, showed no significant difference in muscle strength, but a statistically significant difference in endurance was observed [26]. One study supported the idea that VR therapy during burn wound care can reduce adolescent pain [25]. Three included studies examined the effects of VR on patients with TKR [14,29]. The results of these two articles showed that VR treatment (physical therapy plus VR) did not produce a significant difference in terms of pain, ROM, walking speed, balance, and walking test for patients with total knee replacement (TKR), when compared with conventional therapy [14,16]. The other demonstrated that VAS scales were significantly lower in the experimental group than the control group during acute phase (at 3, 5, and 7 days after TKR) ($p < 0.05$) [29]. However, it did not reach the minimal clinically important difference (MCID) [31]. In the previously described study, VR intervention (one month, three months, six months after TKR) in the chronic phase can improve the functional recovery of the patients with TKR [29]. A study on the treatment of ankle sprains suggested that there is no significant difference in all parameters between VR treatment and traditional treatment [30].

3.6. Effect of Virtual Reality Versus No Intervention

In the included articles, four articles discussed the therapeutic effects of VR and no intervention on chronic cervical pain, burn wound, low back pain, and ankle sprains (Table 4). When applied to chronic cervical pain and burn wound, a statistically significant difference was present in some parameters, as described in the following section. Bohat et al. (2017) [17] studied patients with chronic cervical pain after four weeks of training and found that the VR group had significantly different results from the control group in disability, cervical angular velocity, time to peak velocity, and head follow-up task accuracy. However, in cervical ROM, physical health, and fear of moving the neck, no significant difference was observed [17]. Another study compared the results of a 3-day VR training of low-back pain patients with the results of the non-invasive group, and found no statistically significant difference in lumbar spine flexion ROM and pain improvement [23].

During the dressing application of patients with burn wounds, patients undergoing VR treatment received significantly lower doses of Entonox (analgesic) compared with those in the standard distraction group. However, there is no significant reduction in patient pain [28]. For patients with ankle sprains, no statistically significant difference was observed between the VR treatment and the control group [30].

Table 1. Physiotherapy Evidence Database (PEDro) Score for Included Studies ($n = 14$).

	Huang et al. 2014 [15]	**Pekyavas et al. 2017 [22]**	**Bahat et al. 2015 [18]**	**Bahat et al. 2017 [17]**	**Kim et al. 2014 [24]**	**Thomas et al. 2016 [23]**	**Yilmaz Yelvar et al. 2017 [25]**	**Martinho et al. 2016 [26]**	**Kipping et al. 2012 [28]**	**Jeffs et al. 2014 [27]**	**Fung et al. 2012 [14]**	**Koo et al. 2018 [16]**	**Jin et al. 2018 [29]**	**Punt et al. 2016 [30]**
	2014 Taiwan	2017 Turkey	2015 Australia	2017 Australia	2014 Korea	2016 America	2017 Turkey	2016 Brazil	2012 Australia	2014 America	2012 Canada	2018 Korea	2018 China	2016 Switzerland
Eligibility criteria	Y	N	Y	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Random allocation	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Concealed allocation	N	N	Y	Y	N	Y	N	Y	Y	Y	N	Y	N	Y
Baseline comparability	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Blind subjects	N	N	N	N	N	N	N	N	N	Y	N	N	N	N
Blind therapists	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Blind assessors	Y	N	Y	Y	N	N	Y	N	N	N	Y	Y	N	Y
Adequate follow-up	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y	N	Y	N
Intention-to-treat analysis	Y	N	N	Y	N	N	N	Y	Y	N	N	N	N	Y
Between-group comparisons	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Point estimates and variability	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y
Total score (0–10)	7/10	5/10	7/10	7/10	5/10	6/10	6/10	6/10	7/10	7/10	5/10	6/10	5/10	7/10

Abbreviations: PEDro, Physiotherapy Evidence Database; Y: yes; N: no.

Table 2. Description of Included studies.

Title	Author	Part	Design	Participant (number)		Intervention	Exercise Mode (Frequency or Intensity)	Outcome Measures
				Age (years) = mean (SD)				
Intelligent Frozen Shoulder Rehabilitation	Huang et al. 2014 [15]	Frozen shoulder	RCT	frozen shoulder syndrome > 3 months			20 min/time, 2 times/week (Total 4 weeks)	ROM, CMS assessment
				E	n = 20 Age (years) = 60.65 (11.84)	Hot pack + ultrasonic + VR		
				C	n = 20 Age (years) = 61.45 (12.84)	Hot pack + ultrasonic + traditional exercise training		
Comparison of virtual reality exergaming and home exercise programs in patients with subacromial impingement syndrome and scapular dyskinesis: Short term effect	Pekyavas et al. 2017 [22]	subacromial impingement syndrome(SAIS) & scapular dyskinesis	RCT	18–60 years old; Type II SAIS				VAS(rest, activity, night), SPADI, Neer, Hawkins, SRT, SAT, LSST1-3
				None rotator cuff problem				
				E	n = 15 Age (years) = 40.33 (13.20)	VR + control period (after 6 weeks)	VR: 45 min/day, twice a week, for 6 weeks; Control: 1 month for home exercise	
				C	n = 15 Age (years) = 40.60 (11.77)	Exercise + control period (after 6 weeks)	Exercise: 45 min/day, twice a week, for 6 weeks; Control: 1 month for home exercise	
Cervical Kinematic Training with and without Interactive VR Training for Chronic Neck Pain—a Randomized Clinical Trial	Bahat et al. 2015 [18]	Chronic neck pain	RCT	Neck pain > 3 months, NDI > 10%				VAS, Neck Disability Index, TSK, ROM, Peak velocity, mean velocity, TIP%, Sway SD, Accuracy, Eyes closed balance, singer leg stance, step test
				E	n = 16 Age (years) = 40.63 (14.18)	VR + kinematic training	Total 30 min, at least 3 times a week, for 5 weeks	
				C	n = 16 Age (years) = 41.13 (12.59)	kinematic training (using laser point)		
Remote kinematic training for patients with chronic neck pain: a randomized controlled trial	Bahat et al. 2017 [17]	Chronic neck pain	RCT	Neck pain > 3 months, NDI > 12%				Neck Disability Index, Peak velocity, mean velocity, VAS, EQ5D, TSK, NVP, TTP%, Accuracy, ROM, GPE
				VR	n = 30 Age (years) = 48 (9.5)	VR	1 set 5 min, 20 min/day, 4 times/week, for 4 weeks	
				Laser	n = 30 Age (years) = 48 (12.5)	Laser point training		
				C	n = 30 Age (years) = 48 (13)	Not receive any treatment		
The Effects of VR-Based Wii Fit Yoga on Physical Function in Middle-Aged Female LBP Patients	Kim et al. 2014 [24]	LBP	RCT	LBP > 2 months				VAS, pressure algometer, ODI, RMDQ, FBQ
				E	n = 15 Age (years) = 44.33	VR	30 min/session, 3 session/week, for 4 weeks (1 session had 7 exercise program. 3 min of exercise and 1 min of rest)	
				C	n = 15 Age (years) = 50.46	Trunk stabilizing exercise + physical therapy	2 sets (30 min), 1set included 10 repetitions, physical therapy 30 min	
Feasibility and Safety of a Virtual Reality Dodgeball Intervention for Chronic Low Back Pain: A Randomized Clinical Trial	Thomas et al. 2016 [23]	LBP	RCT	18–50 years old with LBP > 3 months				pain and harm, lumbar spine flexion ROM
				kinesiophobia ≥35				
				E	n = 26 Age (years) = 23.9 (6.8)	VR	3 days (<48 h)	
				C	n = 26 Age (years) = 26.7 (8.5)	Not receive any treatment		

Table 2. *Cont.*

Title	Author	Part	Design	Participant (number)		Intervention	Exercise Mode (Frequency or Intensity)	Outcome Measures
				Age (years) = mean (SD)				
Is physiotherapy integrated virtual walking effective on pain, function, and kinesiophobia in patients with non-specific low-back pain? Randomised controlled trial	Yilmaz Yelvar et al. 2017 [25]	LBP	RCT	non-specific LBP for longer than 2 months				VAS, ODI, TKS, TUG, and 6MWT scores
				E	n = 23 Age (years) = 46.27 (10.93)	VR + Traditional physical therapy	5 times/week, for 2 weeks	
				C	n = 23 Age (years) = 52.81 (11.53)	Traditional physical therapy	5 times/week, for 2 weeks	
The effects of training by virtual reality or gym ball on pelvic floor muscle strength in postmenopausal women: a randomized controlled trial	Martinho et al. 2016 [26]	Pelvic floor muscle	RCT	>50 years old women				Maximum strength, average strength, endurance
				>1 year postmenopausal phase				
				APT-VR	n = 30 Age (years) = 61.9 (8.6)	Abdominopelvic training by VR	1 session 5 min with 90 s resting, for 10 session. Twice a week, for 5 weeks	
				PFMT-GB	n = 30 Age (years) = 61 (8.5)	Pelvic floor muscle training using a gym ball	4 series of 10 fast & sustained (8 s maintain with 16 s resting), each exercise 5 times, twice a week, for 5 weeks	
Virtual reality for acute pain reduction in adolescents undergoing burn wound care: A prospective randomized controlled trial	Kipping et al. 2012 [28]	Burn wound	RCT	11–18 years old				VAS, FLACC scale
				burn wound Total Body Surface Area (TBSA) > 1%				
				E	n = 20 Age (years) = 12.6 (1.3)	VR	Dressing period (3–58 min), only 1 time	
				C	n = 21 Age (years) = 13.5 (1.8)	Another distraction way or no distraction		
Effect of Virtual Reality on Adolescent Pain During Burn Wound Care	Jeffs et al. 2014 [27]	Burn wound	RCT	10–17 years old				Adolescent Pediatric Pain Tool, Spielberger State-Trait Anxiety InventoryFor Children, Pre-Procedure Questionnaire, Post-Procedure Questionnaire
				standard care	n = 10 Age (years) = 18.9 (2.8)	standard care	Dressing period only 1 time	
				passive distraction	n = 10 Age (years) = 12.6 (2.1)	passive distraction watching a movie		
				virtual reality	n = 8 Age (years) = 14.8 (2.0)	virtual reality		
Use of Nintendo Wii Fit™ in the Rehabilitation of Outpatients Following Total Knee Replacement: a Preliminary Randomized Controlled Trial	Fung et al. 2012 [14]	TKR	RCT	requiring twice-weekly physiotherapy treatment for TKR rehabilitation				active knee flexion/extension ROM, 2 min walk test, NPRS, ABCS, LEFS
				Full lower extremity weight bearing				
				E	n = 27 Age (years) = 67.9 (9.5)	physiotherapy + VR	physiotherapy (45 min), 15 min VR until discharge	
				C	n = 23 Age (years) = 68.2 (12.8)	physiotherapy + lower extremity exercise	physiotherapy (45 min), 15 min lower extremity exercise until discharge	

Table 2. *Cont.*

Title	Author	Part	Design	Participant (number)		Intervention	Exercise Mode (Frequency or Intensity)	Outcome Measures
				Age (years) = mean (SD)				
Enhanced Reality Showing Long-Lasting Analgesia after Total Knee Arthroplasty: Prospective, Randomized Clinical Trial	Koo et al. 2018 [16]	TKR	RCT	Full term	n = 20 Age (years) = 63.7 (5.09)	VR + physiotherapy for 2 weeks	VR + PT: 5 days/week, for 2 weeks	VAS, WOMAC, 6 min walk test, Timed-stands test
				Half term	n = 22 Age (years) = 65.0 (6.97)	VR + physiotherapy for 1 week before physiotherapy for 1 week	VR + PT for 1 week	
Virtual reality intervention in postoperative rehabilitation after total knee arthroplasty: a prospective and randomized controlled clinical trial	Jin et al. 2018 [29]	TKR	RCT	E	n = 33 Age (years) = 66.45 (3.49)	VR(begin 2nd days for TKA) + conventional rehabilitation	three sets of 30 repetitions	WOMAC, HSS, VAS, ROM.
				C	n = 33 Age (years) = 66.30 (4.41)	conventional rehabilitation	three sets of 30 repetitions	
Wii Fit™ Exercise Therapy for the Rehabilitation of Ankle Sprains: Its Effect Compared with Physical Therapy or No Functional Exercises at All	Punt et al. 2016 [30]	Ankle sprain	RCT	18–64 years old				FAAM-ADL, FAAM-sport, VAS-rest, VAS-walk
				Grade I or II lateral ankle sprain				
				requiring 4 weeks RICE and can pain free movement				
				VR	n = 30 Age (years) = 34.7 (10.7)	VR	30 min/time, 2 times/week, for 6 weeks	
				Physiotherapy	n = 30 Age (years) = 34.7 (11.3)	modalities, joint mobilization, muscle strengthening, proprioceptive exercise	30 min/time, 9 times/6 weeks	
				C	n = 30 Age (years) = 33.5 (9.5)	Not receive any treatment	-	

Abbreviations: n (number); E (experimental group); C (control group); VR (Virtual reality); min (minute); ROM (range of motion); CMS (Constant-Murley score); VAS (Visual Analog Scale); SPADI (Shoulder Pain and Disability Index); SRT (Scapular Retraction Test); SAT (Scapular Assistance Test); LSST (Lateral Scapular Slide Test); TSK (Tampa scale of kinesiophobia); TIP% (Time to peak velocity percentage); GPE (Global perceived effect); sway SD (standard deviation of the static head sway); EQ-5D(EQ-5D™, http://www.euroqol.org); NVP (Number of velocity peaks); ODI (Oswestry low-back pain disability index); RMDQ (Roland Morris disability questionnaire); FBQ (fear avoidance beliefs questionnaire); NPRS (Numeric Pain Rating Scale); LEFS (Lower Extremity Functional Scale); ABCS (Activity-specific Balance Confidence Scale); WOMAC (Western Ontario and McMaster Universities Osteoarthritis Index); FAAM (Foot and Ankle Ability Measure); ADL (activities of daily living); RICE (rest, ice, compression and elevation); FLACC (Faces, legs, activity, cry, consolability scale); Hospital for Special Surgery knee score (HSS); TKS (TAMPA Kinesiophobia Scale), TUG (timed-up and go test); 6MWT (6-Minute Walk Test); RCT (randomized controlled trial); LBP (low back pain); TKR (Total Knee Replacement).

