Journal of
Clinical Medicine

Diagnosis and Treatment of Musculoskeletal Disorders

Edited by
Jos Runhaar and Marienke van Middelkoop

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

www.mdpi.com/journal/jcm

MDPI

Diagnosis and Treatment of Musculoskeletal Disorders

Diagnosis and Treatment of Musculoskeletal Disorders

Editors

Jos Runhaar
Marienke van Middelkoop

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

Editors
Jos Runhaar
Department of General Practice,
Erasmus MC University Medical Center
Rotterdam
The Netherlands

Marienke van Middelkoop
Department of General Practice,
Erasmus MC University Medical Center
Rotterdam
The Netherlands

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/musculoskeletal_disorders).

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

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

ISBN 978-3-0365-2660-7 (Hbk)
ISBN 978-3-0365-2661-4 (PDF)

Contents

About the Editors

Jos Runhaar works as an associate professor at the Department of General Practice of Erasmus MC. His research focusses on the early diagnosis and treatment of musculoskeletal disorders in primary care. Herewith, his research group aims to improve the diagnosis of musculoskeletal disorders by general practitioners and physiotherapists and to shift the treatment of musculoskeletal disorders to the early disease phase; even to a phase where primary prevention becomes possible.

Marienke van Middelkoop is an associate professor at the department of General Practice. Her research has a strong focus on musculoskeletal complaints in children, adolescents and young adults in primary care. Her research activities have, since 2005, led to over 170 international peer-reviewed publications. She is the coordinator of the Osteoarthritis (OA) Trial Bank, an international initiative of meta-analyses of effect of treatment on predefined subgroups of OA patients from existing studies and she is a deputy editor of the *British Journal of Sports Medicine*.

Journal of
Clinical Medicine

MDPI

Article

What Is the Diagnosis in Patients with Type 2 Diabetes Who Have a Painful Shoulder? Results from a Prospective Cross-Sectional Study

Login Ahmed S. Alabdali [1,2,*], Jasmien Jaeken [3], Nens van Alfen [4], Geert-Jan Dinant [1], Rob A. P. Borghans [5] and Ramon P. G. Ottenheijm [1]

1 Department of Family Medicine, CAPHRI Care and Public Health Research Institute, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands; geertjan.dinant@maastrichtuniversity.nl (G.-J.D.); ramon.ottenheijm@maastrichtuniversity.nl (R.P.G.O.)
2 Ministry of Education, Riyadh 12435, Saudi Arabia
3 Department of Public Health and Primary Care, Catholic University of Leuven, Kapucijnenvoer 33, B-3000 Leuven, Belgium; j.jaeken@alumni.maastrichtuniversity.nl
4 Department of Neurology and Clinical Neurophysiology, Radboud University Medical Center, Donders Institute for Brain Cognition and Behaviour, 6500 HB Nijmegen, The Netherlands; Nens.vanAlfen@radboudumc.nl
5 Department of Radiology, Zuyderland Medical Centre, 6162 BG Sittard-Geleen, The Netherlands; R.borghans@zuyderland.nl
* Correspondence: login.alabdali@maastrichtuniversity.nl

Received: 24 November 2020; Accepted: 15 December 2020; Published: 18 December 2020

Abstract: Background: Patients with diabetes mellitus have higher risk of developing shoulder pathology. However, only adhesive capsulitis is addressed in shoulder pain guidelines as a disorder associated with diabetes. Yet, patients with diabetes are at risk of having several other shoulder disorders, including focal neuropathy. Our aim was to quantify the presence of shoulder disorders using physical examination and ultrasound imaging in patients with type 2 diabetes (T2DM) suffering from shoulder pain in general practice. Methods: In this prospective cross-sectional study, patients with T2DM who had had a painful shoulder for at least four weeks were included. Patients filled out a questionnaire and underwent a physical examination of the shoulders and feet and ultrasound imaging of the shoulder. Results: A total of 66 patients were included, of whom 40.9% (n = 27) had bilateral complaints resulting in 93 symptomatic shoulders. Subacromial pain syndrome was most frequently diagnosed by physical examination (66.6%, 95% CI 51.6–72.0%; $p < 0.0001$), while ultrasound imaging showed that subacromial disorders were statistically significantly the most prevalent (90.3%, 95% CI 81.9–95.2%). Only two patients (3%) were diagnosed with neuropathic shoulder pain. Conclusion: When choosing treatment, general practitioners should be aware that in patients with T2DM the subacromial region is most frequently affected.

Keywords: type 2 diabetes; physical examination; diagnosis; ultrasound; shoulder pain; adhesive capsulitis; subacromial pain syndrome

1. Introduction

Patients with type 2 diabetes mellitus (T2DM) are at increased risk for shoulder disorders, such as adhesive capsulitis and rotator cuff disorders (e.g., tears). The prevalence of adhesive capsulitis ranges from 4 to 30% in patients with diabetes, compared to up to 10% in patients without diabetes [1–4], while rotator cuff tears are also more frequently observed in patients with diabetes [5–8]. Lesser known is that adhesive capsulitis is also associated with diabetic neuropathy [9].

The exact pathophysiology of the increased risk of shoulder disorders in patients with diabetes remains uncertain, but there is evidence that the shoulder can be affected through two pathophysiological pathways: connective tissue damage to the rotator cuff tendons and joint capsule, and peripheral or autonomic neuropathy. Connective tissue damage seems to be caused by abnormal collagen disposition in the periarticular connective tissues, due to the formation of advanced glycosylation end-products. This alters the structural matrix and the mechanical properties of these tissues. Additionally, it is hypothesized that the altered glucose metabolism in patients with diabetes causes functional as well as structural changes to the peripheral nerve system, which ultimately leads to neuropathy [5,10]. About 50% of patients with diabetes mellitus develop diabetic neuropathy [11,12], which can be classified as generalized symmetrical polyneuropathy or (multi)focal asymmetrical neuropathy [11,13]. If patients with subclinical levels of neuropathy are included, the prevalence might exceed 90% [14,15]. However, in patients with diabetes, shoulder involvement has been described in case studies, but is currently regarded as rare [16–18].

While rotator cuff disorders (also named subacromial pain syndrome, abbreviated to SAPS [19,20]) are widely considered the most common cause of shoulder pain in general practice [21–23], diabetes mellitus is currently only associated with adhesive capsulitis in shoulder pain guidelines, while neuropathy is not addressed at all [20,24,25]. However, to ensure optimal treatment outcomes, it is important to identify the underlying disorder in patients with diabetes to prevent the development of chronic shoulder pain. For example, patients with SAPS and neuropathic pain might benefit more from treatment targeted to the neuropathic pain [26]. This seems important especially in patients with T2DM, as shoulder pain might negatively influence physical activity, which is considered to be the cornerstone of diabetes treatment. Therefore, inadequate treatment of shoulder pain might negatively influence diabetes treatment, eventually leading to a vicious circle [27,28].

To our knowledge, no studies have been conducted to establish the frequency of underlying shoulder disorders in patients with T2DM suffering from shoulder pain. In order to prevent the development of chronic shoulder pain in patients with diabetes, more insight is needed into which specific types of shoulder pathology are prevalent in T2DM. Given the complex nature of shoulder pain, the primary aim of this study was to quantify the presence of specific shoulder disorders using physical examination and musculoskeletal ultrasound imaging in patients with type 2 diabetes mellitus (T2DM) in general practice. Secondarily, we sought to examine the relationship of the specific disorders with the presence of a diabetic neuropathy. With this information, we can increase general practitioners (GP) awareness of this specific problem in T2DM.

2. Methods

2.1. Participants and Study Design

This was a prospective cross-sectional study of patients with T2DM suffering from shoulder pain in general practice. Patients were invited to this study while participating in a questionnaire study assessing the prevalence of upper extremity musculoskeletal disorders (manuscript under review). During their annual check-ups in general practice, they had been asked to participate in a questionnaire study, which also included an invitation for the current study that was only intended for patients who had shoulder pain. All eligible participants from the questionnaire study were invited to participate. To increase recruitment, a flyer announcing this study was sent out to physiotherapy and general practices in the Sittard-Geleen area in the Netherlands. During our inclusion period, for which we had a time frame of nine months (March 2018 to November 2018), we included all eligible patients. Inclusion criteria were being aged between 30 and 70 years, and having shoulder pain that had lasted longer than 4 weeks. Exclusion criteria were inability to complete the assessments, and inability to give informed consent. After providing informed consent, patients underwent the following assessments: a questionnaire addressing shoulder pain including co-morbidity and diabetes-related questions, standardized physical examination of the shoulder, neurological examination of the feet, and ultrasound

imaging of the shoulder (see below). All examinations were carried out in the Meditta Medical Center in Echt, a diagnostic center, between March 2018 and November 2018. This study was approved by the Medical Ethics Committee of Zuyderland Medical Centre (METC-Z 17-N-165).

2.2. Questionnaire

The following demographic and shoulder-specific information was collected: an 11-point Numerical Rating Scale, affected side and dominant arm, pain onset (sudden or gradual), neck pain involvement, and if there was a history of rheumatoid arthritis, osteoarthritis, and any other joint inflammation. Additionally, we collected information about the most recent HBA1c value, the current body mass index (BMI), and year of their diabetes diagnosis.

2.3. Ultrasound Imaging

Ultrasound examination of the shoulder was the first examination participants underwent, and was performed by one of the three participating, experienced radiologists, using a Phillips EPIQ 7G (Philips Healthcare, Eindhoven, the Netherlands) with a high-resolution, multi-frequency 4–18 MHz linear transducer. A standardized protocol was performed following the technical guidelines of the European Society of Musculoskeletal Radiology for shoulder scanning [29]. In patients with diabetes, both shoulders were assessed. The following structures were assessed: the tendon of the supraspinatus, infraspinatus, subscapularis, and the long head of biceps, as well as the subacromial-subdeltoid bursa, rotator interval, glenohumeral recess, and acromioclavicular joint (AC joint). We used standardized criteria for subacromial pathology previously defined by us, and added criteria for adhesive capsulitis (see Supplement Table S1) [30–32]. The following disorders were recorded: dislocation of the long head of biceps tendon, biceps tenosynovitis, (calcific) tendinopathy, any kind of tendon tear, bursitis, adhesive capsulitis, paralabral cyst, and acromioclavicular osteoarthritis. Next, they were categorized by their anatomical location as subacromial, glenohumeral, or other disorders, by following a detailed standard operating procedure (see Supplement Table S1 for full details).

2.4. Physical Examination

One trained medical doctor of our study team (LA), blinded to the ultrasound imaging results, carried out a structured physical examination of both shoulders, hands, and a neurological examination of the feet.

Shoulders were examined according to the shoulder pain guidelines of the Dutch College of General Practitioners prevailing at the time of conducting the study [33]. One year later (2019), these guidelines were revised. Therefore, we used the revised version for subsequent classification of the shoulder disorders [20]. Shoulders were inspected for muscle wasting and scapular winging. Active and passive abduction and passive external rotation were carried out to assess for pain and range of motion was visually estimated. Additionally, the Hawkins–Kennedy test and the Neer impingement test were carried out [34,35]. The 10-item Douleur Neuropathique questionnaire was performed to assess neuropathic shoulder pain; a commonly used questionnaire including physical examination for screening and diagnosing neuropathic pain in patients with neurological complaints, valid for the Dutch population, and validated in diabetes patients [36–38]. In the Dutch population, the cut-off score for neuropathic pain is considered ≥5 (out of 10 points) [39]. Physical examination findings were used to diagnose the following disorders: SAPS, a glenohumeral or an "other disorder", and categorized using a mutually exclusive method leading to a single diagnosis (Figure 1) [20]. Glenohumeral disorders were defined by an external rotation range of motion of less than 45 degrees, while SAPS was defined by either a painful abduction (including a painful arc) with or without a limited range of motion during abduction, or a positive Hawkins–Kennedy and Neer test. An "other disorder" was defined as not having any of the two previous disorders.

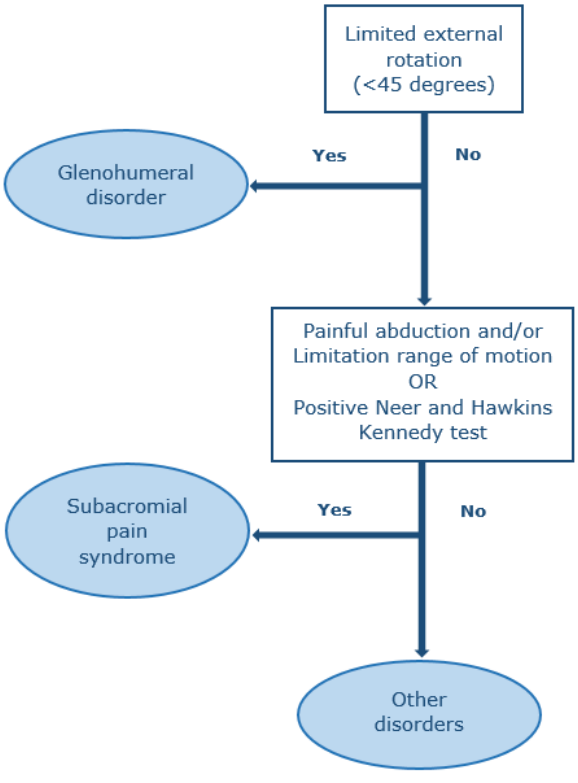

Figure 1. Diagnosis based on physical examination by a mutually excluding method.