Table 3. Effect of virtual reality (VR) versus another intervention.

Study	Outcome Measure		Mean Difference between VR Groups and Another Intervention	Significance of Difference between Groups
Huang et al. 2014 [15]	ROM		8%	Between groups $p < 0.05$
	CMS		NA	Between groups $p < 0.05$
Pekyavas et al. 2017 [22]	Neer	post-intervention/1 month follow-up	NA	$p = 0.02$
	SRT	post-intervention/1 month follow-up	NA	$p = 0.01$
	SAT	post-intervention/1 month follow-up	NA	$p = 0.047$
	VAS (rest, activity, night)	post-intervention/1 month follow-up	NA	-
	SPADI	post-intervention/1 month follow-up	NA	-
	Hawkins	post-intervention/1 month follow-up	NA	-
	LSST1-3	post-intervention/1 month follow-up	NA	-
Bahat et al. 2015 [18]	cervical flexion ROM	post-intervention	NA	Between groups $p < 0.05$
		3 months follow-up	NA	-
	Global Perceived change	post-intervention	NA	-
		3 months follow-up	NA	Between groups $p < 0.05$
	VAS	post-intervention/3 months follow-up	NA	-
	NDI	post-intervention/3 months follow-up	NA	-
	TSK	post-intervention/3 months follow-up	NA	-
	Velocity	post-intervention/3 months follow-up	NA	-
	TIP%	post-intervention/3 months follow-up	NA	-
	Accuracy	post-intervention/3 months follow-up	NA	-
	sway SD	post-intervention/3 months follow-up	NA	-
	Eyes closed balance	post-intervention/3 months follow-up	NA	-
	singer leg stance	post-intervention/3 months follow-up	NA	-
Bahat et al. 2017 [17]	V_{mean} (F,LR)	Post-pre intervention	NA	Between groups $p < 0.05$
	V_{peak} (LR)	Post-pre intervention	NA	Between groups $p < 0.05$
	V_{mean} (F,E,LR)	3 months follow up-pre intervention	NA	Between groups $p < 0.05$
	V_{peak} (E,LR)	3 months follow up-pre intervention	NA	Between groups $p < 0.05$
	VAS	Post-pre/3 months-pre	NA	Between groups $p < 0.05$
	EQ5D	Post-pre/3 months-pre	NA	Between groups $p < 0.05$
	Accuracy (F,RR,LR)	Post-pre intervention	NA	Between groups $p < 0.05$
	Accuracy (F)	3 months follow up-pre intervention	NA	Between groups $p < 0.05$
	TTP%	Post-pre/3 months-pre	NA	Between groups $p < 0.05$
	ROM	Post-pre/3 months-pre	NA	-
	NDI	Post-pre/3 months-pre	NA	-
	TSK	Post-pre/3 months-pre	NA	-
	NVP	Post-pre/3 months-pre	NA	-
	GPE	Post-pre/3 months-pre	NA	-

Table 3. *Cont.*

Study	Outcome Measure	Mean Difference between VR Groups and Another Intervention	Significance of Difference between Groups
Kim et al. 2014 [24]	VAS	NA	Between groups $p < 0.05$
	Pressure algometer	NA	
	ODI	NA	
	FBQ	NA	
	RMDQ	NA	-
Yilmaz Yelvar et al. 2017 [25]	VAS	NA	Between groups $p < 0.05$
	TKS	NA	
	TUG	NA	
	6 MWT scores	NA	
Martinho et al. 2016 [26]	Maximum strength	−0.08	$p = 0.1$
	average strength	0.01	$p = 0.6$
	Endurance	1.83	$p = 0.007$
Jeffs et al. 2014 [27]	Pain	23.7	$p = 0.029$
Fung et al. 2012 [14]	Active knee flexion ROM	−0.33	-
	Active knee extension ROM	−0.6	-
	2 min walk test	2.68	-
	NPPS	16.84	-
	ABCS	14.11	-
	LEFS	31.85	-
Koo et al. 2018 [16]	VAS	NA	-
	WOMAC	NA	-
	6 min walk test	NA	-
	Timed-stands test	NA	-
Jin et al. 2018 [29]	VAS (at 3, 5, 7 days after TKR)	NA	$p < 0.05$
	WOMAC (at 1, 3, 6 months after TKR)	NA	$p < 0.05$
	HSS (at 1, 3, 6 months after TKR)	NA	$p < 0.05$
Punt et al. 2016 [30]	FAAM-ADL	NA	-
	FAAM-sport	NA	-
	VAS-rest	NA	-
	VAS-walk	NA	-

Abbreviations: VR (Virtual reality); ROM (range of motion); CMS (Constant-Murley score); VAS (Visual Analog Scale); SPADI (Shoulder Pain and Disability Index); SRT (Scapular Retraction Test); SAT (Scapular Assistance Test); LSST (Lateral Scapular Slide Test); NDI (Neck Disability Index); TSK (Tampa scale of kinesiophobia); TIP% (Time to peak velocity percentage); GPE (Global perceived effect); sway SD (standard deviation of the static head sway); EQ-5D(EQ-5D™, http://www.euroqol.org); NVP (Number of velocity peaks); ODI (Oswestry low-back pain disability index); RMDQ (Roland Morris disability questionnaire); FBQ (fear avoidance beliefs questionnaire); NPRS (Numeric Pain Rating Scale); LEFS (Lower Extremity Functional Scale); ABCS (Activity-specific Balance Confidence Scale); WOMAC (Western Ontario and McMaster Universities Osteoarthritis Index); FAAM (Foot and Ankle Ability Measure); ADL (activities of daily living); v_{mean} (Mean velocity); Vpeak (Peak velocity); F (Flexion), E (Extension), LR (Left rotation), RR (Right rotation); the-marked mean $p > 0.05$; NA (not available); Hospital for Special Surgery knee score (HSS); TKS (TAMPA Kinesiophobia Scale), TUG (timed-up and go test); 6 MWT (6-Minute Walk Test).

Table 4. Effect of VR versus no intervention.

Study	Outcome Measure	Mean Difference between VR Groups and Control Group	Significance of Difference between Groups
Bahat et al. 2017 [17]	NDI	NA	Between groups $p < 0.05$
	velocity	NA	
	TTP% (F,LR)	NA	
	Accuracy (F,RR)	NA	
	ROM	NA	-
	EQ5D	NA	-
	TSK	NA	-
	NVP	NA	-
Thomas et al. 2016 [23]	ROM	NA	-
	Pain	NA	-
Kipping et al. 2012 [28]	VAS	NA	-
	FLACC (dressing removal)	NA	Between groups $p < 0.05$
Punt et al. 2016 [30]	FAAM-ADL	NA	-
	FAAM-sport	NA	-
	VAS-rest	NA	-
	VAS-walk	NA	-

Abbreviations: VR (Virtual reality); NDI (Neck Disability Index); TIP% (Time to peak velocity percentage); ROM (range of motion); EQ-5D (EQ-5D™, http://www.euroqol.org); TSK (Tampa scale of kinesiophobia); NVP (Number of velocity peaks); VAS (Visual Analog Scale); FLACC (Faces, legs, activity, cry, consolability scale); FAAM (Foot and Ankle Ability Measure); ADL (activities of daily living); the-marked mean $p > 0.05$; NA (not available).

3.7. Effect of Virtual Reality on Acute and Chronic Musculoskeletal Pain

Although associated pain is not itself part of the root disorder, managing the pain of musculoskeletal disorders is a major part of general practice. Of the 14 musculoskeletal studies included, six were for acute pain, including the dressing of the burn wound [27,28], three were for TKR patients [14,16,29], one for patients with ankle sprain [30], and the rest of the eight articles were for chronic musculoskeletal pain patients, including patients with frozen shoulder, SAIS, Neck pain, LBP, and pelvic floor muscle training [15,17,18,22,24–26]. VR treatment seems to reduce the pain of burn patients, or it could reduce the use of analgesics [27,28]. No significant difference in all parameters was observed when TKR and ankle sprain patients received VR treatment as compared to conventional treatment [14,16,29,30]. And there is no MCID for VAS pain in its acute phase [29]. In the included articles, a significant difference in the main outcome was observed for all patients with chronic pain aside from the research conducted by Tomas et al. [23].

4. Discussion

Most virtual reality treatment research applications still focus on the VR treatment of central nervous system problems, such as stroke and cerebral palsy, while only a little research explores the therapeutic effect of VR treatment on patients with musculoskeletal disorders. At present, there is no research on the integration of virtual reality for patients with various musculoskeletal disorders and an analysis of its effects. Therefore, this scoping review searched and integrated multiple musculoskeletal disorders in VR applications. Analysis showed which patients with musculoskeletal disorders had better results after VR treatment.

In general, chronic pain usually lasts for more than 12 weeks, while acute pain usually lasts for 4 to 6 weeks [32]. Therefore, patients in the articles included in this study are those who experience chronic pain due to frozen shoulder (symptoms lasting more than 3 months), SAIS (symptoms lasting at least 2 months), neck pain (symptoms appear for more than 3 months), and LBP (symptoms persist for 2 or 3 months) [15,17,18,22–25]. The study on burn wound care included patients with acute pain due to burns. Fung et al. (2012) included TKR patients in their study under the condition of being able to apply a full load on the lower limbs after 2 weeks of physical therapy post-surgery. Another research

made by Koo et al. studied TKR patients after 2 weeks of physical therapy post-surgery followed by VR treatment. In the preceding two TKR studies, patients belonged to the sub-acute and acute phase, and the pain that they felt was an acute pain [14,16]. As for another study, VR intervention was applied from one days to 6 months after TKR (longitudinal study). In the early postoperative period (3–7 days), VR intervention did not achieve any clinically better analgesic effect than traditional treatment [29]. The study on ankle sprains mentioned that patients with non-repetitive sprains can undergo emergency treatment for 4 weeks without pain followed by VR treatment; therefore, this does not belong to the category of chronic pain. From this systematic review, it was found that subjects that experienced more effective VR intervention tend to be patients with chronic orthopedic pain, or those with acute pain due to burn wounds. Patients with TKR and ankle sprains are not chronic pain patients, and results show that VR treatment is not more effective than other treatments.

Generally, patients suffering from chronic pain have lower levels of fitness than healthy people. This is because pain can affect the motor control strategies of people. Individuals tend to move in the least painful way; however, the least painful way is usually to refrain from moving. This causes a decrease in muscle size and strength; it repeatedly increases pain and stress, eventually producing to a vicious circle [33]. In the included research articles, the motions designed for patients with chronic orthopedic disorders are suitable for the joint movements of patients of this type. For example, in the virtual reality games for patients with frozen shoulder, the actions designed include shoulder elevation, shoulder IR/ER, and a shoulder abduction action, and suitable WII games are selected for shoulder impingement patients (such as the tennis game which involves shoulder capsule stretch, pectoral muscle stretch and shoulder elevation). For patients with other chronic orthopedic disorders, through somatosensory interactive games with larger movements, patients could try actions which they could not achieve. Furthermore, people usually focus on pain or impending pain; therefore, the use of VR is effective in distracting the attention of patients from pain. The distraction produced by VR reduces pain, induces movement, and promotes exercise. It also motivates patients to move. Most users describe that their experience of VR was pleasant, and it can relieve pain as well as reduce anxiety [20,34–36]. Nevertheless, VR treatment done under inadequate supervision may result in less than expected results [17,30]. VR intervention under supervision can increase the motivation or induce patients to receive movement training and boost their concentration. This systematic review found that VR treatment for patients with chronic pain, such as 4 weeks of VR plus modalities (hot pack and ultrasound) on patients with frozen shoulders produced a significant 8% increase in shoulder ROM when compared to traditional exercise training plus modalities [15]. Subacromial impingement syndrome (SAIS) patients underwent 6 weeks of VR training for 45 min/day, twice a week, and showed a significant improvement of SAIS and scapular dyskinesis symptoms than those in the home exercise group.

In addition, this result lasted for one month [22]. Another article that studied chronic cervical pain patients after 5 weeks of VR plus cervical kinematic training (KT) recounted a significant difference in global perceived change (patient self-reported changes in different areas, such as satisfaction, self-reported pain differences) when compared to the only KT group. The experienced outcome lasted for 3 months [18]. One research on low back pain showed that 4 weeks of VR training can alleviate pain, deep tissue pressure algometry, disability, and fear of low back pain. For the TKR patients, after one month to six months of VR intervention, the knee function is better than for those who received the traditional treatment [29]. One of the articles included in this study showed no significant difference in pain and lumbar spine flexion ROM after comparing 3 days of VR treatment for patients with lower back pain, and patients without VR treatment [23]. This scoping review shows that chronic patients may receive at least four weeks of VR treatment in order to experience a significant therapeutic effect. In addition, VR training seems to have a short-term effect for patients with chronic pain in the musculoskeletal system [17,18,22,25]. This is consistent with past research [36–38].

The results of this study show that VR treatment with a hand joystick significantly reduces the pain score of patients when removing dressings from patients with acute burns, or it will reduce the

use of analgesics [27,28]. This is consistent with previous research [39–42]. Hoffman, et al. [40] showed that the use of VR for patients under severe pain can effectively reduce pain by 41%. It is speculated that VR can also be used to distract patients from severe acute pain during dressing change. Therefore, VR can be used to divert attention, thus reducing the use of analgesics.

In summary, VR treatment can reduce pain in acute burn wound care and chronic musculoskeletal disorders. It can effectively distract patients with chronic pain, and allow them to ignore the cumbersome rehabilitation training, consequently improving treatment motivation. In addition, VR treatment may be helpful in the psychological level and the establishment of confidence. For example, patients with burns or chronic disability may have a tendency to fall into depression because of the long course of the disorder. The use of VR can release psychological stress and reduce their fear of pain [19].

Virtual reality is also helpful in the control and perception of muscle movements. This systematic review includes the hard to control PFM, as well as waist and neck movements. The PFM training research included in this article [26] recommended the simple contraction of the lower abdominal muscles in the VR group. Previous studies pointed out that the lower abdominal muscles have a synergistic effect with PFM. Therefore, some scholars have suggested that if the patient does not know how to apply force during PFM contraction, training on abdominal transverse muscle contraction can be done to attain the same purpose [43]. Patients in the VR group interacted with the game screen and performed pelvic movements such as pelvic forward, backward, lateral tilt, and go around motions according to the easy-to-understand motion instructions provided by the Wii game screen. This made it possible for patients to understand how to control their pelvic motion, and at the same time, increased the control and perception of the PFM. The LBP patients included in this article used games that combined Wii and yoga for their training [24]. LBP patients usually have weak deep core muscles [44]. Yoga promotes the strengthening and relaxation of the waist muscles and ligaments. Through yoga, the body can be continuously aligned correctly. At the same time, the patients can clearly see their posture on the screen. The Wii board senses the weight and center of gravity of the body and trains LBP patients according to the steps in the game screen. For rehabilitation that needs repeated feedback and learning of exercises, VR can provide enthusiasm. Patients with neck pain can also use VR glasses to perform target tracking according to the instructions given by the game, and flex, stretch, and rotate the neck. Patients can adjust neck motion through the instant feedback given by the VR glasses [17,18]. The preceding discussions show that VR can be used to increase the control and training of PFM as well as the consciousness of waist and neck motion. Moreover, posture can be adjusted through VR instant feedback.

Depending on the different facilities which possess different visual perception methods, Virtual Reality can be divided into four types: (a) Desktop VR: Mouse, trackball, and joystick are the main computer transmission devices and a common PC screen was used as its output; (b) Simulator VR: In a specific environment, machines and equipment, added to an image screen, provided the Users simulation results; (c) Projection VR: With a large projection screen, several projectors, and stereo sound output devices, simulation scenes were projected around the user; (d) Immersion VR: Specific Input and output devices, such as helmet display, etc., were used in this type of simulation [1]. The result of the five included articles in the current study seemed to show some effectiveness of the immersion VR. VR glasses or VR TV output, 3D shuttle glasses or helmet display were used to allow the user to become fully immersed in the system, and computers were used to provide image or sound feedbacks (five out of five); Three of the five articles showed Wii (belong to the VR type (a) described as above) had achieved some effectiveness. Some patients may have nausea and dizziness due to the problems of the VR device, such as mismatched motion, motion parallax, viewing angle, limited reproduction of a real environment, and the imperfect simulation of human–world interactions. This condition occurring may affect its treatment effectiveness [45]. Facing the current economic development and the increase of the need of clinical care, we believe that it is necessary to explore the clinical effectiveness and applicability of the VR system. This highlights the importance of the ongoing discussions of the MCID on pain relief or on function increase in this article. The challenges in using the truly immersive

VR system include nausea or dizziness caused by immersing in the virtual world and investment costs (facilities, cost, personnel training) [35,45]. All of these also affect whether VR treatment is appropriate in clinical environment implementation.

Limitations

Because this system review includes first-rate RCT studies, fewer articles that compare the effects of VR therapy with other interventions on patients with musculoskeletal disorders are available. In some articles, the lack of raw numeric data makes it impossible to calculate the mean difference between the experimental and control groups. During the article retrieval process, language was also restricted; therefore, some language bias might exist. In addition, very few articles contain the minimal clinically important difference (MCID) on various parameters; hence, further discussion was not made.

5. Conclusions

VR treatment appears to have a significant effect upon pain relief, increased joint mobility, or the motor functions of patients with chronic painful musculoskeletal disorders. VR seems quite effective in relieving the pain of patients with acute burns as well. However, there is insufficient evidence in the current literature; hence, more research is needed to explore the therapeutic effects of VR treatment on musculoskeletal disorders. In the future, VR games maybe used for more patients with chronic musculoskeletal injuries. As to whether different types of VR would affect the effectiveness for rehabilitation results in musculoskeletal disorder patients, this should also be further investigated.

Acknowledgments: This study was partially funded by E-DA Hospital (EDAHT107029) and the Allied Advanced Intelligent Biomedical Research Center (A2IBRC) under the Higher Education Sprout Project of Ministry of Education.

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

References

1. Burdea, G.C.; Coiffet, P. *Virtual Reality Technology*, 2nd ed.; Wiley-IEEE Press: Hoboken, NJ, USA, 2003.
2. Wang, P.; Wu, P.; Wang, J.; Chi, H.-L.; Wang, X. A Critical Review of the Use of Virtual Reality in Construction Engineering Education and Training. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1204. [CrossRef] [PubMed]
3. Levac, D.; Glegg, S.; Colquhoun, H.; Miller, P.; Noubary, F. Virtual Reality and Active Videogame-Based Practice, Learning Needs, and Preferences: A Cross-Canada Survey of Physical Therapists and Occupational Therapists. *Games Health J.* **2017**, *6*, 217–228. [CrossRef] [PubMed]
4. Levac, D.E.; Miller, P.E. Integrating virtual reality video games into practice: Clinicians' experiences. *Physiother. Theory Pract.* **2013**, *29*, 504–512. [CrossRef] [PubMed]
5. Barton, G.J.; Hawken, M.B.; Foster, R.J.; Holmes, G.; Butler, P.B. The effects of virtual reality game training on trunk to pelvis coupling in a child with cerebral palsy. *J. Neuroeng. Rehabil.* **2013**, *10*, 15. [CrossRef] [PubMed]
6. Broeren, J.; Rydmark, M.; Sunnerhagen, K.S. Virtual reality and haptics as a training device for movement rehabilitation after stroke: A single-case study. *Arch. Phys. Med. Rehabil.* **2004**, *85*, 1247–4250. [CrossRef] [PubMed]
7. Chang, Y.J.; Chen, S.F.; Huang, J.D. A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities. *Res. Dev. Disabil.* **2011**, *32*, 2566–2570. [CrossRef] [PubMed]
8. Deutsch, J.E. Use of a low-cost, commercially available gaming console (Wii) for rehabilitation of an adolescent with cerebral palsy. *Phys. Ther.* **2008**, *88*, 1196. [CrossRef] [PubMed]
9. Dos Santos Mendes, F.A.; Pompeu, J.E.; Modenesi Lobo, A.; Guedes da Silva, K.; Oliveira Tde, P.; Peterson Zomignani, A.; Pimentel Piemonte, M.E. Motor learning, retention and transfer after virtual-reality-based training in Parkinson's disease—Effect of motor and cognitive demands of games: A longitudinal, controlled clinical study. *Physiotherapy* **2012**, *98*, 217–223. [CrossRef]
10. Fischer, H.C.; Stubblefield, K.; Kline, T.; Luo, X.; Kenyon, R.V.; Kamper, D.G. Hand rehabilitation following stroke: A pilot study of assisted finger extension training in a virtual environment. *Top. Stroke Rehabil.* **2007**, *14*, 1–12. [CrossRef]

11. Sevick, M.; Eklund, E.; Mensch, A.; Foreman, M.; Standeven, J.; Engsberg, J. Using Free Internet Videogames in Upper Extremity Motor Training for Children with Cerebral Palsy. *Behav. Sci.* **2016**, *6*, 10. [CrossRef]
12. Bonnechere, B.; Jansen, B.; Omelina, L.; Van Sint Jan, S. The use of commercial video games in rehabilitation: A systematic review. *Int. J. Rehabil. Res.* **2016**, *39*, 277–290. [CrossRef]
13. The Definition of Musculoskeletal Disorders. 15 February 2018. Available online: https://www.who.int/news-room/fact-sheets/detail/musculoskeletal-conditions (accessed on 03 April 2019).
14. Fung, V.; Ho, A.; Shaffer, J.; Chung, E.; Gomez, M. Use of Nintendo Wii Fit in the rehabilitation of outpatients following total knee replacement: A preliminary randomised controlled trial. *Physiotherapy* **2012**, *98*, 183–188. [CrossRef] [PubMed]
15. Huang, M.C.; Lee, S.H.; Yeh, S.C.; Chan, R.C.; Rizzo, A.; Xu, W.; Wu, H.L.; Lin, S.H. Intelligent Frozen Shoulder Rehabilitation. *IEEE Intell. Syst.* **2014**, *29*, 22–28. [CrossRef]
16. Koo, K.-I.; Park, D.K.; Youm, Y.S.; Cho, S.D.; Hwang, C.H. Enhanced Reality Showing Long-Lasting Analgesia after Total Knee Arthroplasty: Prospective, Randomized Clinical Trial. *Sci. Rep.* **2018**, *8*, 2343. [CrossRef] [PubMed]
17. Sarig Bahat, H.; Croft, K.; Carter, C.; Hoddinott, A.; Sprecher, E.; Treleaven, J. Remote kinematic training for patients with chronic neck pain: A randomised controlled trial. *Eur. Spine J.* **2017**, *27*, 1309–1323. [CrossRef] [PubMed]
18. Sarig Bahat, H.; Takasaki, H.; Chen, X.; Bet-Or, Y.; Treleaven, J. Cervical kinematic training with and without interactive VR training for chronic neck pain—A randomized clinical trial. *Man. Ther.* **2015**, *20*, 68–78. [CrossRef] [PubMed]
19. Scapin, S.; Echevarria-Guanilo, M.E.; Boeira Fuculo Junior, P.R.; Goncalves, N.; Rocha, P.K.; Coimbra, R. Virtual Reality in the treatment of burn patients: A systematic review. *Burns* **2018**, *44*, 1403–1416. [CrossRef]
20. Hoffman, H.G.; Patterson, D.R.; Carrougher, G.J. Use of virtual reality for adjunctive treatment of adult burn pain during physical therapy: A controlled study. *Clin. J. Pain* **2000**, *16*, 244–250. [CrossRef]
21. Arlati, S.; Colombo, V.; Ferrigno, G.; Sacchetti, R.; Sacco, M. Virtual reality-based wheelchair simulators: A scoping review. *Assist. Technol.* **2019**, 1–12. [CrossRef]
22. Pekyavas, N.O.; Ergun, N. Comparison of virtual reality exergaming and home exercise programs in patients with subacromial impingement syndrome and scapular dyskinesis: Short term effect. *Acta Orthop. Traumatol. Turc.* **2017**, *51*, 238–242. [CrossRef]
23. Thomas, J.S.; France, C.R.; Applegate, M.E.; Leitkam, S.T.; Walkowski, S. Feasibility and Safety of a Virtual Reality Dodgeball Intervention for Chronic Low Back Pain: A Randomized Clinical Trial. *J. Pain* **2016**, *17*, 1302–1317. [CrossRef] [PubMed]
24. Kim, S.S.; Min, W.K.; Kim, J.H.; Lee, B.H. The Effects of VR-based Wii Fit Yoga on Physical Function in Middle-aged Female LBP Patients. *J. Phys. Ther. Sci.* **2014**, *26*, 549–552. [CrossRef] [PubMed]
25. Yilmaz Yelvar, G.; ırak1, Y.C.; Dalkılınc, M.; Demir, Y.P.; Guner, Z.; Boydak, A.E. Is physiotherapy integrated virtual walking effective on pain, function, and kinesiophobia in patients with non-specific low-back pain? Randomised controlled trial. *Eur. Spine J.* **2017**, *26*, 538–545. [CrossRef] [PubMed]
26. Martinho, N.M.; Silva, V.R.; Marques, J.; Carvalho, L.C.; Iunes, D.H.; Botelho, S. The effects of training by virtual reality or gym ball on pelvic floor muscle strength in postmenopausal women: A randomized controlled trial. *Braz. J. Phys. Ther.* **2016**, *20*, 248–257. [CrossRef] [PubMed]
27. Jeffs, D.; Dorman, D.; Brown, S.; Files, A.; Graves, T.; Kirk, E.; Meredith-Neve, S.; Sanders, J.; White, B.; Swearingen, C.J. Effect of virtual reality on adolescent pain during burn wound care. *J. Burn Care Res.* **2014**, *35*, 395–408. [CrossRef]
28. Kipping, B.; Rodger, S.; Miller, K.; Kimble, R.M. Virtual reality for acute pain reduction in adolescents undergoing burn wound care: A prospective randomized controlled trial. *Burns* **2012**, *38*, 650–657. [CrossRef]
29. Jin, C.; Feng, Y.; Ni, Y.; Shan, Z. Virtual reality intervention in postoperative rehabilitation after total knee arthroplasty: A prospective and randomized controlled clinical trial. *Int. J. Clin. Exp. Med.* **2018**, *11*, 6119–6124.
30. Punt, I.M.; Ziltener, J.L.; Monnin, D.; Allet, L. Wii Fit exercise therapy for the rehabilitation of ankle sprains: Its effect compared with physical therapy or no functional exercises at all. *Scand. J. Med. Sci. Sports* **2016**, *26*, 816–823. [CrossRef]