Examination of the feet consisted of two parts: an interview part to evaluate neuropathic symptoms, and a neurological examination to evaluate signs. Participants were asked if they had a burning, tingling, or numbness sensation, and whether they felt like they were walking on cotton wool. The neurological examination included inspection for distal muscle weakness or atrophy, ankle reflexes, touch sensation, and vibration sense. The diagnosis of clinical polyneuropathy was based on a combination of symptoms and signs: presence of at least one symptom in combination with decreased or absent ankle reflexes, and decreased or absent distal sensation (vibration sense and/or touch sensation) with or without distal muscle weakness. For subclinical polyneuropathy, the abovementioned signs had to be present in the absence of symptoms. All examinations including criteria for pathology are presented in the Supplement (see Supplement Table S2) [40–46]. Finally, the hands were examined for the presence of a positive tabletop sign or prayer sign as a manifestation of stiff hand syndrome (see Supplement Table S3) [47–49].

2.5. Statistical Analysis

Continuous variables were tested for normal distribution. Descriptive statistics are presented for continuous variables as means with standard deviations, and categorical variables data are presented as absolute frequencies and percentages. Statistical differences for physical-examination-diagnosed shoulder disorders were calculated using the Chi-square test to compare between the three diagnostic groups. All statistical analysis were performed using IBM SPSS Statistics for Windows (version 25.0, Armonk, NY, USA).

3. Results

Table 1 presents the demographic and shoulder pain characteristics, as well as the results of the neurological and hand examination. A total of 66 patients with T2DM suffering from shoulder pain could be included during the study period. In total, 39 were recruited via the questionnaire study, and 27 via the flyers. The mean age was 63.0 years and 28.8% ($n = 19$) was female. Bilateral complaints were present in 41% of the patients ($n = 27$), resulting in 93 symptomatic shoulders and 39 asymptomatic shoulders.

Table 1. Physical-examination-diagnosed shoulder disorders with the comparison of the demographic and clinical variables in patients with type 2 diabetes mellitus (T2DM), on both patient and shoulder level.

	Total Number of Patients with Shoulder Pain $n = 66$	**Symptomatic Shoulders, $n = 93$**		
		GH	**SAPS**	**Other Disorder**
		$n = 17$ (18.2%, 95% CI: 11.3–27.9) [#]	**$n = 58$ (66.6%, 95% CI: 51.6–72.0) [#]**	**$n = 18$ (16.1%, 95% CI: 12.1–29.1) [#]**
Age mean ± SD IQR range (years)	 63.0 ± 6.9 38–70	 58.8 ± 5.8 48–70	 61.5 ± 7.0 38–70	 61.5 ± 7.6 52–70
Female sex	19 (28.8)	5 (31.3)	11 (25.0)	3 (42.9)
BMI (kg/m^2) mean ± SD range	 28.5 ± 4.2 16–41.5	 28.1 ± 3.7 23.9–35.9	 28.4 ± 4.5 16–41.5	 29.7 ± 3.5 25.4–34.0
Bilateral shoulder pain	27 (40.9)	8 (50)	15 (34.1)	4 (57.1)
Dominant shoulder affected	26 (39.3)	13 (81.3)	36 (81.8)	7 (100)
NRS mean ± SD range	 5.3 ± 2.1 1–9	 6.25 ± 1.9 3–9	 5.2 ± 2.06 1–8	 4.0 ± 2.6 1–8
Pain onset sudden gradual	 13 (19.7) 51 (79.7)	 2 (11.7) 15 (88.3)	 13 (20.9) 44 (70.9)	 2 [a] (11.1) 10 (12.5)
Neck pain	9 (13.6)	2 (12.5)	5 (11.4)	2 (11.1)
Duration of diabetes mean ± SD range (years)	 9.0 ± 5.5 1–27	 8.2 ± 4.3 [a] 1–14	 9.4 ± 5.8 [b] 1–27	 8.2 ± 5.8 2–19
HbA1C (mmol/mol) mean ± SD range	 55.7 ± 7.7 43–75	 55.6 ± 6.2 [a] 44–64	 55.8 ± 8.03 [b] 43–75	 53.6 ± 10.1 [c] 43–70
Rheumatoid arthritis [a]	8 (12.5)	3 (18.8)	4 [a] (10.3)	1 (14.3)
Osteoarthritis [a]	33 (51.6)	9 (56.3)	22 [a] (52.4)	3 (33.3)
Stiff hand syndrome (%)	28 (42.4)	9 (56.3)	17 (38.6)	2 (28.6)
Neuropathic shoulder pain by DN4 (score ≥ 5)	2 (3.0)	0	1 (1.6)	1 (5.2)
Polyneuropathy Clinical Subclinical	 19 (28.8) 37 (56.1)	 5 (29.4) 11 (64.7)	 11 (19.2) 22 (38.5)	 3 (15.7) 4 (21.1)

Values are presented as absolute numbers and percentages unless otherwise stated; SAPS: subacromial pain syndrome; GH: glenohumeral disorder; IQR: interquartile range; BMI: body mass index; SD: standard deviation; NRS: numerical rating scale; DN4: Douleur Neuropathique 4 questionnaire; [a] 2 are missing; [b] 4 are missing; [c] 1 is missing. [#] Statistical comparison between GH vs. SAPS: p-value < 0.0001, SAPS vs. Other disorder: p-value < 0.0001, GH vs. Other disorder: p-value 0.072.

3.1. Physical-Examination-Diagnosed Symptomatic Shoulder Disorders

The results of the shoulder diagnosis based on physical examination are presented in Table 1. SAPS was statistically the most frequently observed disorder in symptomatic shoulders (66.6%, 95% CI 51.6–72.0%), followed by glenohumeral disorders (18.2%, 95% CI 11.3–27.9%; $p < 0.0001$ versus SAPS), and then "other disorder" (16.1%, 95% CI 12.1–29.1%; $p < 0.0001$ versus SAPS). Neuropathic shoulder

pain was observed in two patients (3.2%); one patient was diagnosed with SAPS, the other with an "other disorder".

3.2. Polyneuropathy

Clinical polyneuropathy of the feet was observed in 28.8% ($n = 19$), while subclinical polyneuropathy was present in 56.1% ($n = 37$). In patients with a glenohumeral disorder, polyneuropathy was most commonly observed (29.4%, $n = 5$).

3.3. Ultrasound-Diagnosed Shoulder Disorders

Table 2 presents the results of the ultrasound findings in both symptomatic and asymptomatic shoulders. Subacromial disorders were most frequently diagnosed in both symptomatic and asymptomatic shoulders (90.3% ($n = 84$) and 76.9% ($n = 30$), respectively) with rotator cuff disorders being most prevalent. This was followed by osteoarthritis of the AC joint (59.1% ($n = 55$) and 43.5% ($n = 17$), respectively) and then glenohumeral disorders (8.6%, $n = 8$) in symptomatic shoulders, while these were not observed in asymptomatic shoulders. Of the rotator cuff disorders, the most frequently observed specific disorder was calcific tendinopathy in both symptomatic and asymptomatic shoulders (80.6% ($n = 75$), 95% CI 70.8–87.8% and 66.7% ($n = 26$), 95%CI 49.7–80.4%, respectively), while rotator cuff tears were least commonly observed (38.7% ($n = 36$), 95% CI 28.9–49.4% and 20.5% ($n = 8$), 95% CI 9.8–36.9%, respectively) (these results are not shown in Table 2). Of the eight patients with a glenohumeral disorder, adhesive capsulitis was present in four cases.

When we matched the results of the physical examination and ultrasound diagnosis, in 97% (56/58) of the patients diagnosed with SAPS, a subacromial disorder was observed by ultrasound imaging. Only two out of 17 patients (12%) diagnosed with a glenohumeral disorder showed ultrasound findings matching this diagnosis.

Table 2. Ultrasound-diagnosed disorders in symptomatic and asymptomatic shoulders in patients with T2DM ($n = 132$).

Ultrasound-Diagnosed Disorders	All 66 Patients with 132 Shoulders					
	Symptomatic Shoulders ($n = 93$)			Asymptomatic Shoulders ($n = 39$)		
	n	%	95% CI	*n*	%	95% CI
Subacromial pain disorders	84	90.3	81.9–95.2	30	76.9	60.2–88.2
Subacromial bursitis	13	14.0	7.9–23.1	3	7.7	2.0–21.9
Rotator cuff disorder	84	90.3	81.9–95.2	30	76.9	60.2–88.2
LHBT disorder	10	10.7	5.5–19.3	3	7.6	2.01–21.9
Dynamic impingement	14	15.1	8.7–24.3	3	7.6	2.0–21.9
Glenohumeral disorders	8	8.6	4.1–16.7	0	0	0
Adhesive capsulitis	4	4.3	1.3–11.2	0	0	0
GH effusion only	4	4.3	1.3–11.2	0	0	0
Other disorders						
Acromioclavicular OA	55	59.1	48.4–69.1	17	43.5	28.1–60.2
No disorders	0	0	0	0	0	0

LHBT: Long head of biceps tendon; OA: osteoarthritis; GH: glenohumeral.

4. Discussion

In this study, using physical examination and ultrasound imaging, we found that the subacromial region is the most frequently affected region in patients with T2DM suffering from shoulder pain. Of these subacromial structures, ultrasound imaging shows that the rotator cuff tendons are most frequently affected, with calcific tendinopathy by far the most common specific disorder. Interestingly,

we found that adhesive capsulitis, believed to be frequently diagnosed in patients with T2DM suffering from shoulder pain, was present only in a minority of patients. Moreover, neuropathic shoulder pain can be considered rare as it was present in only two patients.

4.1. Diagnosis by Physical Examination and Ultrasound Imaging

Currently, shoulder pain guidelines only associate diabetes with adhesive capsulitis [20,24,25], while we showed that subacromial disorders are by far the most common in this population. In the absence of prevalence studies, a possible explanation for the current misperception might be that adhesive capsulitis is more prevalent among patients with T2DM compared with patients without diabetes [1–4], but this does not mean that it is the most prevalent shoulder disorder in patients with diabetes.

In line with previous studies conducted in patients with shoulder pain in unselected populations [21–23], we also observed that overall, SAPS is the most common cause of shoulder pain in diabetics. The rotator cuff tendons and subacromial-subdeltoid bursa are considered to be the key sources of pathology in this syndrome [50].

SAPS is a generic term, and incorporates all disorders related to subacromial structures including tendinopathy (tendon degeneration), where calcific tendinopathy is seen as a separate diagnosis, tendon tears, and bursitis. From this spectrum of disorders, calcific tendinopathy was the most common specific disorder in our cohort (81%), which echoes the results of previous studies conducted in general practice [51–53]. However, in these previous ultrasound studies, the frequency did not exceed 50%, but it is known that diabetes is associated with calcific tendinopathy [54]. Additionally, other rotator cuff tendon disorders were more frequently present in our study. These differences might be explained in pathophysiological terms, and although the exact pathophysiology of tendon disorders in patients with diabetes remains uncertain, there is evidence that abnormal tendon collagen disposition alters the structural matrix and the mechanical properties of the tendons [6]. Through this process, the continuum of tendon pathology might be initiated, where a normal tendon changes into a degenerative tendon (called tendinopathy, different stages are described), and ultimately can tear [55]. Part of this process seems reversible through healing responses, but in patients with diabetes and other endocrine disorders, this healing process can fail; calcium deposits then arise due to a mechanism not yet elucidated [54].

In our study, the large majority of our patients clinically diagnosed with SAPS showed ultrasound findings matching SAPS. In clinical practice, ultrasound findings should always be interpreted together with the clinical findings to determine the cause of symptoms. Several shoulder imaging studies have shown that asymptomatic pathology findings are present, for example supraspinatus tears and osteoarthritis of the AC joint [56,57], and these asymptomatic findings are likely to become symptomatic over time [57]. Surprisingly, we also found ultrasound abnormalities of subacromial pathology in 16 out of 17 patients with a clinical glenohumeral disorder, while only 2 out of these 17 patients had ultrasound evidence of glenohumeral pathology. It might be that ultrasound misses the presence of osteoarthritis of the glenohumeral joint, something we did not assess radiographically in this study. The diagnostic accuracy of ultrasound imaging for osteoarthritis of the glenohumeral joint is unknown, but is likely to be suboptimal.

We also observed that 41% of the patients has bilateral complaints. This is in line with previous studies [58,59], and might be explained in the light of the above and systemic effects of diabetes.

4.2. Neuropathy

Although neuropathy is a well-known complication of diabetes, any type of neuropathy of the shoulder is currently regarded as rare [16–18]. In line with this observation, we diagnosed only two patients with neuropathic shoulder pain: both these patients were diagnosed with clinical polyneuropathy, but did not have a glenohumeral disorder on physical examination, or adhesive capsulitis on ultrasound imaging. Overall, polyneuropathy on a clinical and subclinical level was present in 39% and 56% of the patients, respectively. This is broadly consistent with a study showing

that 38% of patients with T2DM and musculoskeletal disorders had polyneuropathy [9]. We do know that diabetic neuropathy increases the risk of developing adhesive capsulitis of the shoulder, a glenohumeral disorder [9]. In our study, any form of polyneuropathy was most prevalent in patients diagnosed with a glenohumeral disorder.

4.3. Strengths and Limitations

Our study had several strengths. We were able to investigate the frequency of the shoulder-diagnosed disorders by two measurements: physical examination and ultrasound imaging. Both indicate that the subacromial region is the most frequently affected structure. The medical doctor who performed the physical examination was blinded to the ultrasound results, which avoided bias for the results of the physical examination.

Our study also had limitations. First, the sample size is small. A larger sample size would have resulted in narrower 95% confidence intervals. Second, we did not include palpation of the AC joint in our protocol, because it was not included in the shoulder pain guidelines of the Dutch College of General Practitioners prevailing at the time of conducting this study. If we had incorporated this in our study, this would also have allowed us to diagnose AC disorders, usually osteoarthritis. In our approach, this disorder is included in the "other disorder" group [33]. It is worth noting that palpation of the AC joint is included in the revised version [20]. Third, neuropathic shoulder pain was diagnosed only by physical examination, while if we had used additional tests, e.g., nerve conducting studies or muscle ultrasound (qualitative or quantitative), it might have increased the number of patients diagnosed with neuropathic shoulder pain. Fourth, the ultrasound diagnostic criteria for glenohumeral disorders such as adhesive capsulitis are debatable as more features are described in the literature, e.g., thickening of the coracohumeral ligament and restriction of external rotation on dynamic scanning [31]. These features are not detected during scanning according to the standardized protocol of the European Society of Musculoskeletal Radiology. For practical purposes, we did not incorporate additional scanning positions necessary to detect these features. This may have introduced an underestimation of the presence of adhesive capsulitis.

4.4. Implications for Practice and Future Research

General practitioners and other healthcare professional involved in the care for patients with shoulder pain should be aware that in patients with T2DM suffering from shoulder pain, the subacromial region is the most frequently affected structure. This knowledge can help make a more accurate diagnosis and can influence treatment decisions. When physical examination is not conclusive enough to establish a diagnosis, GPs may consider ultrasound imaging. This seems to be becoming more usual, as GPs increasingly rely on ultrasound imaging [60,61]. Although we showed that polyneuropathy is frequently present, neuropathic shoulder pain seems rare. However, this finding is based only on physical examination results of the shoulder. To assess whether shoulder pain in T2DM might be caused by neuropathy, studies are needed that use techniques to detect neurological denervation, e.g., nerve conducting studies of qualitative or quantitative muscle ultrasound.

Supplementary Materials: The following are available online at http://www.mdpi.com/2077-0383/9/12/4097/s1, Table S1: Criteria for ultrasound-diagnosed disorders groups, Table S2: Neurological feet examination procedures with diagnostic criteria, Table S3: Diagnostic criteria for stiff hand syndrome.

Author Contributions: Conceptualization, R.P.G.O., G.-J.D., L.A.S.A., and N.v.A.; Methodology, R.P.G.O., G.-J.D., L.A.S.A., N.v.A., and R.A.P.B.; Formal Analysis, L.A.S.A. and J.J.; Clinical investigation, L.A.S.A.; Data curation, L.A.S.A.; Writing the original draft preparation, L.A.S.A.; Writing—review and editing, R.P.G.O., G.-J.D., L.A.S.A., N.v.A., and R.A.P.B.; supervision, R.P.G.O. and G.-J.D.; All authors have read and agreed to the published version of the manuscript.

Funding: L.A.S.A. is funded by the Ministry of Education of the Kingdom of Saudi Arabia.

Acknowledgments: Funding was completely provided by the Ministry of Education of the Kingdom of Saudi Arabia. This funding body was not involved in the design of the study, or writing the manuscript.

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

References

1. Smith, L.L.; Burnet, S.P.; McNeil, J.D. Musculoskeletal manifestations of diabetes mellitus. *Br. J. Sports Med.* **2003**, *37*, 30–35. [CrossRef] [PubMed]
2. Cagliero, E.; Apruzzese, W.; Perlmutter, G.S.; Nathan, D.M. Musculoskeletal disorders of the hand and shoulder in patients with diabetes mellitus. *Am. J. Med.* **2002**, *112*, 487–490. [CrossRef]
3. Thomas, S.J.; McDougall, C.; Brown, I.D.; Jaberoo, M.C.; Stearns, A.; Ashraf, R.; Fisher, M.; Kelly, I.G. Prevalence of symptoms and signs of shoulder problems in people with diabetes mellitus. *J. Shoulder Elb. Surg.* **2007**, *16*, 748–751. [CrossRef] [PubMed]
4. Kidwai, S.S.; Wahid, L.; Siddiqi, S.A.; Khan, R.M.; Ghauri, I.; Sheikh, I. Upper limb musculoskeletal abnormalities in type 2 diabetic patients in low socioeconomic strata in Pakistan. *BMC Res. Notes* **2013**, *6*, 16. [CrossRef]
5. Abate, M.; Schiavone, C.; Salini, V.; Andia, I. Management of limited joint mobility in diabetic patients. *Diabetes Metab. Syndr. Obes.* **2013**, *6*, 197–207. [CrossRef]
6. Lebiedz-Odrobina, D.; Kay, J. Rheumatic manifestations of diabetes mellitus. *Rheum. Dis. Clin. N. Am.* **2010**, *36*, 681–699. [CrossRef]
7. Bhat, T.A.; Dhar, S.A.; Dar, T.A.; Naikoo, M.A.; Naqqash, M.A.; Bhat, A.; Butt, M.F. The Musculoskeletal Manifestations of Type 2 Diabetes Mellitus in a Kashmiri Population. *Int. J. Health Sci. (Qassim)* **2016**, *10*, 57–68. [CrossRef]
8. Abate, M.; Schiavone, C.; Salini, V. Sonographic evaluation of the shoulder in asymptomatic elderly subjects with diabetes. *BMC Musculoskelet. Disord.* **2010**, *11*, 278. [CrossRef]
9. Abourazzak, F.E.; Akasbi, N.; Houssaini, G.S.; Bazouti, S.; Bensbaa, S.; Hachimi, H.; Ajdi, F.; Harzy, T. Articular and abarticular manifestations in type 2 diabetes mellitus. *Eur. J. Rheumatol.* **2014**, *1*, 132–134. [CrossRef]
10. Arkkila, P.E.; Gautier, J.F. Musculoskeletal disorders in diabetes mellitus: An update. *Best Pract. Res. Clin. Rheumatol.* **2003**, *17*, 945–970. [CrossRef]
11. Boulton, A.J. Diabetic neuropathy: Classification, measurement and treatment. *Curr. Opin. Endocrinol. Diabetes Obes.* **2007**, *14*, 141–145. [CrossRef] [PubMed]
12. Dyck, P.J.; Overland, C.J.; Low, P.A.; Litchy, W.J.; Davies, J.L.; Dyck, P.J.; O'Brien, P.C.; Albers, J.W.; Andersen, H.; Bolton, C.F.; et al. Signs and symptoms versus nerve conduction studies to diagnose diabetic sensorimotor polyneuropathy: Cl vs. NPhys trial. *Muscle Nerve* **2010**, *42*, 157–164. [CrossRef] [PubMed]
13. Tesfaye, S.; Boulton, A.J.; Dyck, P.J.; Freeman, R.; Horowitz, M.; Kempler, P.; Lauria, G.; Malik, R.A.; Spallone, V.; Vinik, A.; et al. Diabetic neuropathies: Update on definitions, diagnostic criteria, estimation of severity, and treatments. *Diabetes Care* **2010**, *33*, 2285–2293. [CrossRef] [PubMed]
14. Association, A.D. Standards of Medical Care in Diabetesd2017. *J. Clin. Appl. Res. Educ.* **2017**, *40*, 142.
15. Yagihashi, S.; Mizukami, H.; Sugimoto, K. Mechanism of diabetic neuropathy: Where are we now and where to go? *J. Diabetes Investig.* **2011**, *2*, 18–32. [CrossRef] [PubMed]
16. Pica, E.C.; Verma, K.K. Bilateral brachial plexopathy as an initial presentation in a newly-diagnosed, uncontrolled case of diabetes mellitus. *Singap. Med. J.* **2008**, *49*, e29–e32.
17. Ogawa, K.; Sasaki, H.; Kishi, Y.; Yamasaki, H.; Okamoto, K.; Yamamoto, N.; Hanabusa, T.; Nakao, T.; Nishi, M.; Nanjo, K. A suspected case of proximal diabetic neuropathy predominantly presenting with scapulohumeral muscle weakness and deep aching pain. *Diabetes Res. Clin. Pract.* **2001**, *54*, 57–64. [CrossRef]
18. Wada, Y.; Yanagihara, C.; Nishimura, Y.; Oka, N. A case of diabetic amyotrophy with severe atrophy and weakness of shoulder girdle muscles showing good response to intravenous immune globulin. *Diabetes Res. Clin. Pract.* **2007**, *75*, 107–110. [CrossRef]
19. Diercks, R.; Bron, C.; Dorrestijn, O.; Meskers, C.; Naber, R.; de Ruiter, T.; Willems, J.; Winters, J.; van der Woude, H.J. Guideline for diagnosis and treatment of subacromial pain syndrome: A multidisciplinary review by the Dutch Orthopaedic Association. *Acta Orthop.* **2014**, *85*, 314–322. [CrossRef]
20. Nederlands Huisartsen Genootschap. NHG-Standaard Schouderklachten (Dutch College of General Practitioners Guidelines of Shoulder Complaints). Available online: https://www.nhg.org/standaarden/samenvatting/schouderklachten (accessed on 20 March 2020).

21. Littlewood, C.; May, S.; Walters, S. Epidemiology of Rotator Cuff Tendinopathy: A Systematic Review. *Shoulder Elb.* **2013**, *5*, 256–265. [CrossRef]
22. Littlewood, C.; Bateman, M.; Connor, C.; Gibson, J.; Horsley, I.; Jaggi, A.; Jones, V.; Meakins, A.; Scott, M. Physiotherapists' recommendations for examination and treatment of rotator cuff related shoulder pain: A consensus exercise. *Physiother. Pract. Res.* **2019**, *40*, 87–94. [CrossRef]
23. Van der Windt, D.A.; Koes, B.W.; de Jong, B.A.; Bouter, L.M. Shoulder disorders in general practice: Incidence, patient characteristics, and management. *Ann. Rheum. Dis.* **1995**, *54*, 959–964. [CrossRef]
24. CKS. Shoulder Pain Guideline. Available online: http://www.cks.nhs.uk/shoulder_pain (accessed on 20 March 2020).
25. New Zealand Guidelines Group. The Diagnosis and Management of Soft Tissue Shoulder Injuries and Related Disorders 2020. Available online: https://cdn.ymaws.com/www.alaskachiropracticsociety.com/resource/resmgr/imported/shoulder.pdf (accessed on 20 March 2020).
26. Boer, B.C.; Boetje, J.; Stevens, M.; van den Akker-Scheek, I.; van Raay, J. Adaptation, validity and reliability of the modified painDETECT questionnaire for patients with subacromial pain syndrome. *PLoS ONE* **2019**, *14*, e0211880. [CrossRef] [PubMed]
27. Molsted, S.; Tribler, J.; Snorgaard, O. Musculoskeletal pain in patients with type 2 diabetes. *Diabetes Res. Clin. Pract.* **2012**, *96*, 135–140. [CrossRef] [PubMed]
28. Miksch, A.; Hermann, K.; Rolz, A.; Joos, S.; Szecsenyi, J.; Ose, D.; Rosemann, T. Additional impact of concomitant hypertension and osteoarthritis on quality of life among patients with type 2 diabetes in primary care in Germany—A cross-sectional survey. *Health Qual. Life Outcomes* **2009**, *7*, 19. [CrossRef] [PubMed]
29. ESSR. Ultrasound Shoulder Protocol of the ESSR. 2016. Available online: https://essr.org/content-essr/uploads/2016/10/shoulder.pdf (accessed on 20 March 2020).
30. Lee, J.C.; Sykes, C.; Saifuddin, A.; Connell, D. Adhesive capsulitis: Sonographic changes in the rotator cuff interval with arthroscopic correlation. *Skelet. Radiol.* **2005**, *34*, 522–527. [CrossRef]
31. Tandon, A.; Dewan, S.; Bhatt, S.; Jain, A.K.; Kumari, R. Sonography in diagnosis of adhesive capsulitis of the shoulder: A case-control study. *J. Ultrasound* **2017**, *20*, 227–236. [CrossRef]
32. Ottenheijm, R.P.; Joore, M.A.; Walenkamp, G.H.; Weijers, R.E.; Winkens, B.; Cals, J.W.; de Bie, R.A.; Dinant, G.J. The Maastricht Ultrasound Shoulder pain trial (MUST): Ultrasound imaging as a diagnostic triage tool to improve management of patients with non-chronic shoulder pain in primary care. *BMC Musculoskelet. Disord.* **2011**, *12*, 154. [CrossRef]
33. Geraets, J.J.; de Jongh, A.C.; Boeke, A.J.; Buis, P.A.; Spinnewijn, W.E.; Geijer, R.M.; Goudswaard, A.N. Summary of the practice guideline for shoulder complaints from the Dutch College of General Practitioners. *Ned. Tijdschr. Geneeskd.* **2009**, *153*, A164.
34. Hegedus, E.J.; Goode, A.P.; Cook, C.E.; Michener, L.; Myer, C.A.; Myer, D.M.; Wright, A.A. Which physical examination tests provide clinicians with the most value when examining the shoulder? Update of a systematic review with meta-analysis of individual tests. *Br. J. Sports Med.* **2012**, *46*, 964–978. [CrossRef]
35. Gismervik, S.; Drogset, J.O.; Granviken, F.; Rø, M.; Leivseth, G. Physical examination tests of the shoulder: A systematic review and meta-analysis of diagnostic test performance. *BMC Musculoskelet. Disord.* **2017**, *18*, 41. [CrossRef] [PubMed]
36. Van Seventer, R.; Vos, C.; Giezeman, M.; Meerding, W.J.; Arnould, B.; Regnault, A.; van Eerd, M.; Martin, C.; Huygen, F. Validation of the Dutch version of the DN4 diagnostic questionnaire for neuropathic pain. *Pain Pract. Off. J. World Inst. Pain* **2013**, *13*, 390–398. [CrossRef] [PubMed]
37. Van Seventer, R.; Vos, C.; Meerding, W.; Mear, I.; Le Gal, M.; Bouhassira, D.; Huygen, F.J. Linguistic validation of the DN4 for use in international studies. *Eur. J. Pain* **2010**, *14*, 58–63. [CrossRef] [PubMed]
38. Spallone, V.; Morganti, R.; D'Amato, C.; Greco, C.; Cacciotti, L.; Marfia, G.A. Validation of DN4 as a screening tool for neuropathic pain in painful diabetic polyneuropathy. *Diabet. Med.* **2012**, *29*, 578–585. [CrossRef] [PubMed]
39. Bouhassira, D.; Attal, N.; Alchaar, H.; Boureau, F.; Brochet, B.; Bruxelle, J.; Cunin, G.; Fermanian, J.; Ginies, P.; Grun-Overdyking, A.; et al. Comparison of pain syndromes associated with nervous or somatic lesions and development of a new neuropathic pain diagnostic questionnaire (DN4). *Pain* **2005**, *114*, 29–36. [CrossRef]