31. Tubach, F.; Ravaud, P.; Baron, G.; Falissard, B.; Logeart, I.; Bellamy, N.; Bombardier, C.; Felson, D.; Hochberg, M.; van der Heijde, D.; et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: The minimal clinically important improvement. *Ann. Rheum. Dis.* **2005**, *64*, 29–33. [CrossRef]
32. Treede, R.D.; Rief, W.; Barke, A.; Aziz, Q.; Bennett, M.I.; Benoliel, R.; Cohen, M.; Evers, S.; Finnerup, N.B.; First, M.B.; et al. A classification of chronic pain for ICD-11. *Pain* **2015**, *156*, 1003–1007. [CrossRef]
33. Carpenter, D.M.; Nelson, B.W. Low back strengthening for the prevention and treatment of low back pain. *Med. Sci. Sports Exerc.* **1999**, *31*, 18–24. [CrossRef] [PubMed]
34. Sharar, S.R.; Alamdari, A.; Hoffer, C.; Hoffman, H.G.; Jensen, M.P.; Patterson, D.R. Circumplex Model of Affect: A Measure of Pleasure and Arousal During Virtual Reality Distraction Analgesia. *Games Health J.* **2016**, *5*, 197–202. [CrossRef] [PubMed]
35. Mosadeghi, S.; Reid, M.W.; Martinez, B.; Rosen, B.T.; Spiegel, B.M. Feasibility of an Immersive Virtual Reality Intervention for Hospitalized Patients: An Observational Cohort Study. *JMIR Ment. Health* **2016**, *3*, e28. [CrossRef] [PubMed]
36. Pourmand, A.; Davis, S.; Marchak, A.; Whiteside, T.; Sikka, N. Virtual Reality as a Clinical Tool for Pain Management. *Curr. Pain Headache Rep.* **2018**, *22*, 53. [CrossRef] [PubMed]
37. Garrett, B.; Taverner, T.; Masinde, W.; Gromala, D.; Shaw, C.; Negraeff, M. A rapid evidence assessment of immersive virtual reality as an adjunct therapy in acute pain management in clinical practice. *Clin. J. Pain* **2014**, *30*, 1089–1098. [CrossRef] [PubMed]
38. Jones, T.; Moore, T.; Choo, J. The Impact of Virtual Reality on Chronic Pain. *PLoS ONE* **2016**, *11*, e0167523. [CrossRef] [PubMed]
39. Chan, E.; Foster, S.; Sambell, R.; Leong, P. Clinical efficacy of virtual reality for acute procedural pain management: A systematic review and meta-analysis. *PLoS ONE* **2018**, *13*, e0200987. [CrossRef] [PubMed]
40. Hoffman, H.G.; Patterson, D.R.; Seibel, E.; Soltani, M.; Jewett-Leahy, L.; Sharar, S.R. Virtual reality pain control during burn wound debridement in the hydrotank. *Clin. J. Pain* **2008**, *24*, 299–304. [CrossRef]
41. Morris, L.D.; Louw, Q.A.; Crous, L.C. Feasibility and potential effect of a low-cost virtual reality system on reducing pain and anxiety in adult burn injury patients during physiotherapy in a developing country. *Burns* **2010**, *36*, 659–664. [CrossRef] [PubMed]
42. Maani, C.V.; Hoffman, H.G.; Morrow, M.; Maiers, A.; Gaylord, K.; McGhee, L.L.; DeSocio, P.A. Virtual reality pain control during burn wound debridement of combat-related burn injuries using robot-like arm mounted VR goggles. *J. Trauma* **2011**, *71*, S125–S130. [CrossRef]
43. Junginger, B.; Baessler, K.; Sapsford, R.; Hodges, P.W. Effect of abdominal and pelvic floor tasks on muscle activity, abdominal pressure and bladder neck. *Int. Urogynecol. J.* **2010**, *21*, 69–77. [CrossRef] [PubMed]
44. Hibbs, A.E.; Thompson, K.G.; French, D.; Wrigley, A.; Spears, I. Optimizing Performance by Improving Core Stability and Core Strength. *Sports Med.* **2008**, *38*, 995–1008. [CrossRef] [PubMed]
45. Park, M.; Im, H.; Kim, D.Y. Feasibility and user experience of virtual reality fashion stores. *Fash. Text.* **2018**, *5*, 1–17. [CrossRef]

Journal of
Clinical Medicine

MDPI

Review

Patient Experience in Home Respiratory Therapies: Where We Are and Where to Go

Cátia Caneiras [1,2,†,*], Cristina Jácome [3,4,†], Sagrario Mayoralas-Alises [5,6], José Ramon Calvo [6], João Almeida Fonseca [3,7,8], Joan Escarrabill [9,10,11] and João Carlos Winck [12]

1 Institute of Environmental Health (ISAMB), Faculty of Medicine, Universidade de Lisboa, 1649-028 Lisboa, Portugal
2 Healthcare Department, Praxair Portugal Gases, 2601-906 Lisboa, Portugal
3 CINTESIS-Center for Health Technologies and Information Systems Research, Faculty of Medicine, University of Porto, 4200-450 Porto, Portugal; cjacome@med.up.pt (C.J.); fonseca.ja@gmail.com (J.A.F.)
4 Respiratory Research and Rehabilitation Laboratory (Lab3R), School of Health Sciences (ESSUA), University of Aveiro, 3810-193 Aveiro, Portugal
5 Service of Pneumology, Hospital Universitario Moncloa, 28008 Madrid, Spain; sarimayoralas@gmail.com
6 Healthcare Department, Praxair Spain, 28020 Madrid, Spain; jose_ramon_calvo@praxair.com
7 MEDCIDS-Department of Community Medicine, Health Information and Decision, Faculty of Medicine, University of Porto, 4200-450 Porto, Portugal
8 Allergy Unit, Instituto and Hospital CUF, 4460-188 Porto, Portugal
9 Hospital Clínic de Barcelona, 08036 Barcelona, Spain; ESCARRABILL@clinic.cat
10 Master Plan for Respiratory Diseases (Ministry of Health) & Observatory of Home Respiratory Therapies (FORES), 08028 Barcelona, Spain
11 REDISSEC Health Services Research on Chronic Patients Network, Instituto de Salud Carlos III, 28029 Madrid, Spain
12 Faculty of Medicine, University of Porto, 4200-319 Porto, Portugal; jcwinck@mail.telepac.pt
* Correspondence: ccaneiras@gmail.com; Tel.: +351-912-556-452
† These authors contributed equally to this work.

Received: 5 March 2019; Accepted: 23 April 2019; Published: 24 April 2019

Abstract: The increasing number of patients receiving home respiratory therapy (HRT) is imposing a major impact on routine clinical care and healthcare system sustainability. The current challenge is to continue to guarantee access to HRT while maintaining the quality of care. The patient experience is a cornerstone of high-quality healthcare and an emergent area of clinical research. This review approaches the assessment of the patient experience in the context of HRT while highlighting the European contribution to this body of knowledge. This review demonstrates that research in this area is still limited, with no example of a prescription model that incorporates the patient experience as an outcome and no specific patient-reported experience measures (PREMs) available. This work also shows that Europe is leading the research on HRT provision. The development of a specific PREM and the integration of PREMs into the assessment of prescription models should be clinical research priorities in the next several years.

Keywords: Long-term oxygen therapy; home mechanical ventilation; patient-reported experience measures; quality of care; healthcare; sustainability

1. Introduction

Long-term oxygen therapy (LTOT) and/or home mechanical ventilation (HMV) are well-established therapies for patients with chronic respiratory failure, such as those with chronic obstructive pulmonary disease (COPD), neuromuscular diseases, and obstructive sleep apnea (OSA), among others. These therapies represent key services in the home respiratory therapy (HRT) provided to these patients.

Increasing numbers of patients receiving HRT are reported not only in Europe but also worldwide [1–5]. Thus, HRT is imposing a major impact on clinical care and healthcare systems. Over the next several years, the main challenge will be to ensure a sustainable healthcare system to continue to guarantee access to HRT while maintaining the quality of care.

According to the World Health Organization, quality of care is defined as "the extent to which health care services provided to individuals and patient populations improve desired health outcomes. In order to achieve this, health care must be safe, effective, timely, efficient, equitable and people-centered" [6]. A necessary step in the process of maintaining and improving quality is to monitor and evaluate the quality of healthcare in routine clinical practice. Based on the reactive, disease-focused, and biomedical model, the indicators of quality have been mainly restricted to traditional clinical metrics. A number of studies conducted over the last few decades have addressed the beneficial effects of HRT on morbidity, mortality, and adverse outcomes, as well as the variations in HRT provision among countries [5,7,8]. However, these metrics alone do not provide a complete picture of HRT quality.

The patient's experience of treatment is a cornerstone of high-quality healthcare [9]. Only by analyzing the relational and functional aspects of the patient experience is it possible to assess the extent to which patients are receiving care that is in line with their preferences, needs, and values. The integration of the patient experience with healthcare delivery and quality evaluation are key steps in moving toward patient-centered and personalized care [10]. As Doyle et al. suggested, the patient's experience is the third pillar of quality, along with clinical safety and effectiveness [11]. However, it is only in recent years that patients' perceptions of healthcare provision have started to receive attention.

This review approaches the assessment of the patient experience in the clinical context of HRT while highlighting the European contribution to this emerging body of knowledge.

2. Patient Experience in the Context of HRT

The patient experience in the context of HRT is reviewed with a focus on two main areas: (1) HRT prescription models and the inclusion of the patient experience as an outcome of these models and (2) methods used to assess the patient experience. To address these two aims, a narrative review was conducted. The search, although not systematic in nature, included searches in electronic databases (PubMed, Medline, ISI Web of Knowledge and Google Scholar), as well as hand searches (expert consultation and a review of the reference lists in the included papers). The databases were searched between July and December 2018 using topic-related terms, such as oxygen therapy, home mechanical ventilation, noninvasive mechanical ventilation, home respiratory therapy, home treatment, chronic respiratory insufficiency, chronic respiratory failure, epidemiology, prescription, quality control, outcomes, patient experience, patient perspective, carers, caregivers, patient-reported experience measure, questionnaires, interviews, and focus groups. There was no time restriction in the literature search, although it was limited to English, Portuguese, or Spanish.

2.1. Prescription Models of HRT

There are a number of studies that have assessed the prescription of HRT. Table 1 summarizes 15 relevant studies on this topic. The majority of the studies (n = 9) were conducted from 2009 onward and primarily assessed the prescription of HMV (n = 10) [4,5,12–19], followed by LTOT (n = 6) [19–24]. The estimated prevalence of HMV (from 2.5 to 23/100,000 population) and of LTOT (from 31.6 to 102/100,000 population) were variable among distinct regions or countries. The estimated prevalence of HMV in Europe was 6.6 per 100,000 people, and Portugal was one of the countries with the highest prevalence [5].

Three studies reported the assessment of HRT prescription at a regional level (Catalan, Spain; Hong Kong, China; Tasmania, Australia), eight at a national level (Sweden, Canada, Poland, Denmark, England, Australia, France, Spain), and four at an international level (two countries, seven countries, 13 European countries, 16 European countries).

Table 1. Studies assessing the prescription of home respiratory therapies.