40. England, J.D.; Gronseth, G.S.; Franklin, G.; Miller, R.G.; Asbury, A.K.; Carter, G.T.; Cohen, J.A.; Fisher, M.A.; Howard, J.F.; Kinsella, L.J.; et al. Distal symmetric polyneuropathy: A definition for clinical research: Report of the American Academy of Neurology, the American Association of Electrodiagnostic Medicine, and the American Academy of Physical Medicine and Rehabilitation. *Neurology* **2005**, *64*, 199–207. [CrossRef]
41. Alport, A.R.; Sander, H.W. Clinical approach to peripheral neuropathy: Anatomic localization and diagnostic testing. *Continuum* **2012**, *18*, 13–38. [CrossRef]
42. Yang, Z.; Zhang, Y.; Chen, R.; Huang, Y.; Ji, L.; Sun, F.; Hong, T.; Zhan, S. Simple tests to screen for diabetic peripheral neuropathy. *Cochrane Database Syst. Rev.* **2014**. [CrossRef]
43. Feng, Y.; Schlösser, F.J.; Sumpio, B.E. The Semmes Weinstein monofilament examination as a screening tool for diabetic peripheral neuropathy. *J. Vasc. Surg.* **2009**, *50*, 675–682. [CrossRef]
44. Smieja, M.; Hunt, D.L.; Edelman, D.; Etchells, E.; Cornuz, J.; Simel, D.L. Clinical examination for the detection of protective sensation in the feet of diabetic patients. International Cooperative Group for Clinical Examination Research. *J. Gen. Intern. Med.* **1999**, *14*, 418–424. [CrossRef]
45. Kästenbauer, T.; Sauseng, S.; Brath, H.; Abrahamian, H.; Irsigler, K. The value of the Rydel-Seiffer tuning fork as a predictor of diabetic polyneuropathy compared with a neurothesiometer. *Diabet. Med.* **2004**, *21*, 563–567. [CrossRef] [PubMed]
46. Liniger, C.; Albeanu, A.; Bloise, D.; Assal, J.P. The tuning fork revisited. *Diabet. Med.* **1990**, *7*, 859–864. [CrossRef] [PubMed]
47. Fisher, L.; Kurtz, A.; Shipley, M. Association between cheiroarthropathy and frozen shoulder in patients with insulin-dependent diabetes mellitus. *Br. J. Rheumatol.* **1986**, *25*, 141–146. [CrossRef] [PubMed]
48. Hill, N.E.; Roscoe, D.; Stacey, M.J.; Chew, S. Cheiroarthropathy and tendinopathy in diabetes. *Diabet. Med.* **2019**, *36*, 939–947. [CrossRef] [PubMed]
49. Gokcen, N.; Cetinkaya Altuntas, S.; Coskun Benlidayi, I.; Sert, M.; Nazlican, E.; Sarpel, T. An overlooked rheumatologic manifestation of diabetes: Diabetic cheiroarthropathy. *Clin. Rheumatol.* **2019**, *38*, 927–932. [CrossRef]
50. Lewis, J. Rotator cuff related shoulder pain: Assessment, management and uncertainties. *Manual Ther.* **2016**, *23*, 57–68. [CrossRef]
51. Ottenheijm, R.P.; Cals, J.W.; Weijers, R.; Vanderdood, K.; de Bie, R.A.; Dinant, G.J. Ultrasound imaging for tailored treatment of patients with acute shoulder pain. *Ann. Fam. Med.* **2015**, *13*, 53–55. [CrossRef]
52. Cadogan, A.; Laslett, M.; Hing, W.A.; McNair, P.J.; Coates, M.H. A prospective study of shoulder pain in primary care: Prevalence of imaged pathology and response to guided diagnostic blocks. *BMC Musculoskelet. Disord.* **2011**, *12*, 119. [CrossRef]
53. Ottenheijm, R.P.; van't Klooster, I.G.; Starmans, L.M.; Vanderdood, K.; de Bie, R.A.; Dinant, G.J.; Cals, J.W. Ultrasound-diagnosed disorders in shoulder patients in daily general practice: A retrospective observational study. *BMC Fam. Pract.* **2014**, *15*, 115. [CrossRef]
54. Oliva, F.; Via, A.G.; Maffulli, N. Physiopathology of intratendinous calcific deposition. *BMC Med.* **2012**, *10*, 95. [CrossRef]
55. Cook, J.L.; Purdam, C.R. Is tendon pathology a continuum? A pathology model to explain the clinical presentation of load-induced tendinopathy. *Br. J. Sports Med.* **2009**, *43*, 409–416. [CrossRef] [PubMed]
56. Girish, G.; Lobo, L.G.; Jacobson, J.A.; Morag, Y.; Miller, B.; Jamadar, D.A. Ultrasound of the shoulder: Asymptomatic findings in men. *AJR Am. J. Roentgenol.* **2011**, *197*, W713–W719. [CrossRef] [PubMed]
57. Lawrence, R.L.; Moutzouros, V.; Bey, M.J. Asymptomatic Rotator Cuff Tears. *JBJS Rev.* **2019**, *7*, e9. [CrossRef] [PubMed]
58. Arkkila, P.E.; Kantola, I.M.; Viikari, J.S.; Ronnemaa, T. Shoulder capsulitis in type I and II diabetic patients: Association with diabetic complications and related diseases. *Ann. Rheum Dis.* **1996**, *55*, 907–914. [CrossRef] [PubMed]
59. Noël, E.; Thomas, T.; Schaeverbeke, T.; Thomas, P.; Bonjean, M.; Revel, M. Frozen shoulder. *Joint Bone Spine* **2000**, *67*, 393–400. [PubMed]
60. Naunton, J.; Harrison, C.; Britt, H.; Haines, T.; Malliaras, P. General practice management of rotator cuff related shoulder pain: A reliance on ultrasound and injection guided care. *PLoS ONE* **2020**, *15*, e0227688. [CrossRef]

61. Ottenheijm, R.P.; Hesselmans, N.J.; Kemper, A.; Moser, A.; de Bie, R.A.; Dinant, G.J.; Cals, J.W. GPs' perspectives on the diagnostic work-up in patients with shoulder pain: A qualitative study. *J. Eval. Clin. Pract.* **2014**, *20*, 239–245. [CrossRef]

Journal of
Clinical Medicine

MDPI

Review

Recommendations for Diagnosis and Treatment of Lumbosacral Radicular Pain: A Systematic Review of Clinical Practice Guidelines

Ahmad Khoshal Khorami [1], Crystian B. Oliveira [2,3], Christopher G. Maher [3,4], Patrick J. E. Bindels [1], Gustavo C. Machado [3,4], Rafael Z. Pinto [5], Bart W. Koes [1,6] and Alessandro Chiarotto [1,*]

[1] Department of General Practice, Erasmus MC, University Medical Center Rotterdam, 3000 CA Rotterdam, The Netherlands; a.khorami@erasmusmc.nl (A.K.K.); p.bindels@erasmusmc.nl (P.J.E.B.); b.koes@erasmusmc.nl (B.W.K.)
[2] Physical Therapy Department, Faculty of Medicine, University of Western São Paulo (UNOESTE), Presidente Prudente, Sao Paulo CEP 19060-900, Brazil; crystianboliveira@gmail.com
[3] Institute for Musculoskeletal Health, Sydney Local Health District, Sydney, NSW 2050, Australia; christopher.maher@sydney.edu.au (C.G.M.); gustavo.machado@sydney.edu.au (G.C.M.)
[4] Sydney School of Public Health, Faculty of Medicine and Health, University of Sydney, Sydney, NSW 2050, Australia
[5] Department of Physical Therapy, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte 30000-000, Brazil; rafaelzambelli@gmail.com
[6] Center for Muscle and Joint Health, University of Southern Denmark, 5230 Odense, Denmark
* Correspondence: a.chiarotto@erasmusmc.nl; Tel.: +31-1-0704-1038

check for updates

Citation: Khorami, A.K.; Oliveira, C.B.; Maher, C.G.; Bindels, P.J.E.; Machado, G.C.; Pinto, R.Z.; Koes, B.W.; Chiarotto, A. Recommendations for Diagnosis and Treatment of Lumbosacral Radicular Pain: A Systematic Review of Clinical Practice Guidelines. *J. Clin. Med.* **2021**, *10*, 2482. https://doi.org/10.3390/jcm10112482

Academic Editor: Jo Nijs

Received: 12 May 2021
Accepted: 28 May 2021
Published: 3 June 2021

Abstract: The management of patients with lumbosacral radicular pain (LRP) is of primary importance to healthcare professionals. This study aimed to: identify international clinical practice guidelines on LRP, assess their methodological quality, and summarize their diagnostic and therapeutic recommendations. A systematic search was performed (August 2019) in MEDLINE, PEDro, National Guideline Clearinghouse, National Institute for Health and Clinical Excellence (NICE), New Zealand Guidelines Group (NZGG), International Guideline Library, Guideline central, and Google Scholar. Guidelines presenting recommendations on diagnosis and/or treatment of adult patients with LRP were included. Two independent reviewers selected eligible guidelines, evaluated quality with Appraisal of Guidelines Research & Evaluation (AGREE) II, and extracted recommendations. Recommendations were classified into 'should do', 'could do', 'do not do', or 'uncertain'; their consistency was labelled as 'consistent', 'common', or 'inconsistent'. Twenty-three guidelines of varying quality (AGREE II overall assessment ranging from 17% to 92%) were included. Consistent recommendations regarding diagnosis are ('should do'): Straight leg raise (SLR) test, crossed SLR test, mapping pain distribution, gait assessment, congruence of signs and symptoms. Routine use of imaging is consistently not recommended. The following therapeutic options are consistently recommended ('should do'): educational care, physical activity, discectomy under specific circumstances (e.g., failure of conservative treatment). Referral to a specialist is recommended when conservative therapy fails or when steppage gait is present. These recommendations provide a clear overview of the management options in patients with LRP.

Keywords: lumbar radicular pain; clinical practice guidelines; AGREE II; diagnosis; treatment

1. Introduction

Low back pain (LBP) is globally not only a major medical problem but also a major economic problem [1]. Despite intensified research efforts on LBP management, the population burden and disability related to this disorder is increasing [2–4]. According to The Global Burden of Disease 2017 study, Years Lived with Disability due to LBP have globally increased by 54% between 1990 and 2015 [5]. LBP affects many people, especially

female individuals and those aged 40–80 years, with a mean point prevalence of 11.9%, and 1-month prevalence of 23.2% [6]. Among patients with LBP seeking care in primary care, approximately 36% also report low back-related leg pain below the knee [7].

Low back-related leg pain is either radicular or referred (non-specific) pain. The former is described as radiating pain where a spinal nerve root is involved causing leg pain along the spinal nerve accompanied by numbness and tingling, muscle weakness and loss of reflexes. The latter is described as pain spreading down the legs arising from structures such as disc, joints or ligaments [8]. In the literature, multiple terms are used for lumbosacral radicular pain (LRP), with lumbar disc herniation being the most common diagnosis for this condition [9]. A large majority of patients with LRP tend to have a favourable prognosis in terms of pain and disability, but the time to recovery is usually longer than that of patients with LBP without concomitant radicular pain [7,10–12]. Therefore, an accurate assessment of these patients is needed to provide adequate management and treatment at an early stage of presentation.

Clinical practice guidelines are developed for implementing strong evidence into clinical care, to improve quality of care and reduce variation in decision making of healthcare practitioners. Over the last decades, an increasing number of guidelines have been developed in different countries for patients with LBP [13]. In most of these guidelines, diagnostic triage is recommended (i.e., classification into specific or non-specific LBP) [13], and some guidelines include diagnostic and therapeutic recommendations for different types of LBP. Additionally, an increasing number of guidelines containing specific recommendations for LRP have been issued over the past years.

Since 2001, overviews of clinical guidelines for the management of patients with LBP have been conducted and updated [13–17]. However, these overviews have focused on clinical recommendations for patients with acute or chronic non-specific LBP. No systematic review has been conducted on clinical practice guidelines for patients with LRP. Moreover, while the methodological quality of guidelines on non-specific LBP has been reviewed up to 2009 for acute LBP [14] and up to 2018 for chronic LBP [18], a quality assessment of guidelines focusing only on LRP has never been performed. Since 2009, the Appraisal of Guidelines Research & Evaluation (AGREE) II instrument can be used to assess the methodological quality of clinical practice guidelines [19].

The aim of this study was to retrieve all existing guidelines formulating recommendations on the clinical management of patients with LRP, to assess their methodological quality using the AGREE II tool, and to summarize their diagnostic and therapeutic recommendations.

2. Methods

2.1. Review Registration and Reporting

The protocol for the review was registered on the International Prospective Register of Systematic Reviews (PROSPERO) with ID number CRD42020138738. This review was reported in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [20].

2.2. Literature Search

Literature searches were performed in the following electronic databases from their inception up to August 2019: MEDLINE via OVID, PEDro, National Guideline Clearinghouse, National Institute for Health and Clinical Excellence (NICE), New Zealand Guidelines Group (NZGG), International Guideline Library, and Guideline central. No language restrictions were applied. We also conducted web searches on Google Scholar using recommendations described elsewhere [21], but only the first 10 pages were screened because they retrieve the most relevant search results [21]. Furthermore, backward citation tracking on the reference lists of previous relevant reviews on the topic was performed [13,18]. Appendix A provides the search strategy used in each database.

2.3. Guidelines Selection

Clinical practice guidelines providing recommendations regarding diagnosis and treatment of LRP were included. Only guidelines formulating recommendations based on evidence were included. A guideline was included regardless of the type of professional association (e.g., physiotherapy, chiropractic, multidisciplinary), geographical location, and date of publication. The following terms were considered as LRP synonyms: sciatica, radiculopathy, nerve root compromise, nerve root compression, lumbar radicular syndrome, disc herniation, radiculitis, nerve root pain, and nerve root entrapment. Documents briefly mentioning LRP for diagnostic triage without providing further details on its management were excluded. When multiple versions of a guideline issued by a similar professional association were available, only the most recent version was selected. Clinical practice guidelines available in English, German, Portuguese, Spanish, Italian or Dutch were included because the author team could understand these languages. If guidelines in other languages were found, translators for extracting the required information from the guidelines were sought. If this method failed, documents for which no translators could be identified were excluded. Title/abstracts and full-texts were screened by two independent reviewers (AKK and CBO) and disagreements were discussed in online consensus meetings. If disagreements could not be solved, a third reviewer (AC) arbitrated.

2.4. Quality Assessment

The methodological quality of clinical practice guidelines was assessed using the AGREE II tool [22]. AGREE II is an update of the previous AGREE instrument to improve its measurement properties (i.e., reliability and validity), to refine its items and to improve the supporting documentation (i.e., original training manual and user's guide) [19]. AGREE II consists of 23 items categorized in six domains: scope and purpose; stakeholder involvement; rigor of development; clarity of presentation; applicability; editorial independence. Each item is scored on a 7-point Likert scale from one (i.e., strongly disagree) to seven (i.e., strongly agree). AGREE II also includes two global items. The first global item is scored similarly to the 23 items and evaluates a guideline overall quality. The second item evaluates the recommendation for use which is rated using a three-point scale (i.e., yes, yes with modification, or no). A previous study [23] using four appraisers found good to excellent inter-rater reliability (intraclass correlation coefficient [$ICC_{2,1}$] ranging from 0.66 to 0.93) for the 23 items and the first global item. The AGREE II manual recommends the appraisal of the guidelines by at least two reviewers [19]. Thus, two independent authors (AKK and CBO) performed the online AGREE II training [24] and followed the AGREE II manual to assess each included guideline [19]. To calculate the score for each domain all the scores of the individual items in a domain from both appraisers were summed up and scaled as a percentage of the maximum possible score. The overall assessment score is the mean score for the 6 different domains. To investigate the reliability among the assessors, we calculated inter-rater reliability from the two appraisers of the scores obtained for each domain and first global item using the $ICC_{2,1}$ and 95% confidence interval. Inter-rater agreement was classified as; 0.75–1.00, excellent; 0.60–0.74, good; 0.40–0.59, fair and <0.40 as poor [25]. Reliability analyses were performed using IBM SPSS Statistics 25 with a two-way random effects model for the domain scores and the first global item scores.

There is not consensus on categorizing guidelines as high, average or low quality guidelines depending on AGREE II scores. However, several methods are provided in the AGREE II manual. A previously described threshold of $\geq$60% for an acceptably high score on a domain [26] was adopted in this review; for the overall quality of the included guidelines: high-quality guideline when $\geq$5 domains were scored $\geq$60%, average quality guideline when 3 or 4 domains were scored $\geq$ 60% and low-quality guideline when $\leq$2 domains scored $\geq$ 60%.

2.5. Data Extraction and Synthesis

Two authors (AKK and CBO) independently extracted data from the included guidelines. In cases of no consensus, a third reviewer (AC) was consulted. The following information was extracted using a standardized form: recommendations regarding diagnosis (e.g., history, physical examination) and treatment (e.g., patient education, pharmacological intervention). Regarding surgical treatment options only recommendations for discectomy were extracted as this procedure is the most common applied for disc herniation. As radicular pain is also covered in most generic LBP guidelines, recommendations were extracted when it was clearly specified that they concerned radicular pain, or if recommendations for LBP in general also applied to radicular pain. If available, the level of evidence considered to formulate each recommendation was also extracted.

In order to assess the type and direction of the recommendations, the extracted recommendations were firstly classified into one of the categories: 'Should do', 'Could do', 'Do not do' and 'Uncertain'. This classification was dependent on the terminology used for a recommendation in the guideline (see Appendix B, Table A1). Secondly, to determine the consistency of a recommendation the following categories were identified using a modified version of the approach previously adopted by Lin et al. [27]:

a. Consistent recommendations: from the guidelines including recommendation for a specific approach, the majority (≥80%) indicate as 'should do', 'could do', 'do not do', or 'uncertain', but without conflicting recommendations across guidelines. Conflicting recommendations are present when at least one 'should do' or 'could do', and at least one 'do not do' is applied for the same recommendation in different guidelines.
b. Common recommendations: from the guidelines including recommendation for a specific approach, most (between 50% and 80%) indicate as 'should do', 'could do', 'do not do', or 'uncertain', but with no conflicting recommendations across guidelines.
c. Inconsistent recommendations: a recommendation for one approach indicates 'should do' or 'could do', and another recommendation for the same approach indicates 'do not do' or 'uncertain', both recommendations issued by different guidelines; the same applies if a recommendation for an approach is 'uncertain', and another recommendation for the same approach is 'do not do'.

To determine the consistency across recommendations, only the options included at least two different guidelines were considered.

3. Results

3.1. Guidelines Selection

Our literature searches identified 3032 records. After screening of title/abstract and full texts, 23 eligible guidelines were included (Figure 1). The characteristics of these guidelines are presented in Table 1. The 23 included guidelines [28–50] were developed in 10 different countries from 3 continents (i.e., North America, Europe and Asia): United States (n = 12, 52%), Canada (n = 2), Belgium (n = 1), Denmark (n = 1), Korea (n = 1), Philippines (n = 1), Italy (n = 1), the Netherlands (n = 1) and Norway (n = 1). One guideline was a joint European guideline. One guideline was written in Dutch [45], one updated guideline in Norwegian [44] and the other 21 in English. The professional entities involved in developing the guidelines vary in different countries (Table 1). Most guidelines (n = 14, 61%) are from a specific (medical) professional association (e.g., general practitioners, pain physicians, radiologists, chiropractors, physiotherapists).

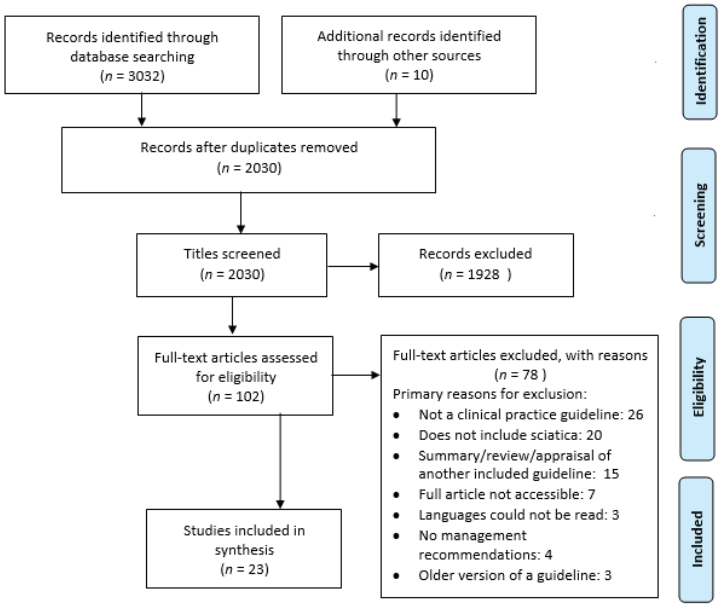

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.

Table 1. Characteristics of included clinical practice guidelines on recommendations for patients with lumbar radicular pain (n = 23).

Title Guideline	Year	Country	Professional Bodies/ Abbreviation
Low back disorders	2016	USA	American College of Occupational and Environmental Medicine (ACOEM [28])
Diagnostic Imaging for Low Back Pain: Advice for High-Value Health Care From the American College of Physicians	2011	USA	American College of Physicians (ACP [29])
Diagnosis and Treatment of Low Back Pain: A Joint Clinical Practice Guideline from the American College of Physicians and the American Pain Society	2007	USA	American College of Physicians and the American Pain Society (ACP-APS [30])
ACR Appropriateness Criteria Low Back Pain	2016	USA	American College of Radiology (ACR [31])
Interventional Therapies, Surgery, and Interdisciplinary Rehabilitation for Low Back Pain	2009	USA	American Pain Society (APS [32])
Low Back Pain: Clinical Practice Guidelines Linked to the International Classification of Functioning, Disability, and Health from the Orthopedic Section of the American Physical Therapy Association	2012	USA	Orthopedic Section of the American Physical Therapy Association (APTA [33])
An Update of Comprehensive Evidence-Based Guidelines for Interventional Techniques in Chronic Spinal Pain. Part II: Guidance and Recommendations	2013	USA	American Society of Interventional Pain Physicians (ASIPP [34])
Spinal Manipulative Therapy and Other Conservative Treatments for Low Back Pain: A Guideline From the Canadian Chiropractic Guideline Initiative	2018	Canada	Canadian Chiropractic Guideline Initiative (CCGI [35])

Table 1. *Cont.*

Title Guideline	Year	Country	Professional Bodies/ Abbreviation
Low Back Pain medical Treatment Guidelines	2014	USA	Department of Labor and Employment, Division of Worker's Compensation (DLE-DWC [36])
European guidelines for the management of chronic non-specific low back pain	2004	Europe	European Commission, Research Directorate-General, department of Policy, Coordination and Strategy (EG [37])
Adult Acute and Subacute Low Back Pain Diagnosis Algorithm	2018	USA	Institute for Clinical Systems Improvement (ICSI [38])
Diagnostic therapeutic flow-charts for low back pain patients: the Italian clinical guideline	2006	Italy	Italian Health Ministry-Care and Research Institute Fondazione Don Carlo Gnocchi ONLUS of Milan (IHM [39])
Low back pain and radicular pain: assessment and management	2017	Belgian	Belgian Health Care Knowledge Center (KCE [40])
Korean medicine clinical practice guideline for lumbar herniated intervertebral disc in adults	2017	Korea	The Korea Institute of Oriental Medicine (KIOM [41])
Clinical Guideline for the Diagnosis and Treatment of Lumbar Disc Herniation with Radiculopathy	2012	USA	North American Spine Society (NASS [42])
National Clinical Guideline: interventions for recent onset lumbar radiculopathy	2016	Denmark	National Board of Health (Denmark) (NBHD [43])
Acute low back pain Interdisciplinary clinical guidelines	2007	Norway	The Norwegian Back Pain Network and an interdisciplinary working group (NBPN [44])
NHG-Standaard Lumbosacraal radiculair syndroom	2015	Netherlands	Dutch General Practitioners Society (NHG [45])
Low back pain and sciatica in over 16s: assessment and management	2016	UK	National Institute for Health and Care Excellence (NICE [46])
Clinical Practice Guidelines on the Diagnosis and Management of LBP	2011	Philippines	Philippine Academy of Rehabilitation Medicine PARM [47])
Evidence-Informed Primary Care Management of Low Back Pain	2015	Canada	Institute of Health Economics Toward Optimized Practice (TOP [48])
Acute Low Back Pain	2010	USA	The University of Michigan Health System (UMHS [49])
VA/Do clinical practice guideline for diagnosis and treatment of low back pain	2017	USA	Department of Veterans Affairs and Department of Defense (Va/Dod [50])

3.2. Quality Assessment and Inter-Rater Agreement

In Table 2, the AGREE II scores for each domain and the overall assessment scores are displayed for each included guideline. The overall quality score was variable, ranging from 17% to 92%, with the NICE guideline having the highest score and the Department of Labor and Employment, Division of Worker's Compensation (DLW-DWC)the lowest score. Of the 23 included guidelines, 15 (65%) scored above the $\geq$ 60% threshold in the overall assessment score. According to our described classification criteria [26], ten [32,34,38,40,42–44,46,48,50] (43%) of the guidelines are classified as high quality, seven [29,30,33,35,37,39,41] (30%) as average quality, and six [28,31,36,45,47,49] (26%) as low quality. Five guidelines (American Pain Society (APS) and Institute for Clinical Systems Improvement (ICSI) from USA, NICE from UK, Belgian Health Care Knowledge Center (KCE) from Belgium, and Institute of Health Economics Toward Optimized Practice (TOP) from Canada) displayed a score $\geq$ 60% in all 6 domains (Table 2).

Table 2. Appraisal of Guidelines Research & Evaluation (AGREE) II domain scores and quality of the eligible guidelines (n = 23).

Guideline	AGREE Domains							
Guideline *	Scope and Purpose, %	Stakeholder Involvement, %	Rigor of Development, %	Clarity of Presentation, %	Applicability, %	Editorial Independence, %	Overall Assessment, %	Quality
ACOEM	58	44	50	100	23	92	61	Low
ACP	86	36	11	83	46	100	60	Average
ACP-APS	97	47	76	94	46	96	76	Average
ACR	14	17	81	67	0	17	33	Low
APS	100	67	94	94	60	96	85	High
APTA	81	47	72	89	0	0	48	Average
ASIPP	83	64	72	67	29	96	68	High
CCGI	92	72	74	58	33	67	66	Average
DLE-DWC	17	11	0	72	2	0	17	Low
EG	72	64	70	92	33	38	61	Average
ICSI	83	69	76	92	60	92	79	High
IHM	61	58	68	83	50	25	58	Average
KCE	97	75	88	86	77	100	87	High
KIOM	78	28	76	81	6	83	59	Average
NASS	83	61	75	78	17	92	68	High
NBHD	86	47	91	97	60	96	80	High
NBPN	89	83	83	94	29	100	80	High
NHG	83	56	58	89	25	33	57	Low
NICE	97	94	95	97	71	100	92	High
PRAM	72	25	31	81	38	0	41	Low
TOP	94	78	84	94	73	96	87	High
UMHS	56	25	57	94	6	92	55	Low
VA/Dod	94	92	77	83	44	67	76	High
Mean	77	55	68	85	36	69	65	

* Table 1 provides the extensive names of the included guidelines.