Author, Year	Region or Country, Years Analyzed	Aim	Method	Data Collection	Results
Ekström et al., 2017 [20]	Sweden, 1987–2015	Long-term oxygen therapy (LTOT): incidence, prevalence, and the quality of prescription and management	Data from the Swedevox registry between 1 January 1987 and 31 December 2015	Data: Birth date, Sex, Primary/secondary causes of LTOT, Follow-up, Stop date and stop cause, PaO_2 air and $PaCO_2$ air, PaO_2 oxygen and $PaCO_2$ oxygen, FEV_1 and VC, World Health Organization performance status, Height and weight, Never/Past/Current smoker, Maintenance treatment with oral corticosteroids, Oxygen dose, Oxygen duration.	23,909 patients on LTOT. 48 respiratory or medicine units. Incidence of LTOT increased from 3.9 to 14.7/100,000 inhabitants over the study time period. In 2015, 2596 patients had ongoing therapeutic LTOT in the registry, a prevalence of 31.6/100,000. Adherence to prescription recommendations and fulfilment of quality criteria were stable or improved over time. Of patients starting LTOT in 2015, 88% had severe hypoxemia and 97% had any degree of hypoxemia; 98% were prescribed oxygen for ≥15 hours/day; 76% had both stationary and mobile oxygen equipment; 75% had a mean PaO_2 > 8.0 kPa breathing oxygen; and 98% were non-smokers.
Rose et al., 2015 [12]	Canada, 2012–2013	Home mechanical ventilation (HMV): national data profiling	Survey administered via a web link from August 2012 to April 2013 to service providers delivering care/services to ventilator-assisted individuals requiring daily noninvasive ventilation (NIV) or invasive mechanical ventilation via tracheostomy at home.	Survey content: provider characteristics, including services and education provided; user characteristics (age, ventilation type, primary disorder, duration of ventilation); criteria for initiation and monitoring ventilation effectiveness; equipment (ventilators and interfaces used, ventilator servicing arrangements and backup); training and education (audience, structure, topics, ongoing competency assessment); liaisons and transitions (referral, barriers to transition); follow-up (structure, frequency, location).	Response rate 152/171 (89%). 4334 ventilator-assisted individuals: an estimated prevalence of 12.9/100,000 population. 73% receiving NIV and 18% receiving intermittent mandatory ventilation (9% not reported). Services were delivered by 39 institutional providers and 113 community providers. Various models of ventilator servicing were reported. 64% of providers stated that caregiver competency was a prerequisite for home discharge, but repeated competency assessment and retraining were offered by 45%. Barriers to home transition: insufficient funding for paid caregivers, equipment, and supplies; a shortage of paid caregivers; negotiating public funding arrangements.

Table 1. *Cont.*

Author, Year	Region or Country, Years Analyzed	Aim	Method	Data Collection	Results
Escarrabill et al., 2015 [13]	Catalan Health Service (Spain), 2008–2011	HMV: prevalence and variability in prescriptions	Catalan Health Service (CatSalut) billing database, between 2008 and 2011.	Not reported (NR)	240,760 patients received some type of HRT funded by the public system. 75.8% used continuous positive airway pressure equipment, 17.3% used various forms of oxygen supply, 4.2% used nebulized therapy, 2.5% used HMV, and 0.2% used miscellaneous treatments. 6,867 patients received HMV, 23 users per 100,000 population. Rates of HMV increased by 39% over the study period
Nasiłowski et al., 2015 [14]	Poland, 2000–2010	HMV: trends over the last decade	Questionnaire designed specifically for the study was sent to the heads of nine HMV centers	Survey Content: Center details: location, area of activity (uniregional/multiregional), and year of initiating HMV. Number of subjects treated with HMV in each consecutive year. Overall number of treated subjects, divided into five disease categories: (1) neuromuscular diseases, (2) lung diseases (chronic obstructive pulmonary disease (COPD), bronchiectasis, cystic fibrosis, interstitial diseases), (3) chest-wall diseases (scoliosis, thoracoplasty, ankylosing spondylitis, post-tuberculosis sequelae), (4) hypoventilation syndromes (due to obesity, central congenital hypoventilation syndrome, central sleep apnea), (5) other diseases. Technique of ventilation (invasive and noninvasive). Number of new cases; Overall number of subjects treated with NIV or tracheostomy. Age of the treated subjects, Site where ventilation was initiated: intensive care unit, respiratory department, neurology department, general medicine department, home, or other.	Nine HMV centers, 1495 subjects Center experience 9 ± 3 years (6–13 years) One center was dedicated specifically to children, Two solely treated adults, and other centers treated subjects irrespective of age. In 2010, prevalence of HMV reached almost 2.5 subjects/100,000. The majority of subjects on HMV suffered from neuromuscular diseases (100% in 2000–2002 to 51% in 2010). Subjects with a diagnosis of respiratory failure due to pulmonary conditions appeared in 2004, and the number of subjects rapidly increased beginning in 2007. In 2010, they accounted for almost 25% of all HMV cases. Hypoventilation syndromes were the third main diagnostic group (4% until 2008, reaching 11% in 2010). Proportion of chest-wall diseases remained ~3%. In 2000 and 2001, ventilation via tracheostomy was exclusively used. The first subjects on NIV were treated in 2002. The number of subjects on NIV was 1/3 in 2004 and then leveled off for the following five years, followed by a rapid increase until 2010, when the proportions of subjects treated with NIV and tracheostomy equalized. Since 2008, the number of new cases treated noninvasively surpassed the number of new cases treated with invasive ventilation, and in 2010, the total number of subjects in both groups was virtually the same.

Table 1. *Cont.*

Author, Year	Region or Country, Years Analyzed	Aim	Method	Data Collection	Results
Garner et al., 2013 [15]	Australia and New Zealand, 2002–2004	HMV	HMV centers that had prescribed HMV for more than three months to more than five adult patients. A designed survey.	Survey Content: (1) Institutional details: location, type (e.g., tertiary), funding (e.g., government), patient catchment, years of service; (2) Criteria for HMV prescription by disease group (e.g., COPD); (3) HMV service details: number of patients receiving HMV, staffing levels, methods of implementation by location/tests utilized/staff involved, methods of follow-up by location/tests utilized/staff involved (0–3 grading from never to always), annual clinic attendances, presence of an outreach service; (4) Individual patient data (if available): age, gender, primary indication for HMV, duration of therapy, adherence to therapy, interface, machine settings (mode, inspiratory positive airway pressure, expiratory positive airway pressure, back-up rate); (5) Local database: current database for that center, data collected, what data should be collected, support for creation of a national database, center willing to participate; (6) Problems encountered with setting up an HMV service.	28 centers (82%) responded, providing data on 2725 patients. Prevalence of HMV was 9.9 patients/100,000 in Australia and 12.0 patients/100,000 in New Zealand. Variation existed among Australian states (range 4–13 patients/100,000) correlating with population density ($r = 0.82$, $p < 0.05$). The commonest indications for treatment were obesity hypoventilation syndrome (31%) and neuromuscular disease (30%). COPD was an uncommon indication (8%). No consensus on indications for commencing treatment was found.

Table 1. *Cont.*

Author, Year	Region or Country, Years Analyzed	Aim	Method	Data Collection	Results
Ringbaek et al., 2013 [21]	Denmark, 2001–2010	(LTOT: incidence, prevalence, treatment modalities, and survival in COPD.	Danish Oxygen Register in the period from 01 January 2001 to 31 December 2010: information on patients on home oxygen therapy, their prescriptions, and termination of therapy. National Health Services Central Register: information on diagnosis for LTOT and on vital status up to 31 December 2011.	NR	On 31 Dec 2001, a total of 2247 COPD patients (42.0/100,000) were receiving LTOT. The number of patients on LTOT had increased constantly to reach a prevalence of 48.1/100,000 in 2010. Incidence of oxygen therapy increased insignificantly from 30.5 to 32.2/100,000. The majority of COPD patients were women and older than 70 years of age. The mean age of patients who started LTOT during the study period increased from 73.4 ± 9 years to 74.8 ± 9.7 years. Most of the COPD patients were prescribed oxygen therapy by a hospital doctor immediately after an acute hospitalization, and the number of prescriptions from general practitioners was continuously declining toward zero during the study period. An increasing number of the COPD patients were prescribed oxygen at least 15 h daily and had delivered oxygen concentrator and mobile oxygen, whereas, in general, the oxygen flow remained low (≤1.5 L/minute). Compared with men, women started LTOT more often in connection with hospitalization and more often stopped LTOT within the first 6 months. Women were prescribed a lower oxygen flow than men and the treatment was more often specified to take place for 15–24 h per day.

Table 1. *Cont.*

Author, Year	Region or Country, Years Analyzed	Aim	Method	Data Collection	Results
Mandal et al., 2013 [16]	England, NA	HMV: prevalence of sleep and ventilation diagnostic and treatment services	A short survey delivered by email to 101 NHS Hospitals	Survey content: 10-item survey, focused on diagnostic services and HMV provision: (a) availability of diagnostics, (b) funding; (c) patient groups.	76 (68%) responses received; 42 (55%) trusts reported the provision of an HMV service. Only 65% of units charged for the delivery of an HMV service, with 12% of these services commissioned by an external provider. Median set-up frequency for the units charging was 42 patients per annum (interquartile range 23–73), whereas those units that failed to charge had a median of 11 (interquartile range 4–22). Of all the HMV set-ups, 67% were for obesity-related respiratory failure and COPD, with the other restrictive lung conditions forming the remainder
Serginson et al., 2009 [22]	Australia, 2004–2005	LTOT: prescription and costs	Data from all LTOT services in Australian Government's departments and health services (state and federal) Centralized departments managing state budgets for LTOT provided costs (for the financial year 2004–2005) and patient numbers (point prevalence in 2005). If centralized data were not available, regional departments administering LTOT services were contacted.	Data: Costs were defined as "equipment only" (fees paid to oxygen companies) or "equipment and administrative" (wages and non-labor costs of administering programs included).	20,127 patients (100/100,000) through 59 different services at a cost of over $31 million. Prescription rates for LTOT per 100,000 population within each state ranged from 44 to 133, a threefold difference. Costs of LTOT per patient prescribed per year funded by individual states and territories ranged from $1014 to $2574. The cost of oxygen concentrators averaged $85 per month (range, $29–$109), portable oxygen ranged from $16 to $35 per month without refills, and, with a conserver included, $55 (two refills) to $166 unlimited refills) per month. All services provided concentrators for home use. Portable oxygen was funded in all states, except one (where it was limited to children and patients waiting for heart or lung transplants).

Table 1. *Cont.*

Author, Year	Region or Country, Years Analyzed	Aim	Method	Data Collection	Results
Jones et al., 2007 [23]	Tasmania (Australia), 2002–2004	LTOT	Records of all patients receiving Tasmanian Government-funded LTOT between December 2002 and April 2004	Data: Recipient demographics, Indications for LTOT, Oxygen prescription, Time to follow-up. The service provider provided usage reports and costs.	April 2004: 490 patients receiving LTOT Rate of 102/100,000; Median age at prescription of LTOT was 71.5 (range 0.7–97.2) years, and 54% of patients were female. Oxygen was prescribed for 267 patients (54%) during hospitalization, although only 192 of these patients (72%) met criteria for oxygen use at this time. LTOT was prescribed by respiratory physicians for 248 patients (51%) and by other hospital physicians for most of the remaining patients (39%). Data on indications were available for 430 patients (88%), and COPD accounted for 48% of prescriptions, but this proportion varied regionally. Median time to reassessment was 5.5 (range, 0.1–116) months, but varied between regions. Usage data were available for 175 patients (41%) using oxygen concentrators in April 2004. Of these 175 patients, 122 (70%) were prescribed oxygen for COPD. In this group, the median use was 18.3 (range, 0.38–24) hours per day; however, 36 (30%) had a median use < 15 hours/day.

Table 1. *Cont.*

Author, Year	Region or Country, Years Analyzed	Aim	Method	Data Collection	Results
Lloyd-Owen et al., 2005 [5]	16 European countries (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Norway, Poland, Portugal, Spain, Sweden, UK), 2001–2002	HMV: patterns of use across Europe	Questionnaire of center details, HMV user characteristics and equipment choices sent to selected HMV centers	Survey Content: Center (type of institution and year of starting HMV), Number of HMV users on 01 July 2001, Users' characteristics (sex, age, and time on HMV). Users' causes for respiratory failure: (1) Lung: lung and airway diseases: COPD, cystic fibrosis, bronchiectasis, pulmonary fibrosis, and pediatric diseases, including bronchopulmonary dysplasia; (2) Thor: thoracic cage abnormalities: early-onset kyphoscoliosis, tuberculosis sequelae such as thoracoplasty, obesity hypoventilation syndrome, and sequelae of lung resection; (3) Neur: neuromuscular diseases: muscular dystrophy, motor neuron disease (including amyotrophic lateral sclerosis), post-polio kyphoscoliosis, central hypoventilation, spinal cord damage, and phrenic nerve paralysis. Type of ventilator and interface used.	329 centers completed surveys, 21,526 HMV users; Estimated prevalence of HMV was 6.6/100,000 in the 16 European countries. Differences between countries in the relative proportions of (1) lung and neuromuscular patients using HMV and (2) the use of tracheostomies in lung and neuromuscular HMV users. Lung users were linked to an HMV duration of <1 year, thoracic cage users with 6–10 years of ventilation and neuromuscular users with a duration of ≥6 years. Almost all of the HMV users had positive pressure ventilators, with only 0.005% (79 users) having other types. Volume preset positive pressure ventilators were used the least for lung problems and most frequently for neurological problems (% volume: Lung 15%; Thor 28%; Neur 41%). Overall, 13% of the survey population had ventilation via a tracheostomy with the highest percentage in neuromuscular patients (Neur 24%; Thor 5%; Lung 8%).
Chu et al., 2004 [17]	Hong Kong (China), 2002	HMV	Survey to consultants of respiratory medicine in all adult medical departments of Hong Kong Hospital Authority hospitals to report their adult patients (>18 years) who had ever been managed by HMV	Survey content: demographic data, mode of ventilation (non-invasive or tracheostomy ventilation), underlying disease, indications for HMV, time of starting ventilation, time and reason of stopping ventilation, if any, in the follow-up period.	249 cases reported to the survey from 14 centers of adult respiratory medicine; 156 males (62.7%) and 93 females (37.3%) with a mean age of 62.7 ± 13.8 years; 80% of HMV cases were under the care of six major centers. 197 cases were continuing with HMV, corresponding to ~2.9 HMV users per 100,000 population. The majority (n = 236, 94.8%) were treated by noninvasive ventilation (NIV), with the remaining 13 patients (5.2%) receiving tracheostomy ventilation. All NIVs were provided by bilevel pressure-support ventilators. All tracheostomized cases were put on HMV after repeated failures to wean. The disease conditions for which HMV was prescribed: COPD (121, 48.6%); Complicated obstructive sleep apnea/obesity hypoventilation syndrome (43, 17.2%); and Restrictive thoracic disorders (85, 34.1%).