The domains with the highest scores were "Clarity of presentation" and "Scope and purpose" with mean scores of 85% and 77%, respectively (Table 2). The domains that were less well addressed by guideline developers were "Applicability" and "Stakeholder involvement" with mean scores of 36% and 55%, respectively. "Editorial independence" is the domain with the highest variability among the guidelines ranging from 0% to 100% (Table 2).

Table 3 indicates inter-rater agreement for AGREE II domains and overall scores. The $ICC_{2,1}$ was 'excellent' for the domains "Scope and purpose", "Stakeholder involvement", "Rigor of development", "Applicability", "Editorial independence" and "Overall rating", and 'fair' for the domain "Clarity of presentation". Due to the small sample of assessed guidelines (n = 23), ICC 95% confidence intervals were broad, especially for the domain with the lowest inter-rater agreement (Table 3).

Table 3. Inter-rater agreement for AGREE II domains and overall rating.

Domain	ICC * (95% CI)
Scope and purpose	0.847 (0.631 to 0.936)
Stakeholder involvement	0.820 (0.2563 to 0.926)
Rigor of development	0.858 (0.636 to 0.943)
Clarity of presentation	0.549 (-0.044 to 0.811)
Applicability	0.874 (0.695 to 0.948)
Editorial independence	0.901 (0.762 to 0.959)
Overall rating	0.785 (0.493 to 0.910)

* ICC; Intraclass Correlation Coefficient.

3.3. Recommendations for Diagnosis

Table 4 and supplementary Table S4 describe the recommendations for physical examination and other diagnostic procedures in each clinical practice guideline.

Table 4. Guideline recommendations for physical examination and other diagnostic procedures.

Physical Examination	Guideline *	Consistency	Clinical Inference
Femoral stretch test	NASS NBPN PARM	Inconsistent	None
Straight leg test	NASS PARM IHM NHG ACP-APS NBPN	Consistent	Should do
Crossed straight leg test	NASS PARM IHM NHG NBPN	Consistent	Should do
Muscle testing	NASS PARM IHM NHG ACP-APS NBPN	Common/consistent	Should/Could do
Sensory testing	NASS PARM IHM ACP-APS NBPN	Common	Could do

Table 4. *Cont.*

Physical Examination	Guideline *	Consistency	Clinical Inference
Reflex tests (ankle and knee tendon)	PARM	Inconsistent	None
	NBPN		
	ACP-APS		
	NASS		
Mapping pain distribution	PARM	Consistent	Should do
	IHM		
Slump test	PARM	Inconsistent	None
	NASS		
Wasserman test	PARM	Consistent	Could do
	IHM		
Gait	PARM	Consistent	Should do
	IHM		
Agreement of signs and symptoms	PARM	Consistent	Should do
	IHM		
Diagnostics			
Imaging			
Routinely offering imaging in primary care or absent of red flags	KCE	Consistent	Do not do
	ICSI		
	NICE		
	ACOEM		
	CCGI		
	ACR		
	NBPN		
Computed Tomography (CT)/ Magnetic resonance imaging (MRI) routinely in first 4–6 weeks	ICSI	Consistent	Do not do
CT when history and physical examination findings consistent with disc herniation, after 4–6 weeks of low back pain if surgery is considered, severe or progressive neurologic signs and symptoms present	ACOEM	Consistent	Should do
	NHG		
	PRSM		
	IHM		
	NBPN		
	NASS		
	VA/Dod		
	TOP		
	PRAM		
	ACOEM		
	IHM		
	ACP-APS		
	NBPN		
	ACR		
	NASS		
MRI when history and physical examination findings consistent with disc herniation, radiculopathy persists after six weeks, if surgery is considered, severe or progressive neurologic signs and symptoms present, where an epidural glucocorticosteroid injection is being considered	VA/DoD	Common	Should do
	TOP		
	NBHD		
	EG		
	PRAM		
	ACOEM		
	IHM		
	ACP-APS		
	UMHS		
	NBPN		
	ACR		
	ACP		

Table 4. *Cont.*

Physical Examination	Guideline *	Consistency	Clinical Inference
Others			
EMG	EG PRAM ACOEM IHM UMHS	Inconsistent	None
Sensory nerve somatosensory evoked potentials (SEP)	NASS DLE-DWC	Inconsistent	None
Discography	ACOEM DLE-DWC	Inconsistent	None
Diagnostic medial branch block	APS ACOEM	Inconsistent	None

* Table 1 provides the extensive names of the included guidelines.

3.3.1. Physical Examination

A minority of guidelines (6 out of 23; 26%); [30,39,42,44,45,47] made recommendations concerning physical examination (Table 4). The consistent recommendation for 'should do' in the physical examination are: performing straight leg raise test (SLR) [30,39,42,44,45,47], crossed SLR test [39,42,44,45,47], mapping pain distribution [39,47], steppage gait (inability to lift the foot while walking due to the weakness of muscles that cause dorsiflexion of the ankle joint) assessment [39,47], and agreement of signs and symptoms [39,47].

For the recommendation 'muscle testing', the guidelines are evenly distributed with 'should do' and 'could do' [30,39,42,44,45,47]. The inconsistent recommendations concerned the performance of: femoral stretch test [42,44,47], reflex tests [30,38,42,44,47] and slump test [42,47].

3.3.2. Diagnostics

A majority of the guidelines (19 out of 23, 82%; [28–31,35–40,42–50]) made recommendations concerning diagnostics (Table 4 and supplementary Table S4). A consistent recommendation for 'should do' in imaging is to perform a computed tomography (CT) scan when history and physical examination findings are consistent with disc herniation, after 4–6 weeks of pain, if surgery is considered or severe or progressive neurologic signs and symptoms are present [28,30,31,39,42,44,47,50]. A common recommendation for 'should do' concerns magnetic resonance imaging (MRI) when history and physical examination findings are consistent with disc herniation, radiculopathy persists after six weeks, if surgery is considered, severe or progressive neurologic signs and symptoms are present or where an epidural glucocorticosteroid injection is being considered [28–31,37,39,42–44,47–50].

Consistent recommendations for 'do not do' are: routinely offering imaging in primary care or in the absence of red flags [28,31,35,38,40,44,46] and routine computed tomography/magnetic resonance imaging (CT/MRI) scans in the first 4–6 weeks [28,38,39,44,45,47]. There were inconsistent recommendations regarding the use of electromyography (EMG) [28,37,39,47,49], sensory nerve somatosensory evoked potentials (SEP) [36,42], discography [28,36] and diagnostic medial branch block [28,32].

3.4. Guideline Recommendations for Treatment

Table 5 and supplementary Table S5 describes the three treatment categories (i.e., non-invasive, pharmacological and invasive) provided by each clinical practice guideline.

Table 5. Therapeutic recommendations from guidelines for lumbar radicular pain.

Non-Invasive Interventions	Guideline *	Consistency	Clinical Inference
Bed rest	ACOEM PARM IHM NHG NBPN	Inconsistent	None
Physical activity	NBHD PARM IHM ACP-APS NICE NHG NBPN	Common	Should do
Educational care	NICE ICSI NHG ACP-APS NBPN	Consistent	Should do
Multidisciplinary approach/rehabilitation program/Psychological therapy	NICE KCE UMHS NBPN	Common	Could do
Alternative medicine			
Acupuncture	ICSI PARM ACOEM IHM NICE KIOM NBPN	Inconsistent	None
Manual therapies			
Traction	NASS KCE PRAM ACOEM DLE-DWC NICE NBPN APTA	Inconsistent	None
Manipulation/mobilisation/soft-tissue techniques	NASS KCE ICSI NBHD PARM ACOEM DLE-DWC IHM NICE NHG CCGI NBPN APTA	Inconsistent	None
Massage	PRAM ACOEM IHM NBPN	Inconsistent	None

Table 5. *Cont.*

Non-Invasive Interventions	Guideline *	Consistency	Clinical Inference
Devices (e.g., belts, corset, foot orthotics etc.)	NICE KCE ACOEM NBPN	Consistent	Do not do
Exercise/physical therapies	NASS KCE ACOEM NBHD CCGI NBPN APTA	Consistent	Could do
Electrotherapies *TENS/PENS/interferential therapy*	NICE KCE PRAM ACOEM IHM NBPN	Consistent	Do not do
Therapeutic ultrasound	KCE PARM NICE	Inconsistent	None
Heat/cold/infrared therapies	ICSI PRAM UMHS ACOEM	Inconsistent	None
Pharmacological interventions			
Paracetamol	KCE PARM ACOEM IHM UMHS NHG NBPN	Inconsistent	None
Non-steroidal anti-inflammatory drugs (NSAIDs)	KCE ICSI PARM ACOEM IHM NHG NBPN	Inconsistent	None
Opioids	NICE KCE PARM IHM NHG NBPN	Inconsistent	None
Paracetamol + opioids	IHM NHG NBPN	Inconsistent	None
Anti-epilepticum	KCE VA/Dod PARM ACOEM NASS NICE NHG	Inconsistent	None

Table 5. *Cont.*

Non-Invasive Interventions	Guideline *	Consistency	Clinical Inference
Muscle relaxants	KCE ICSI PARM ACOEM IHM NHG NBPN	Inconsistent	None
Antidepressants	NASS NICE KCE ACOEM NHG	Inconsistent	None
Corticosteroids	VA/Dod PARM ACOEM DLE-DWC NHG ACP-APS	Inconsistent	None
Antibiotics	KCE ACOEM	Inconsistent	None
Cannabis	NICE NHG	Consistent	Do not do
Invasive Treatments			
Surgery	NASS APS NHG DLE-DWC IHM ACOEM NBPN	Common	Could do
Injection therapies *Epidural injections*	NICE NASS KCE VA/Dod ICSI NBHD PRAM ACOEM ACP NHG NBPN APS	Inconsistent	None
Referral	PRAM IHM NHG	Consistent	Should do

* Table 1 provides the extensive names of the included guidelines.

3.4.1. Non-Invasive Treatments

Most guidelines (16 out of 23; 70%; [28,30,33,35,36,38–47,49]) included recommendations regarding non-invasive treatments. The consistent recommendation for 'should do' is: educational care [30,38,44–46] and the common recommendation for 'should do' is physical activity [30,39,43–47,49]. Exercise/physical therapies [28,33,35,40,43,44] is consistently recommended as 'could do'. The consistent recommendations for 'do not do' are: devices (e.g., belts, corset, foot orthotics etc.) [28,40,44,46] and transcutaneous electrical nerve stimulation (TENS)/percutaneous electrical nerve stimulation (PENS)/interferential

therapy [28,39,40,44,46,47]. The inconsistent recommendations are: bed rest [28,39,44,45,47,49], acupuncture [28,38,39,41,43,44,46,47], traction [28,33,36,38,42,44,46,47], manipulations/ mobilisations/soft tissue techniques [28,33,35,36,38–40,42–47], massage [28,39,44,47], therapeutic ultrasound [40,46,47] and heat/cold/infrared therapies [28,38,47,49].

3.4.2. Pharmacological Interventions

Nine guidelines (out of 23; 39%; [28,38–40,44–47,49]) included recommendations regarding pharmacological interventions. The only consistent recommendation is for 'do not do' for medicinal cannabis [45,46]. For all the other medications, such as paracetamol, Non-steroidal anti-inflammatory drugs (NSAIDs), opioids, anticonvulsants, muscle relaxants, antidepressants, corticosteroids and antibiotics, the recommendations are highly inconsistent. For example, while six guidelines [28,38–40,45,47] tend to suggest NSAIDs as an option, one guideline [44] found no evidence to make a recommendation. Six guidelines [28,39,44,45,47,49] suggest paracetamol, but KCE [40] advices against it. NICE [46] advices against the use of strong opioids such as morphine or long term use of Tramadol in a non-specializing setting, but the Dutch General Practitioners Society (NHG) [45] guideline suggests morphine or fentanyl for chronic pain. Four guidelines [39,44,46,47] recommend a weak opioid such as tramadol as an option, where KCE [40] advises against any opioid.

3.4.3. Invasive Treatments and Referral

Thirteen guidelines (out of 23; 57%; [28,32,36,38–40,42–47,50]) included recommendations regarding invasive treatments and three guidelines (out of 23; 13%; [39,45,47]) regarding referral. The common recommendation in invasive therapy for 'could do' is discectomy when conservative therapy fails or when progressive/persistent disability is present [28,32,36,39,42–45]. The recommendation regarding epidural injections is inconsistent across the guidelines [28,29,32,38,40,42–47,50]. The consistent recommendation for 'should do' is referral to a specialist when there is no improvement of symptoms with conservative therapy, or immediately when there is steppage gait [39,45,47].

4. Discussion

This systematic review provides a summary of the diagnostic and therapeutic recommendations from 23 international clinical practice guidelines for LRP. The consistent and common recommendations for 'should do' for physical examination are performing the SLR test, the crossed SLR test, mapping pain distribution, steppage gait assessment, and evaluating congruence of signs and symptoms. Regarding imaging, guidelines recommend CT scan or MRI under specific circumstances (e.g., physical examination findings are consistent with disc herniation, after 4–6 weeks of pain, surgery or epidural injections are considered, severe or progressive neurologic signs and symptoms present), and do not recommend the routine use of any form of imaging. The consistent and common therapeutic recommendations for 'should do' are: providing educational care and physical activity, referral to a specialist when conservative therapy fails or when steppage gait is present.

This systematic review provides a methodological quality assessment of the 23 selected guidelines using the AGREE II tool. The overall quality of the guidelines ranged from low to high. High quality guidelines (5 out of 9) are from (national) health care institutes [38,40,43,46,48], two from a pain society [32,34], one from a department of veterans affair [50], one from interdisciplinary back pain network [44] and one from the spine society [42]. In this study, "Clarity of presentation" and "Scope and purpose" were the AGREE domains with the highest scores, and "Applicability" and "Stakeholder involvement" with the lowest scores. These finding are in accordance with another critical appraisal of the quality of LBP guidelines [51]. This previous study assessed 5 guidelines [33,38,46,48,49] which were also included in this review (although we included the latest version of the ICSI [38] guideline from 2018). Compared to the previous review, four out of five of these guidelines were classified as the same quality (i.e., high). One guideline [49] was classified as low quality in

this review but of average quality in the other. This discrepancy could be due to a difference in the amount of AGREE II appraisers which could lead to a higher ICC ratio (two appraisers in this study vs. 4 appraisers in Doniselli et al. [51]. Four guidelines [33,46,48,49] included in this review were also included in another review of LBP guidelines [18], where the domain scores were generally similar for the most of the domains. If we apply the same threshold for the domain scores, these 4 guidelines would be categorized with the same quality as in our study. Based on the AGREE II scores of this and earlier studies, it is important that guideline developers take potential barriers of implication of recommendations into account and provide criteria for monitoring and auditing (i.e., AGREE II applicability). Additionally, it is important to include individuals from all relevant professions and take the view of the target population into account (i.e., AGREE II stakeholder involvement).

Although many organizations tend to develop a new guideline, studies have been undertaken on adopting and adapting existed good quality guidelines for saving time and other resources. Schünemann H. et al. [52] has developed the "GRADE-ADOLOPMENT" approach for adopting, adapting and de novo development of recommendations for guideline productions. This approach allows guideline developers to quickly and efficiently create recommendations appropriate for their context where the evidence is taken into account. The Belgian guideline KCE is a good example where the UK comprehensive guideline (NICE) has been adopted and adapted for the Belgian population. Harstall C. et al. [53] have described a multidisciplinary adaption process for creating a single overarching evidence-based clinical practice guideline for patients with LBP. The adolopment approach could facilitate other national and international guideline developers to save time and other important resources and we suggest this approach for LBP when resources are limited and good quality guidelines already exists.

The guidelines commonly recommend physical examination such as performing a SLR test, crossed SLR test, muscle and sensory testing, and reflex test for lumbar radicular pain (Table 4). However, the sensitivity and specificity of these tests have been questioned. A systematic review [54] showed poor diagnostic performance of most physical tests to identify disc herniation when used in isolation, while better performance could be obtained when tests are combined. Two stretch tests that have shown high diagnostic accuracy in patients with LRP are the SLR test and the slump test [55] where the later test was found to be more sensitive. However, according to our finding the recommendation regarding the slump test is inconsistent and six guidelines suggested the SLR test (Table 4). Professionals involved in developing or updating guidelines should consider also the slump test for recommendation. The guidelines commonly/consistently recommend physical activity and exercises/physical therapy. These recommendations were also taken up by a recent narrative review [56] where it is concluded that both physical activity and structured exercises might be beneficial elements in conservative management of patients with lumbar radicular pain.

Besides the consistent and common recommendations for diagnostic and therapeutic options for LRP, we have also identified inconsistent recommendations among the guidelines. Regarding non-invasive treatments, there are inconsistent recommendations on advising bed rest, acupuncture, traction, manipulations/mobilisations/soft tissue techniques, massage, therapeutic ultrasound and heat/cold therapy. Three guidelines ([28,39,47], two low quality and one average quality) do not recommend bed rest with the exception of 2–4 days in severe cases, while two other guidelines ([44,45], one high and one low quality) recommend bed rest for a few days to relieve pain. However, other high quality guidelines [32,34,38,40,42,43,46,48,50] do not include any recommendations on bed rest. Therefore, the question could be raised how strong bed rest recommendation is as a non-invasive therapeutic option. For pharmacological interventions, there are inconsistent recommendations for paracetamol, NSAIDs, opioids, anticonvulsants, muscle relaxants, antidepressants, corticosteroids and antibiotics (Table 5). These inconsistencies in guideline recommendations are not surprising considering that there is still uncertain evidence regarding their effectiveness for patients with LRP [57]. Inconsistent recommendations for

the use of paracetamol is not surprising. In fact, the publication of the placebo controlled trial of paracetamol (PACE) study from 2014, probably explain these differences [58], as no effect was detected favoring paracetamol on pain and speed recovery in patients with acute LBP with or without leg pain. It can be noticed that guidelines published earlier than 2014 recommend paracetamol, whereas recently published guidelines are being careful with making this recommendation. Only the NHG guideline from 2015 [45] still suggests paracetamol as the first analgesic choice. This could lead to a conclusion that most guidelines might have followed the PACE study results considering that a randomized controlled study in patients with LRP is missing. Nevertheless, such study should be conducted as it would provide clearer indications on the efficacy of this drug for future LRP guidelines. A Cochrane review from 2016 [59] showed no significant efficacy of NSAIDs for pain reduction in treating patients with LRP. However, two guidelines [38,40] issued after 2016 suggest NSAIDs to be considered. More research is needed to evaluate the effects of NSAIDs on this condition to reach more firm conclusions. In the category invasive treatments, the recommendation for epidural injections is inconsistent. A recently published Cochrane review [60] concluded that there are small effects of epidural injections in the treatment of lumbar radicular pain, which are mainly evident at short-term follow up.

To the best of our knowledge, this is the first systematic review of clinical practice guidelines focusing on diagnosis and treatment of LRP. A strength of this study is using the AGREE II tool to methodologically assess the quality of the included guidelines and consistent findings in AGREE with other two recent reviews [18,51]. We have transparently reported all recommendations in every single guideline in the Tables (Supplementary Tables S4 and S5). A limitation of this review is that it was not possible to assess the eligibility of three guidelines due to the review team not being able to read and understand the language of these guidelines (one Hebrew and two Croatian). Another limitation is that AGREE II appraisal of the NHG and the Norwegian Back Pain Network (NBPN) guidelines was done by only one appraiser because one of the two reviewers could not understand the Dutch or Norwegian language. In assessing the consistency and clinical inference of the recommendation we have not taken the publication date of a guideline into account. This approach could raise the consideration that this could lead to potentially biased recommendations as an older guideline is equally weighted as a recently published guideline based on more recent evidence. Nevertheless, we performed a sensitivity analysis in a sample of five recommendations from Tables 4 and 5 (i.e., straight leg test, performing CT scan, bed rest, physical activity and paracetamol) in which guidelines published before 2010 were excluded. The consistency of recommendations and clinical inference remained the same. Therefore, the publication date of guideline would not influence the results consistently. We acknowledge the limitation of the AGREE II consortium not setting specific cut-off scores for domains or guidelines of high versus low quality. The cut-off scores used in this review were taken from a small study [26], but they were also used in other recent reviews of clinical practice guidelines using the AGREE II [61–63]. More research on the optimal cut-off scores for the domain and total scores is needed.

This review has highlighted the lack of homogeneity in the manner in which clinical practice guidelines formulate the strength of their recommendations, and the related level of evidence per recommendation. This makes it difficult, at times, to compare and contrast the strength of each recommendation from different guidelines. In our study, we have defined our own terminology by grouping different terms used by the different guidelines (Appendix B). In our view, this point could be addressed by issuing a standardized terminology for strength of a recommendation when developing a clinical practice guideline. Moreover, to date, no standardized threshold value has been suggested to classify high- and low-quality guidelines using the AGREE II appraisal. Introducing a standardized quality classification system based on the AGREE II domain scores is a point of consideration for the future.

5. Conclusions

Twenty-three clinical practice guidelines for patients with LRP were retrieved and their overall quality ranged from low to high according to the AGREE II tool. These guidelines recommend physical examinations perform the SLR test, crossed SLR test, mapping pain distribution, steppage gait assessment, and agreement of signs and symptoms. Imaging is only recommended under specific circumstances, and its routine use is consistently not recommended. For treatment, the recommendations are: providing educational care, prescribing physical activity, and referring to a specialist when conservative therapy fails, or when steppage gait is present. These consistent recommendations should be adopted by healthcare professionals and healthcare systems worldwide to implement the most effective care.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/jcm10112482/s1, Table S4: Diagnostic Recommendations from guidelines for lumbar radicular pain. And Table S5: Therapeutic recommendations from guidelines for lumbar radicular pain.

Author Contributions: Conceptualization, A.K.K., C.B.O., C.G.M., R.Z.P., B.W.K., A.C.; methodology, A.K.K., C.B.O., C.G.M., B.W.K., A.C.; formal analysis, A.K.K., A.C.; investigation, A.K.K., C.B.O., A.C.; resources, C.G.M., R.Z.P., B.W.K.; data curation, A.K.K., C.B.O., A.C.; writing—original draft preparation, A.K.K.; writing—review and editing, A.K.K., C.B.O., C.G.M., R.Z.P., P.J.E.B., G.C.M., B.W.K., A.C.; visualization, A.K.K., A.C.; supervision, B.W.K., A.C.; project administration, A.K.K.; funding acquisition, B.W.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable to this article.

Informed Consent Statement: Informed consent is not applicable to this article.

Data Availability Statement: Data sharing is not applicable to this article.

Acknowledgments: The authors wish to thank Maarten F.M. Engel and Elise Krabbendam from the Erasmus MC Medical Library for developing and updating the search strategies. We would also like to acknowledge Kjersti Storheim for performing the AGREE II assessment and extracting the recommendations from the NBPN Norwegian guideline.

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

Appendix A

Search strategy performed in the electronic databases.
Medline (via Ovid)

1. "Guideline*".ab,ti.
2. Exp Guideline/or exp Practice Guideline/;
3. Clinical practice guideline.ab,ti.
4. 1 or 2 or 3;
5. Back disorder*.tw,kf.;
6. Dorsalgia.tw,kf.;
7. Exp Back Pain/;
8. (Backache* or back ache*).tw,kf.;
9. ((Lumb* or back) adj3 pain).tw,kf.;
10. Coccyx.tw,kf.;
11. Coccydynia.tw,kf.;
12. Sciatic*.tw,kf.;
13. Exp sciatic neuropathy/;
14. Spondylosis.tw,kf.;
15. Lumbago.tw,kf.;
16. Ischialgia.tw,kf.;
17. (Discitis or diskitis).tw,kf.;
18. ((Disc* or disk*) adj3 degenerat*).tw,kf.;

19. ((Disc* or disk*) adj3 prolapse*).tw,kf.;
20. ((Disc* or disk*) adj3 herniat*).tw,kf.
21. Intervertebral Disc/;
22. Exp Intervertebral Disk Displacement/;
23. Exp Intervertebral Disc Degeneration/;
24. Lumbar Vertebrae/;
25. Nerve Compression Syndromes/;
26. Spinal Osteophytosis/;
27. Radiculopathy/;
28. Polyradiculopathy/;
29. Radicul*.tw,kf.;
30. polyradicul*.tw,kf.;
31. arachnoiditis.tw,kf.;
32. exp Spinal Nerve Roots/;
33. or/5–32
34. 4 and 33

PEDro

We searched the following terms applying the filter of "practice guidelines"; Low back pain; sciatic*; radicul*.

National Institute for Health and Clinical Excellence (NICE), New Zealand Guidelines Group (NZGG), International Guideline Library

We performed separated terms in these electronic databases using each of the following combination of terms: low back pain OR sciatic* OR radicul*.

Guideline central

We performed separated searches entering each of the following terms: low back pain; sciatic*; radicular; radiculopathy.

Scholar.google.com

"Guideline" OR "practice guideline" AND "low back pain" OR "sciatic*" OR "radicul*".

Appendix B

Table A1. Recommendation classification.

Recommendation Classification	Definition	Examples of Terminology Used in Guidelines
Should do	Strong recommendations based on strong evidence	Offer, should occur, provide, always do, give, use, apply, (strongly) endorsed, endorsed, must be considered, should be considered, provide, promote and facilitate, important to do, carried out, primary choice, must be done, usually appropriate
Could do	When a recommendation could be 'Considered'	May include, recommend, practitioner might, suggest, may be used, advice, give, may be considered, conditional recommendation, evaluate, can be evaluated, can be tried, may be appropriate, suggested, should be carefully considered
Do not do	When a recommendation should not be offered	Should refrain from, do not routinely offer, not appropriate, should not, do not give, do not use, do not, not indicated, do not start
Uncertain	Inconclusive recommendation	'We are unable to recommend for or against', inconclusive evidence, uncertain, no basis for recommending, Insufficient evidence to make a recommendation for or against