Table 1. *Cont.*

Author, Year	Region or Country, Years Analyzed	Aim	Method	Data Collection	Results
Fauroux et al., 2003 [18]	France, 2000	Domiciliary non-invasive mechanical ventilation (NIMV) in children	Anonymous national cross-sectional Survey A postal questionnaire sent by the Paediatric Group of the National Home Care Organization (ANTADIR) in 1999 to all 64 senior pediatric respiratory, neurology, and intensive care physicians in France. Patients aged < 18 years and receiving home NIMV were included in the study.	All physicians taking care of children with NIMV were sent a second questionnaire in 2000. The specific information requested on each patient included: Sex and date of birth; Primary and secondary diagnosis; Symptoms that justified NIMV; Age at onset of NIMV; Type of nasal mask, ventilatory mode, and concurrent use of oxygen therapy; Investigations performed before initiating of NIMV and during follow-up.	102 patients from 15 centers: 4/15 centers cared for 84% of patients; 7% of patients were under 3 years; 35% were 4–11 years; and 58% were >12 years. Underlying diagnoses included neuromuscular disease (34%), obstructive sleep apnea and/or craniofacial abnormalities (30%), cystic fibrosis (17%), congenital hypoventilation (9%), scoliosis (8%), and other disorders (2%). NIMV was started because of nocturnal hypoventilation (67%), acute exacerbation (28%), and/or failure to thrive (21%). Volume-targeted ventilation was preferred in restrictive disorders (56%) and central hypoventilation (56%), while pressure support ventilation (PSV) was preferred in cystic fibrosis (71%). Patients with obstructive sleep apnea and/or craniofacial abnormalities were ventilated with continuous positive airway pressure (45%) or bilevel PSV (52%).

Table 1. *Cont.*

Author, Year	Region or Country, Years Analyzed	Aim	Method	Data Collection	Results
Wijkstra et al., 2001 [24]	Seven countries (Brazil, Canada, France, Italy, Spain, Netherlands, USA), NR	LTOT: prescription	Questionnaire mailed to 100 randomly selected respirologists from a list of respiratory specialists belonging to a professional organization in each country	Characteristics of the respirologists: Date of birth; How many years they had been practicing respiratory medicine; Number of patients for whom they prescribed oxygen for the first time or for renewal purposes over the previous month. Prescription of oxygen at rest; Whether they prescribed a standard oxygen flow rate for all their patients or whether they individualized flow rates with or without specific testing of each patient; How the recommended oxygen flow at rest was chosen (either tested at rest or tested during exercise); The position (sitting, semirecumbent, supine) in which the patients were tested, the target level of arterial oxygen saturation (SaO_2) used to establish an oxygen prescription and the percentage of time during the measure in which this target had to be achieved. Prescription of oxygen during sleep and exercise; How they prescribed oxygen during sleep and exercise; The type of exercise test (walking, laboratory testing) used to establish the exercise prescription; The target level of saturation during exercise and the percentage of time during the test in which this target had to be achieved.	81% of respondents individualized the oxygen prescription at rest. Resting SaO_2 was most commonly targeted at 90–91%. The approach to night prescription varied. Respirologists in Canada and the USA increased the resting SaO_2 by 1–2 L/min during sleep, while those in Spain used the resting flow for the night prescription (62%). Respirologists in the Netherlands, France, and Italy individualized the night prescription more frequently. Although oxygen during exercise was individualized in most countries (74%), significant differences remained among countries. 62% of respirologists (62%) aimed to achieve an SaO_2 of 90–91% during exercise, while 70% of all respirologists tried to achieve the desired SaO_2 for 90% of the test.

Table 1. *Cont.*

Author, Year	Region or Country, Years Analyzed	Aim	Method	Data Collection	Results
de Lucas Ramos et al., 2000 [4]	Spain, 1998–1999	HMV: prescription	Questionnaire mailed to the respiratory medicine departments of 200 hospitals in the public health system	Survey Content: Center name, Year of initiation of the MV program, Number of patients in the first year, Number of patients in the current year. Diagnosis: Neuromuscular disease, Thoracic cage disease, Hypoventilation-obesity syndrome, COPD, Other. Ventilation type: Volumetric, BI-level Positive Airway Pressure, Interface, Nasal mask, Conventional, Personalized, Tracheostomy, Mouthpiece.	43 hospitals, 1821 patients; 813 patients had restrictive disease due to thoracic cage disease, 452 neuromuscular disease, 271 hypoventilation-obesity syndrome, 162 COPD, and 123 other diseases/conditions. 965 (53%) used pressure support devices and 856 (47%) used volumetric ventilators. 1320 conventional nasal mask, 336 personalized nasal mask, 118 tracheostomy, 41 facial mask, six mouthpiece.

Table 1. *Cont.*

Author, Year	Region or Country, Years Analyzed	Aim	Method	Data Collection	Results
Fauroux et al., 1994 [19]	13 European countries (Belgium, Denmark, England, France, Germany, Ireland, Italy, Netherlands, Norway, Poland, Spain, Sweden, Switzerland), 1992	Home care of chronic respiratory insufficiency	Questionnaire at the end of 1992.	Questionnaire content: Home treatments (LTOT, HMV); Prescribers; Practical organization of home care (supply of material, supervision of patients and equipment). Information on patients: diagnostic information (either obstructive, restrictive, or mixed pulmonary disease); Age; Sex; Equipment supplied; Service provided; Therapeutic schedules.	Information was easier to obtain for LTOT than for HMV. In all countries, both adults and children received LTOT at home for lung diseases and other less common problems, such as chest-wall deformities and sequelae of tuberculosis. Oxygen concentrators were used preferentially in all countries except Italy (80% of the patients received liquid oxygen), Denmark, Spain, and the Netherlands (cylinders were used by 80% of the patients). Both adults and children received HMV at home for chronic lung disease, neuromuscular disease, chest-wall deformities, and central hypoventilation in all countries, except in Denmark and Poland, where this treatment is almost unknown in the home. Home ventilator treatment was generally performed by volume-cycled ventilators. National prescription rules existed in some parts of Spain, Switzerland, and Belgium. In other countries, such as Germany, prescriptions relied on recommendations elaborated by specialists or international guidelines. Service and equipment were provided by national organizations, health services, commercial companies, or hospitals. Home supervision of the patient was performed by a nurse and/or a doctor and equipment maintenance by a technician.

Most studies included both children and adult patients in their analysis. Only one of the studies specifically focused on a pediatric population [18]. Questionnaires, having been used in 10 studies, were the preferred method of data collection. In five studies, existing databases from HRT registries or health services were used. Irrespective of the data collection method used, data on users (age, sex, and diagnosis), type and duration of respiratory therapy, and equipment and interfaces were the most commonly recorded. None of the 15 studies reported the patient's experience with HRT.

2.2. Assessment of Patient Experience

Assessing the patient experience has become a common approach to describing healthcare from the patient's point of view, evaluating the process of care, and measuring the outcome of care [25–27]. Both quantitative and qualitative methods are being used to assess patients' perception. Self-reported questionnaires, individual interviews, and focus groups are among the most frequently used methods of collecting data.

2.2.1. Patient-Reported Experience Measures

The development of self-reported questionnaires, namely, patient-reported experience measures (PREMs) and patient-reported outcome measures (PROMs), has exponentially increased in the last several years. These two types of questionnaires collect information about the patient's perspective but with distinct purposes. A PREM evaluates patients' perception of their personal experience of the healthcare received, while a PROM assesses the perception of their health status and health-related quality of life [10,28]. A combination of PROMs and PREMs is essential to fully understand the performance of healthcare systems. Moreover, both measures are useful to provide a patient-centered perspective of healthcare, but PREMs are more adequate to assess experience with healthcare.

Distinct instruments to assess the patient's experience with healthcare are available. Table 2 summarizes 14 instruments designed to assess the patient's experience with the provision of care in different clinical settings [29–34], hospital [35–38], primary care [39,40], intermediate care [41], and community [33,41]. The majority of such instruments are generic and designed to be used for a diverse range of health conditions. However, two of the described questionnaires were specifically developed for patients with chronic diseases [29,34], and one was intended particularly for patients with COPD [30]. The majority of PREMs were developed to target adult patients and tested in patients who were at least 15 years old. Only two developed instruments were tested with the carers of children [31,39]. English is the most common language used, with some instruments also in Norwegian [31,38,39], Italian [35,41], and Spanish [29]. Most instruments already had some of their psychometric properties explored, namely, their reliability and validity.

None of the instruments above were specifically designed to assess the patient's experience with HRT. However, a recent European Respiratory Society (ERS)/European Lung Foundation (ELF) survey was conducted across 11 European countries and assessed the attitudes and preferences of 687 patients on HMV and those of 100 carers [42]. A questionnaire was specifically developed for this study in eight languages (English, German, Dutch, Spanish, Italian, Portuguese, Greek, and French) and explored four areas: (1) patients' demographic and clinical characteristics; (2) issues influencing compliance, such as interface comfort, abilities to travel, sleep, and socialize with a ventilator, type and technical functioning of the ventilator (e.g., alarms, ability to operate and change settings, on/off switches, and electricity consumption); (3) support, training, and education; and (4) requests for improved devices and support.

Today, it is possible to evaluate a patient's perception of the HRT received using one of the described PREMs. Nevertheless, in the near future, the aim should be to develop a specific PREM to assess patients' personal experience with HRT.

Table 2. Instruments designed to assess patient's experience with the provision of care.

Instrument	Population	Setting	Language	Concepts	Structure	Measurement Properties
CEFIT: Care Experience Feedback Improvement Tool [36]	Tested in 802 patients (≥18 years) with healthcare experience.	Hospital	English	Safe, Timely, System navigation, Caring, Effective.	Five questions scored using a five-point scale, from 1 (never) to 5 (always).	Reliability Validity
COPD PREM9: disease-specific patient-reported experience measure in COPD [28,30]	Tested in 174 adult patients with COPD.	Clinical settings (e.g., pulmonary rehabilitation, nurse-led clinics, or GP annual reviews)	English	Everyday life with COPD, Everyday care in COPD, Self-management of COPD, exacerbations.	Nine questions scored using a six-point scale, from 0 (good experience) to 5 (bad experience).	Reliability Validity
GS-PEQ: Generic Short Patient Experiences Questionnaire [31]	Tested in 1324 patients (including outpatients undergoing rehabilitation and carers of children).	Services provided in a range of specialist healthcare (in- and out-patient)	Norwegian	Outcome, Clinician services, User involvement, Incorrect treatment, Information, Organization, Accessibility.	10 questions scored using a five-point scale, from 1 (Not at all/Not important) to 5 (To a very large extent/Of utmost importance).	Not reported
howRwe (how are we doing?)questionnaire: short generic patient experience questionnaire [32,43]	Tested in 828 patients in an orthopedic pre-operative assessment clinic [32] and in 90 adult patients (≥18 years) from general practices (10 with COPD) [43].	Generic, applicable without change across all patient categories and care settings, including primary, secondary, community, emergency, domiciliary, and social care.	English Dutch	Clinical care (kindness and communication), Organization of care (promptness and organization).	Four items scored using a four-point scale from 0 (poor) to 3 (excellent).	Reliability Validity

Table 2. *Cont.*

Instrument	Population	Setting	Language	Concepts	Structure	Measurement Properties
Health Services OutPatient Experience (HSOPE): global outcome measure of perceived patient-centeredness of the outpatient healthcare pathway [35]	Tested in 1532 adult outpatients (≥16 years) receiving care (including rehabilitation).	Hospital	Italian	Perceived technical effectiveness of the staff, Information on modalities of the outpatient visit, on the visit outcomes, and the course of the healthcare pathway, Relational aspects of outpatient–staff interaction, Involvement in decision making.	10 statements scored using a five-point Likert scale from 1 (never) to 5 (always) 1 item scored using a 10-point scale from 1 (very dissatisfied) to 10 (very satisfied). Three sociodemographic questions (sex, age, and residence). One question about suggestions to improve outpatient visits.	Reliability Validity
Intermediate care-IC-PREMs: Bed-Based Patient-Reported Experience Measure [41]	Tested in 1832 adult patients.	Bed-based IC services	English Italian	Goal Setting, Empowerment, Self-Management, Care-Planning, Transitions, Decision Making, Communication.	15 questions scored using two, three, or four response categories.	Reliability Validity
IC-PREMs: home-based (and reablement-based) Patient-Reported Experience Measure [41]	Tested in 4627 adult patients.	Home-based or reablement IC services	English Italian	Goal Setting, Empowerment, Self-Management, Care-Planning, Transitions, Decision Making, Communication.	15 questions scored using two, three, or four response categories.	Reliability Validity

Table 2. *Cont.*

Instrument	Population	Setting	Language	Concepts	Structure	Measurement Properties
IEXPAC, Instrument for Evaluation of the Experience of Chronic Patients [29]	Tested in 356 patients (≥16 years) with chronic diseases (20% with COPD).	Health and social services	Spanish	Type and scope of patient and professional interactions oriented to patient activation. Patient's self-management capacity of his/her wellbeing resulting from the interventions received. New relational model of the patient with the system through the internet or with partners in group intervention.	11 + 1 items scored using a five-point scale from 0 (never) to 10 (always). Since 2018, a new version with 11 + 4 items is used, with three additional items.	Reliability Validity
LifeCourse experience tool [33]	Tested in 607 adult patients with emergency department and in-patient utilization, advanced primary diagnosis of heart failure, cancer, or dementia.	Home, Nursing Homes, Assisted living	English	Care Team, Communication, Care Goals.	22 items scored using a four-point scale from 1 (Never or Strongly Disagree) to 4 (Always or Strongly Agree).	Reliability Validity
Multidimensional Semantic Patient Experience Measurement Questionnaire [37]	Tested in 60 patients (≥15 years) undergoing a magnetic resonance scan.	Hospital	English	Evaluation/valence, Potency/control, Activity/arousal, Novelty.	12 rating scales using a seven-point bipolar attribute rating scales: 'extremely', 'quite', 'slightly', 'neither', 'slightly', 'quite', and 'extremely'.	Reliability

Table 2. *Cont.*

Instrument	Population	Setting	Language	Concepts	Structure	Measurement Properties
PEQ: Patient experience questionnaire 2001 [39]	Tested in 1092 patients (1–91 years)/carers	Primary care	Norwegian	Communication, Emotions, Short-term outcome, Barriers, Relations with auxiliary staff.	Total 18 items: Four items using a five-point scale from 1 ('no more' or 'nothing') to 5 ('much more' or 'a lot'). 10 items using a five-point scale from 1 (disagree completely) to 5 (agree completely). Four items were formed on seven-point scales.	Reliability Validity
PEQ: Patient Experiences Questionnaire 2004 [38]	Tested in 19578 patients (≥16 years) with experience with surgical wards and wards of internal medicine	Hospital	Norwegian	Information on future complaints, Nursing services, Communication, Information examinations, Contact with next-of-kin, Doctor services, Hospital and equipment, Information medication, Organization, General satisfaction.	35 items with 10-point ordinal response scales from 1 (negative) to 10 (positive).	Reliability Validity
PACIC: Patient Assessment of Chronic Illness Care [34,44]	Tested in 4108 adult patients with diabetes, chronic pain, heart failure, asthma, coronary artery disease.	Chronic care management	English	Patient activation, Delivery system design, Goal setting, Problem solving, Follow-up/coordination. Focuses on the receipt of patient-centered care and self-management behaviors.	20 items using a five-point scale from 1 (Almost Never) to 5 (Almost Always).	Reliability Validity

Table 2. *Cont.*

Instrument	Population	Setting	Language	Concepts	Structure	Measurement Properties
ACES-SF: Ambulatory Care Experiences Survey [40]	Tested in 49,861 adult patients.	Primary care	English	Quality of physician–patient interaction, Health promotion support, Care coordination, Organizational access, Office staff interactions, An additional item to assess patients' willingness to recommend the physician to family and friends.	18 items using continuous responses: Never, Almost never, sometimes, Usually, Almost always, Always; or Yes, definitely, Yes, somewhat, No, definitely not; or Definitely yes, Probably yes, Not sure, Probably not, Definitely not.	Reliability Validity

2.2.2. Individual Interviews and Focus Groups

Qualitative studies that explore the experience of patients receiving HRT are still limited in the literature. Nevertheless, the literature review revealed some studies that explored the experience of patients living with COPD, pulmonary fibrosis, and OSA. These studies specifically focused on patients' needs and the adaptation process to respiratory therapies. Two studies explored the patient's experience with LTOT [45,46], and the others assessed the patient's experience with non-invasive ventilation [47–51]. These studies were conducted in the United States of America [45,47], New Zealand [48,49], the United Kingdom [50], Sweden [51], and Spain [46] and included both adult patients and carers. Two reviews were also found on the needs of patients with COPD and were also used in the present analysis [52,53].

From the analysis of these studies, it was possible to clearly identify education, training, support, and carer involvement as important key-points in facilitating a patient's treatment experience and subsequent adherence. Below, each one of these four key-points is described in detail.

Education: on the basis of the perspectives of patients, it is apparent that education is crucial for defining clear expectations about the treatment and motivating patient adherence. The main education topics raised by patients receiving respiratory therapies are related to disease self-management (e.g., COPD, OSA); physical effects and potential clinical benefits of the respiratory therapy; risks of not using the respiratory therapy; guidance on the use and function of equipment (e.g., continuous positive airway pressure (CPAP) devices, oxygen concentrators, how to use pulse oximeters and adjust flow with exertion); side effects and guidance on its management (skin protection, dry mouth, nasal congestion, irritated eyes); traveling with equipment; follow-up appointments; and assistance with financial elements (e.g., how to claim electricity costs) [45,46,49,50].

Training: formal training on appropriate equipment use has been suggested to be an important strategy for improving adherence [46–51]. Healthcare professionals need to introduce the device, explore possible practical problems, and give advice/help to solve these problems. In their initial experiences with respiratory therapy, patients should have a hands-on demonstration for setting up the device, trialing different masks/pressures, making mask adjustments, conquering different side-effects, and finding the best position for the tubing or machine (also considering the loudness of the device). Regular follow up visits or phone calls are important to assess practical problems being experienced (e.g., pressure from the mask, mask leakage, disturbing noise, and difficulties changing sleeping positions) and to discuss effective strategies to address them.

Support: establishing a trustworthy relationship with healthcare professionals after the initiation of respiratory therapy is perceived as helpful by patients, and these relationships positively influence their adherence [46]. Healthcare professionals need to foster a non-judgmental environment in which patients have opportunities to ask questions, share concerns and feelings, feel listened to, and feel understood. This is particularly important following the initiation of therapy [47], as questions or concerns are more likely to arise during the first days or weeks of treatment [49,52]. These opportunities can arise during regular follow-up visits, scheduled follow-up phone calls, and through access to a 24-h hotline [47].

Carer involvement: carers provide substantial care (emotional, physical) to the individual on a daily basis and, most of the time, live in the same house as the patient. On the basis of their important role in patients' lives, carer involvement has been found to be essential to patients receiving HRT [45–48,50–53]. Patients recognize that carers play a major role in their treatment by helping them manage the disease and adapt to the equipment (e.g., verbal reminders, encouragement, setting up the machine, making mask adjustments, reassurance of therapy benefits). Carers themselves recognize their need for information regarding aspects of the disease and benefits of the HRT [47]. Carer involvement is thus perceived by all stakeholders as an essential component of education and training from the beginning of treatment [45,47,48,50–53], and it is generally associated with positive results, namely, the patients' adoption and adherence to HRT [47,53].

3. Discussion

This comprehensive review is a first critical step toward the assessment of the patient experience in the clinical context of HRT. It demonstrates that research in this area is still limited, with no example of an HRT prescription model that incorporates the patient experience as an outcome and with no specific PREM available. This review also shows that European countries have been involved in HRT provision research from an early stage.

Most of the research on the assessment of HRT prescription models has been conducted within the last decade and mainly in European countries, highlighting the emergent interest and Europe's leading position in this area of health research. In addition, HMV has attracted more attention from the scientific community in comparison with LTOT. Questionnaires were found to be the preferred method for data collection, however, existing databases from HRT registries or health services have also been used. Databases in comparison with questionnaires have the advantage of generating more representative data and may be a method of choice in future studies. The patient experience has not been examined in the assessment of the prescription models presented. While this reality was expected from the oldest studies, it was quite a surprising result for those from the last decade. These results show that, until now, the assessment of patients' perceptions has not been seen as a priority in the assessment of prescription models. Unfortunately, this is also a reality in other health contexts and settings [10]. The Organisation for Economic Co-operation and Development (OECD) and Europe in "Health at a Glance: Europe 2018" reported critical gaps in the data on patient-reported experience, and they recommended collecting data on the patient experience from any doctor in ambulatory care settings [10]. Thus, future studies on the provision of HRT should address this important gap in the literature.

To address this gap, we need to be aware of the current methods being used to assess the patient experience. Different instruments used at distinct levels of healthcare are available and described in this review. These instruments were developed to be completed by adult patients and, in some cases, by carers of children. In our opinion, although the carers' perspective is, of course, incredibly valuable, it should do not replace the children's experience. The development of PREMs for pediatric populations is crucial to the collection of information on the experience and outcome of children's care. Additionally, as previously mentioned, none of the instruments have been specifically designed to assess the patient's experience with HRT. The development of a specific PREM for this health context should be a research priority in the upcoming years. The most commonly assessed domains in the described instruments, including the ERS/ELF survey, together with the key facilitators of the patient's treatment experience, can be used as important sources of data to inform the development of a comprehensive instrument. Access to information and support, implementation of effective and clear communication, active participation in shared decision making, enhanced accessibility and navigability across the healthcare system for patients and families, particularly across transitional care, and management of polypharmacy are known to influence the patient experience in other healthcare settings and could be topics of interest to be included in future PREMs for patients on HRT [54]. Future studies should explore which of these raised topics are indeed meaningful for patients and carers.

On the basis of qualitative studies, it was found that education, training, support, and carer involvement were important key-points in facilitating the patient's treatment experience and adherence. This knowledge comes mainly from the perspective of adult patients with COPD, pulmonary fibrosis, and OSA receiving CPAP and from their carers. These studies were conducted in five countries (three from Europe) [45–53]. Thus, this evidence may not completely apply to the experience of younger patients (including children) and that of their carers or to patients with other diseases and other treatment modalities (e.g., Bilevel Positive Pressure Airway, LTOT) and from other countries/continents. Considering these identified gaps, the experience of other patients receiving HRT could be explored in future studies. The identified key-points may inform the development process of semi-structured guides of focus groups or individual interviews to be used in these exploratory studies.

4. Conclusions

To the authors' best knowledge, this is the first published work to review the emerging topic of the patient experience in the clinical context of HRT and give important insights into the status of this clinical research area while also pointing out possible directions in which to move to realize patient-centered care. The assessment of the patient experience is in its early stages, and further research is needed to integrate these measures with routine healthcare delivery and the core set of healthcare quality indicators, as well as and to drive quality improvements in HRT.

Author Contributions: Conceptualization, C.C., C.J., J.A.F., J.E., and J.C.W.; writing—original draft preparation, C.C. and C.J.; writing-review and editing, S.M.A., J.R.C., J.A.F., J.E., and J.C.W.; project administration, C.C., S.M.A., J.R.C., and J.C.W.

Funding: C.J. has a post-doctoral grant (SFRH/BPD/115169/2016) funded by Fundação para a Ciência e a Tecnologia (FCT), co-financed by the European Social Fund (POCH) and Portuguese national funds from MCTES (Ministério da Ciência, Tecnologia e Ensino Superior).

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

References

1. Cranston, J.; Crockett, A.; Moss, J.; Alpers, J. Domiciliary oxygen for chronic obstructive pulmonary disease. *Cochrane Database Syst. Rev.* **2005**, *4*, CD001744. [CrossRef] [PubMed]
2. Annane, D.; Orlikowski, D.; Chevret, S. Nocturnal mechanical ventilation for chronic hypoventilation in patients with neuromuscular and chest wall disorders. *Cochrane Database Syst. Rev.* **2014**, *12*, CD001941. [CrossRef] [PubMed]
3. Melo, I. Alguns dados sobre a Assistência Respiratória Domiciliária em Portugal. *Rev. Port. Pneumol.* **1997**, *3*, 481–492. [CrossRef]
4. De Lucas Ramos, P.; Rodríguez González-Moro, J.M.; Santa-Cruz Siminiani, A.; Cubillo Marcos, J.M.; Paz González, L. Estado actual de la ventilación mecánica domiciliaria en España: Resultados de una encuesta de ámbito nacional. *Arch. Bronconeumol.* **2000**, *36*, 545–550. (In Spanish) [CrossRef]
5. Lloyd-Owen, S.J.; Donaldson, G.C.; Ambrosino, N.; Escarabill, J.; Farre, R.; Fauroux, B.; Robert, D.; Schoenhofer, B.; Simonds, A.K.; Wedzicha, J.A. Patterns of home mechanical ventilation use in Europe: Results from the Eurovent survey. *Eur. Respir. J.* **2005**, *25*, 1025–1031. [CrossRef] [PubMed]
6. World Health Organization. What is Quality of Care and Why Is It Important? Available online: http://www.who.int/maternal_child_adolescent/topics/quality-of-care/definition/en/ (accessed on 11 April 2019).
7. Simonds, A.K. Home Mechanical Ventilation: An Overview. *Ann. Am. Thorac. Soc.* **2016**, *13*, 2035–2044. [CrossRef]
8. Dogan, O.T.; Turkyilmaz, S.; Berk, S.; Epozturk, K.; Akkurt, I. Effects of long-term non-invasive home mechanical ventilation on chronic respiratory failure. *Curr. Med. Res. Opin.* **2010**, *26*, 2229–2236. [CrossRef]
9. NHS Department of Health. *High Quality Care for All: NHS Next Stage Review Final Report*; Norwich, UK, 2008. Available online: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/228836/7432.pdf (accessed on 11 April 2019).
10. OECD/EU. Health at a Glance: Europe 2018: State of Health in the EU Cycle. OECD Publishing: Paris, France, 2018. Available online: https://doi.org/10.1787/health_glance_eur-2018-en (accessed on 11 April 2019).
11. Doyle, C.; Lennox, L.; Bell, D. A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. *BMJ Open* **2013**, *3*, e001570. [CrossRef]
12. Rose, L.; McKim, D.A.; Katz, S.L.; Leasa, D.; Nonoyama, M.; Pedersen, C.; Goldstein, R.S.; Road, J.D. Home mechanical ventilation in Canada: A national survey. *Respir. Care* **2015**, *60*, 695–704. [CrossRef]
13. Escarrabill, J.; Tebe, C.; Espallargues, M.; Torrente, E.; Tresserras, R.; Argimon, J. Variability in home mechanical ventilation prescription. *Arch. Bronconeumol.* **2015**, *51*, 490–495. [CrossRef]
14. Nasilowski, J.; Wachulski, M.; Trznadel, W.; Andrzejewski, W.; Migdal, M.; Drozd, W.; Pytel, A.; Suchanke, R.; Czajkowska-Malinowska, M.; Majszyk, T.; et al. The evolution of home mechanical ventilation in poland between 2000 and 2010. *Respir. Care* **2015**, *60*, 577–585. [CrossRef] [PubMed]

15. Garner, D.J.; Berlowitz, D.J.; Douglas, J.; Harkness, N.; Howard, M.; McArdle, N.; Naughton, M.T.; Neill, A.; Piper, A.; Yeo, A.; et al. Home mechanical ventilation in Australia and New Zealand. *Eur. Respir. J.* **2013**, *41*, 39–45. [CrossRef] [PubMed]
16. Mandal, S.; Suh, E.; Davies, M.; Smith, I.; Maher, T.M.; Elliott, M.W.; Davidson, A.C.; Hart, N. Provision of home mechanical ventilation and sleep services for England survey. *Thorax* **2013**, *68*, 880–881. [CrossRef]
17. Chu, C.M.; Yu, W.C.; Tam, C.M.; Lam, C.W.; Hui, D.S.; Lai, C.K. Home mechanical ventilation in Hong Kong. *Eur. Respir. J.* **2004**, *23*, 136–141. [CrossRef]
18. Fauroux, B.; Boffa, C.; Desguerre, I.; Estournet, B.; Trang, H. Long-term noninvasive mechanical ventilation for children at home: A national survey. *Pediatr. Pulmonol.* **2003**, *35*, 119–125. [CrossRef] [PubMed]
19. Fauroux, B.; Howard, P.; Muir, J.F. Home treatment for chronic respiratory insufficiency: The situation in Europe in 1992. The European Working Group on Home Treatment for Chronic Respiratory Insufficiency. *Eur. Respir. J.* **1994**, *7*, 1721–1726. [CrossRef]
20. Ekstrom, M.; Ahmadi, Z.; Larsson, H.; Nilsson, T.; Wahlberg, J.; Strom, K.E.; Midgren, B. A nationwide structure for valid long-term oxygen therapy: 29-year prospective data in Sweden. *Int. J. Chron. Obstruct. Pulmon. Dis.* **2017**, *12*, 3159–3169. [CrossRef] [PubMed]
21. Ringbaek, T.J.; Lange, P. Trends in long-term oxygen therapy for COPD in Denmark from 2001 to 2010. *Respir. Med.* **2014**, *108*, 511–516. [CrossRef]
22. Serginson, J.G.; Yang, I.A.; Armstrong, J.G.; Cooper, D.M.; Matthiesson, A.M.; Morrison, S.C.; Gair, J.M.; Cooper, B.; Zimmerman, P.V. Variability in the rate of prescription and cost of domiciliary oxygen therapy in Australia. *Med. J. Aust.* **2009**, *191*, 549–553.
23. Jones, A.; Wood-Baker, R.; Walters, E.H. Domiciliary oxygen therapy services in Tasmania: Prescription, usage and impact of a specialist clinic. *Med. J. Aust.* **2007**, *186*, 632–634. [PubMed]
24. Wijkstra, P.J.; Guyatt, G.H.; Ambrosino, N.; Celli, B.R.; Güell, R.; Muir, J.F.; Préfaut, C.; Mendes, E.S.; Ferreira, I.; Austin, P.; et al. International approaches to the prescription of long-term oxygen therapy. *Eur. Respir. J.* **2001**, *18*, 909–913. [CrossRef] [PubMed]
25. Sitzia, J.; Wood, N. Patient satisfaction: A review of issues and concepts. *Soc. Sci. Med.* **1997**, *45*, 1829–1843. [CrossRef]
26. Donabedian, A. Evaluating the Quality of Medical Care. *Milbank Q.* **2005**, *83*, 691–729. [CrossRef] [PubMed]
27. Tarlov, A.R.; Ware, J.E., Jr.; Greenfield, S.; Nelson, E.C.; Perrin, E.; Zubkoff, M. The Medical Outcomes Study. An application of methods for monitoring the results of medical care. *JAMA* **1989**, *262*, 925–930. [CrossRef] [PubMed]
28. Hodson, M.; Andrew, S.; Michael Roberts, C. Towards an understanding of PREMS and PROMS in COPD. *Breathe* **2013**, *9*, 358–364. [CrossRef]
29. Mira, J.J.; Nuno-Solinis, R.; Guilabert-Mora, M.; Solas-Gaspar, O.; Fernandez-Cano, P.; Gonzalez-Mestre, M.A.; Contel, J.C.; Del Rio-Camara, M. Development and Validation of an Instrument for Assessing Patient Experience of Chronic Illness Care. *Int. J. Integr. Care* **2016**, *16*, 13. [CrossRef] [PubMed]
30. Hodson, M. Development of a Patient Reported Experience Measure in Chronic Obstructive Pulmonary Disease (COPD). Ph.D. Thesis, University of Portsmouth, Hampshire, UK, 2018.
31. Sjetne, I.S.; Bjertnaes, O.A.; Olsen, R.V.; Iversen, H.H.; Bukholm, G. The Generic Short Patient Experiences Questionnaire (GS-PEQ): Identification of core items from a survey in Norway. *BMC Health Serv. Res.* **2011**, *11*, 88. [CrossRef]
32. Benson, T.; Potts, H.W. A short generic patient experience questionnaire: howRwe development and validation. *BMC Health Serv. Res.* **2014**, *14*, 499. [CrossRef]
33. Fernstrom, K.M.; Shippee, N.D.; Jones, A.L.; Britt, H.R. Development and validation of a new patient experience tool in patients with serious illness. *BMC Palliat. Care* **2016**, *15*, 99. [CrossRef]
34. Schmittdiel, J.; Mosen, D.M.; Glasgow, R.E.; Hibbard, J.; Remmers, C.; Bellows, J. Patient Assessment of Chronic Illness Care (PACIC) and improved patient-centered outcomes for chronic conditions. *J. Gen. Intern. Med.* **2008**, *23*, 77–80. [CrossRef]
35. Coluccia, A.; Ferretti, F.; Pozza, A. Health Services OutPatient Experience questionnaire: Factorial validity and reliability of a patient-centered outcome measure for outpatient settings in Italy. *Patient Relat. Outcome Meas.* **2014**, *5*, 93–103.

36. Beattie, M.; Shepherd, A.; Lauder, W.; Atherton, I.; Cowie, J.; Murphy, D.J. Development and preliminary psychometric properties of the Care Experience Feedback Improvement Tool (CEFIT). *BMJ Open* **2016**, *6*, e010101. [CrossRef]
37. Kleiss, J.A. Preliminary Development of a Multidimensional Semantic Patient Experience Measurement Questionnaire. *Herd* **2016**, *10*, 52–64. [CrossRef]
38. Pettersen, K.I.; Veenstra, M.; Guldvog, B.; Kolstad, A. The Patient Experiences Questionnaire: Development, validity and reliability. *Int. J. Qual. Health Care* **2004**, *16*, 453–463. [CrossRef]
39. Steine, S.; Finset, A.; Laerum, E. A new, brief questionnaire (PEQ) developed in primary health care for measuring patients' experience of interaction, emotion and consultation outcome. *Fam. Pract.* **2001**, *18*, 410–418. [CrossRef]
40. Rodriguez, H.P.; von Glahn, T.; Grembowski, D.E.; Rogers, W.H.; Safran, D.G. Physician effects on racial and ethnic disparities in patients' experiences of primary care. *J. Gen. Intern. Med.* **2008**, *23*, 1666–1672. [CrossRef]
41. Teale, E.A.; Young, J.B. A Patient Reported Experience Measure (PREM) for use by older people in community services. *Age Ageing* **2015**, *44*, 667–672. [CrossRef]
42. Masefield, S.; Vitacca, M.; Dreher, M.; Kampelmacher, M.; Escarrabill, J.; Paneroni, M.; Powell, P.; Ambrosino, N. Attitudes and preferences of home mechanical ventilation users from four European countries: An ERS/ELF survey. *ERJ Open Res.* **2017**, *3*, 00015–02017. [CrossRef]
43. Hendriks, S.H.; Rutgers, J.; van Dijk, P.R.; Groenier, K.H.; Bilo, H.J.G.; Kleefstra, N.; Kocks, J.W.H.; van Hateren, K.J.J.; Blanker, M.H. Validation of the howRu and howRwe questionnaires at the individual patient level. *BMC Health Serv. Res.* **2015**, *15*, 447. [CrossRef]
44. Glasgow, R.E.; Wagner, E.H.; Schaefer, J.; Mahoney, L.D.; Reid, R.J.; Greene, S.M. Development and validation of the Patient Assessment of Chronic Illness Care (PACIC). *Med. Care* **2005**, *43*, 436–444. [CrossRef]
45. Holm, K.E.; Casaburi, R.; Cerreta, S.; Gussin, H.A.; Husbands, J.; Porszasz, J.; Prieto-Centurion, V.; Sandhaus, R.A.; Sullivan, J.L.; Walsh, L.J.; et al. Patient Involvement in the Design of a Patient-Centered Clinical Trial to Promote Adherence to Supplemental Oxygen Therapy in COPD. *Patient* **2016**, *9*, 271–279. [CrossRef]
46. Clèries, X.; Solà, M.; Chiner, E.; Escarrabill, J. Aproximación a la experiencia del paciente y sus cuidadores en la oxigenoterapia domiciliaria. *Arch. Bronconeumol.* **2016**, *52*, 131–137. (In Spanish) [CrossRef] [PubMed]
47. Luyster, F.S.; Dunbar-Jacob, J.; Aloia, M.S.; Martire, L.M.; Buysse, D.J.; Strollo, P.J. Patient and partner experiences with obstructive sleep apnea and CPAP treatment: A qualitative analysis. *Behav. Sleep Med.* **2016**, *14*, 67–84. [CrossRef] [PubMed]
48. Ward, K.; Gott, M.; Hoare, K. Making choices about CPAP: Findings from a grounded theory study about living with CPAP. *Collegian* **2017**, *24*, 371–379. [CrossRef]
49. Gibson, R.; Campbell, A.; Mather, S.; Neill, A. From diagnosis to long-term treatment: The experiences of older New Zealanders with obstructive sleep apnoea. *J. Prim. Health Care.* **2018**, *2*, 140–149. [CrossRef] [PubMed]
50. Gale, N.K.; Jawad, M.; Dave, C.; Turner, A.M. Adapting to domiciliary non-invasive ventilation in chronic obstructive pulmonary disease: A qualitative interview study. *Palliat. Med.* **2015**, *29*, 268–277. [CrossRef] [PubMed]
51. Brostrom, A.; Nilsen, P.; Johansson, P.; Ulander, M.; Stromberg, A.; Svanborg, E.; Fridlund, B. Putative facilitators and barriers for adherence to CPAP treatment in patients with obstructive sleep apnea syndrome: A qualitative content analysis. *Sleep Med.* **2010**, *11*, 126–130. [CrossRef]
52. Gardener, A.C.; Ewing, G.; Kuhn, I.; Farquhar, M. Support needs of patients with COPD: A systematic literature search and narrative review. *Int. J. Chron. Obstruct. Pulmon. Dis.* **2018**, *13*, 1021–1035. [CrossRef] [PubMed]
53. Katsenos, S.; Constantopoulos, S.H. Long-Term Oxygen Therapy in COPD: Factors Affecting and Ways of Improving Patient Compliance. *Pulm. Med.* **2011**, *2011*, 325362. [CrossRef]
54. Fujisawa, R.; Klazinga, N.S. Measuring patient experiences (PREMS): Progress made by the OECD and its member countries between 2006 and 2016. *OECD Health Work. Pap.* **2017**. [CrossRef]

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
Tel. +41 61 683 77 34
Fax +41 61 302 89 18
www.mdpi.com

Journal of Clinical Medicine Editorial Office
E-mail: jcm@mdpi.com
www.mdpi.com/journal/jcm

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland

Tel: +41 61 683 77 34
Fax: +41 61 302 89 18

www.mdpi.com

ISBN 978-3-03936-863-1