References

1. Hartvigsen, J.; Hancock, M.J.; Kongsted, A.; Louw, Q.; Ferreira, M.L.; Genevay, S.; Hoy, D.; Karppinen, J.; Pransky, G.; Sieper, J.; et al. What low back pain is and why we need to pay attention. *Lancet* **2018**, *391*, 2356–2367. [CrossRef]
2. Freburger, J.K.; Holmes, G.M.; Agans, R.P.; Jackman, A.M.; Darter, J.D.; Wallace, A.S.; Castel, L.D.; Kalsbeek, W.D.; Carey, T.S. The Rising Prevalence of Chronic Low Back Pain. *Arch. Intern. Med.* **2009**, *169*, 251–258. [CrossRef] [PubMed]
3. Hoy, D.; March, L.; Brooks, P.; Woolf, A.; Blyth, F.; Vos, T.; Buchbinder, R. Measuring the global burden of low back pain. *Best Pr. Res. Clin. Rheumatol.* **2010**, *24*, 155–165. [CrossRef] [PubMed]
4. O'Connell, N.E.; Ward, S.P. Low Back Pain: What Have Clinical Guidelines Ever Done for Us? *J. Orthop. Sports Phys. Ther.* **2018**, *48*, 54–57. [CrossRef]
5. Disease, G.B.D.; Injury, I.; Prevalence, C. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: A systematic analy-sis for the Global Burden of Disease Study 2017. *Lancet* **2018**, *392*, 1789–1858.
6. Hoy, D.; Bain, C.; Williams, G.; March, L.; Brooks, P.; Blyth, F.; Woolf, A.; Vos, T.; Buchbinder, R. A systematic review of the global prevalence of low back pain. *Arthritis Rheum.* **2012**, *64*, 2028–2037. [CrossRef] [PubMed]
7. Hill, J.C.; Konstantinou, K.; Egbewale, B.E.; Dunn, K.M.; Lewis, M.; Van Der Windt, D. Clinical Outcomes Among Low Back Pain Consulters With Referred Leg Pain in Primary Care. *Spine* **2011**, *36*, 2168–2175. [CrossRef]
8. Merskey, H.; Bogduk, N. *Classification of Chronic Pain: Descriptions of Chronic Pain Syndromes and Definitions of Pain Terms*; IASP Press: Seattle, WA, USA, 1994.
9. Lin, C.-W.; Verwoerd, A.J.H.; Maher, C.; Verhagen, A.; Pinto, R.; Luijsterburg, P.A.J.; Hancock, M. How is radiating leg pain defined in randomized controlled trials of conservative treatments in primary care? A systematic review. *Eur. J. Pain* **2014**, *18*, 455–464. [CrossRef]
10. Hider, S.L.; Whitehurst, D.G.T.; Thomas, E.; Foster, N.E. Pain location matters: The impact of leg pain on health care use, work disability and quality of life in patients with low back pain. *Eur. Spine J.* **2014**, *24*, 444–451. [CrossRef] [PubMed]
11. Konstantinou, K.; Hider, S.L.; Jordan, J.L.; Lewis, M.; Dunn, K.M.; Hay, E.M. The impact of low back-related leg pain on outcomes as compared with low back pain alone: A systematic review of the literature. *Clin. J. Pain.* **2013**, *29*, 644–654. [CrossRef]
12. van Tulder, M.W.; Koes, B.W.; Bouter, L.M. A cost-of-illness study of back pain in The Netherlands. *Pain* **1995**, *62*, 233–240. [CrossRef]
13. Oliveira, C.B.; Maher, C.G.; Pinto, R.Z.; Traeger, A.C.; Lin, C.-W.C.; Chenot, J.-F.; Van Tulder, M.; Koes, B.W. Clinical practice guidelines for the management of non-specific low back pain in primary care: An updated overview. *Eur. Spine J.* **2018**, *27*, 2791–2803. [CrossRef]
14. Bouwmeester, W.; van Enst, A.; van Tulder, M. Quality of low back pain guidelines improved. *Spine* **2009**, *34*, 2562–2567. [CrossRef] [PubMed]
15. Koes, B.W.; van Tulder, M.; Lin, C.W.; Macedo, L.G.; McAuley, J.; Maher, C. An updated overview of clinical guide-lines for the management of non-specific low back pain in primary care. *Eur. Spine J.* **2010**, *19*, 2075–2094. [CrossRef] [PubMed]
16. Koes, B.W.; van Tulder, M.W.; Ostelo, R.; Kim Burton, A.; Waddell, G. Clinical guidelines for the management of low back pain in primary care: An international comparison. *Spine* **2001**, *26*, 2504–2513. [CrossRef]
17. Pillastrini, P.; Gardenghi, I.; Bonetti, F.; Capra, F.; Guccione, A.; Mugnai, R.; Violante, F.S. An updated overview of clinical guidelines for chronic low back pain management in primary care. *Jt. Bone Spine* **2012**, *79*, 176–185. [CrossRef]
18. Meroni, R.; Piscitelli, D.; Ravasio, C.; Vanti, C.; Bertozzi, L.; de Vito, G.; Perin, C.; Guccione, A.A.; Cerri, C.G.; Pillastrini, P. Evidence for managing chronic low back pain in primary care: A re-view of recommendations from high-quality clinical practice guidelines. *Disabil. Rehabil.* **2021**, *43*, 1029–1043. [CrossRef]
19. The AGREE Next Steps Consortium. Appraisal of Guidelines for Research & Evaluation (AGREE) II Manual. 2009. Available online: http://www.agreetrust.org/wp-content/uploads/2013/06/AGREE_II_Users_Manual_and_23-item_Instrument_ENGLISH.pdf (accessed on 14 May 2009).
20. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med.* **2009**, *6*, e1000097. [CrossRef] [PubMed]
21. Ripple, A.S. Expert googling: Best practices and advanced strategies for using google in health sciences li-braries. *Med. Ref. Serv. Q.* **2006**, *25*, 97–107. [CrossRef]
22. Brouwers, M.C.; Kho, M.; Browman, G.P.; Burgers, J.S.; Cluzeau, F.; Feder, G.; Fervers, B.; Graham, I.D.; Grimshaw, J.; Hanna, S.E.; et al. AGREE II: Advancing guideline development, reporting and evaluation in health care. *Can. Med. Assoc. J.* **2010**, *182*, E839–E842. [CrossRef]
23. Uzeloto, J.S.; Moseley, A.M.; Elkins, M.R.; Franco, M.R.; Pinto, R.Z.; Freire, A.P.C.F.; Ramos, E.M.C. The quality of clinical practice guidelines for chronic respiratory diseases and the reliability of the AGREE II: An observational study. *Physiotherapy* **2017**, *103*, 439–445. [CrossRef] [PubMed]
24. Zens, T.J.; Cartmill, R.; Muldowney, B.L.; Fernandes-Taylor, S.; Nichol, P.; Kohler, J.E. Practice Variation in Umbili-cal Hernia Repair Demonstrates a Need for Best Practice Guidelines. *J. Pediatr.* **2019**, *206*, 172–177. [CrossRef] [PubMed]
25. Cicchetti, D.V. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol. Assess.* **1994**, *6*, 284. [CrossRef]

26. Yan, J.; Min, J.; Zhou, B. Diagnosis of pheochromocytoma: A clinical practice guideline appraisal using AGREE II instrument. *J. Eval. Clin. Pract.* **2012**, *19*, 626–632. [CrossRef] [PubMed]
27. Lin, I.; Wiles, L.; Waller, R.; Goucke, R.; Nagree, Y.; Gibberd, M.; Straker, L.; Maher, C.G.; O'Sullivan, P.P.B. What does best practice care for musculoskeletal pain look like? Eleven consistent recommendations from high-quality clinical practice guidelines: Systematic review. *Br. J. Sports Med.* **2020**, *54*, 79–86. [CrossRef] [PubMed]
28. American College of Occupational and Environmental Medicine. *Low Back Disorders*; Reed Group Ltd.: Broomfield, CO, USA, 2016.
29. Chou, R.; Qaseem, A.; Owens, D.K.; Shekelle, P. Diagnostic Imaging for Low Back Pain: Advice for High-Value Health Care From the American College of Physicians. *Ann. Intern. Med.* **2011**, *154*, 181–189. [CrossRef] [PubMed]
30. Chou, R.; Qaseem, A.; Snow, V.; Casey, D.; Cross, J.T., Jr.; Shekelle, P.; Owens, D.K. Diagnosis and treatment of low back pain: A joint clinical practice guideline from the American College of Physicians and the American Pain Society. *Ann. Intern. Med.* **2007**, *147*, 478–491. [CrossRef]
31. Patel, N.D.; Broderick, D.F.; Burns, J.; Deshmukh, T.K.; Fries, I.B.; Harvey, H.B.; Holly, L.; Hunt, C.H.; Jagadeesan, B.D.; Kennedy, T.A.; et al. ACR Appropriateness Criteria Low Back Pain. *J. Am. Coll. Radiol.* **2016**, *13*, 1069–1078. [CrossRef]
32. Chou, R.; Loeser, J.D.; Owens, D.K.; Rosenquist, R.W.; Atlas, S.J.; Baisden, J.; Carragee, E.J.; Grabois, M.; Murphy, D.R.; Resnick, D.K.; et al. Interventional Therapies, Surgery, and Interdisciplinary Rehabilitation for Low Back Pain. *Spine* **2009**, *34*, 1066–1077. [CrossRef]
33. Delitto, A.; George, S.Z.; Van Dillen, L.; Whitman, J.M.; Sowa, G.; Shekelle, P.; Denninger, T.R.; Godges, J.J. Low back pain. *J. Orthop. Sports Phys. Ther.* **2012**, *42*, A1–A57. [CrossRef]
34. Manchikanti, L.; Abdi, S.; Atluri, S.; Benyamin, R.M.; Boswell, M.V.; Buenaventura, R.M.; Bryce, D.A.; Burks, P.A.; Caraway, D.L.; Calodney, A.K.; et al. An update of comprehensive evidence-based guidelines for interven-tional techniques in chronic spinal pain. Part II: Guidance and recommendations. *Pain Physician* **2013**, *16*, S49–S283. [PubMed]
35. Bussières, A.E.; Stewart, G.; Al-Zoubi, F.; Decina, P.; Descarreaux, M.; Haskett, D.; Hincapié, C.; Pagé, I.; Passmore, S.; Srbely, J.; et al. Spinal manipulative therapy and other conservative treatments for low back pain: A guideline from the Canadian Chiropractic Guideline Initiative [with consumer summary]. *J. Manip. Physiol. Ther.* **2018**, *41*, 265–293. [CrossRef] [PubMed]
36. Colorado Division of Workers Compensation USDoH, Human Services PHSAfHR, Quality. Low Back Pain–Medical Treatment Guidelines. 2014. Available online: https://www.colorado.gov/pacific/sites/default/files/Rule_XVII_Exhibit_A.pdf (accessed on 10 March 2014).
37. Airaksinen, O.; Brox, J.I.; Cedraschi, C.; Hildebrandt, J.; Klaber-Moffett, J.; Kovacs, F.; Mannion, A.F.; Reis, S.; Staal, J.B.; Ursin, H.; et al. European guidelines for the management of chronic non-specific low back pain [with consumer summary]. *Eur. Spine J.* **2006**, *15*, S192–S300. [CrossRef] [PubMed]
38. Thorson, D.; Campbell, R.; Massey, M.; Mueller, B.; McCathie, B.; Richards, H.; Peterson, S.; Kramer, C.; Ginkel, T.; Dvorkin, J.; et al. *Adult Acute and Subacute Low Back Pain*, 16th ed.; Institute for Clinical Systems Improvement (ICSI): Bloomington, MN, USA, 2018.
39. Negrini, S.; Giovannoni, S.; Minozzi, S.; Barneschi, G.; Bonaiuti, D.; Bussotti, A.; D'Arienzo, M.; Di Lorenzo, N.; Mannoni, A.; Mattioli, S.; et al. Diagnostic therapeutic flow-charts for low back pain patients: The Italian clinical guidelines. *Eur. Med.* **2006**, *42*, 151–170.
40. van Wambeke, P.; Desomer, A.; Ailliet, L.; Berquin, A.; Demoulin, C.; Depreitere, B.; Dewachter, J.; Dolphens, M.; Forget, P.; Fraselle, V.; et al. *Low Back Pain and Radicular Pain: Evaluation and Management*; Good Clinical Practice (GCP), KCE Reports 287. D/2017/10.273/36; Belgian Health Care Knowledge Centre (KCE): Brussels, Belgium, 2017.
41. Jun, J.H.; Cha, Y.; Lee, J.A.; Choi, J.; Choi, T.-Y.; Park, W.; Chung, W.; Shin, B.-C.; Lee, M.S. Korean medicine clinical practice guideline for lumbar herniated intervertebral disc in adults: An evidence based approach. *Eur. J. Integr. Med.* **2017**, *9*, 18–26. [CrossRef]
42. Kreiner, D.S.; Hwang, S.W.; Easa, J.E.; Resnick, D.K.; Baisden, J.L.; Bess, S.; Cho, C.H.; DePalma, M.J.; Dougherty, P.; Fernand, R.; et al. An evidence-based clinical guideline for the diagnosis and treatment of lumbar disc herniation with radiculopathy. *Spine J.* **2014**, *14*, 180–191. [CrossRef] [PubMed]
43. Danish Health Authority. National Clinical Guideline for the Nonsurgical Treatment of Recent Onset Lumbar Nerve Root Compression (Lumbar Radiculopathy). 2016. Available online: https://www.sst.dk/da/udgivelser/2016/~{}/media/B9D3E068233A4F7E95F7A1492EBC4484.ashx (accessed on 6 June 2016).
44. Laerum, E.; Dullerud, R.; Kirkesola, G.; Nygaard, Ø.P.; Skouen, J.S.; Stig, L.-C.; Werner, E. The Norwegian Back Pain Network. Acute Low Back Pain: Interdisciplinary Clinical Guidelines. 2007. Available online: https://chiro.org/Low_Back_Pain/FULL/Acute_Low_Back_Pain_Interdisciplinary.pdf (accessed on 20 June 2016).
45. Nederlands Huisartsen Genootschap Workgroup. Standaard Lumbosacraal Radiculair Syndroom. 2015. Available online: https://www.henw.org/artikelen/nhg-standaard-lumbosacraal-radiculair-syndroom-tweede-herziening (accessed on 6 September 2018).
46. UK NGC. Evidence Review for Pharmacological Management of Sciatica: Low Back Pain and Sciatica in over 16s: Assessment and Management: Evidence Review A. 2020. Available online: https://pubmed.ncbi.nlm.nih.gov/33090751/ (accessed on 20 March 2021).

47. Philippine Academy of Rehabilitation Medicine. Clinical Practice Guidelines on the Diagnosis and Management of Low Back. 2011. Available online: http://www.ncbi.nlm.nih.gov/books/nbk562934/ (accessed on 7 February 2021).
48. Low Back Pain Working Group Toward Optimized Practice. Evidence-Informed Primary Care Management of Low Back Pain. 2017. Available online: https://parm.org.ph/pdf/lbp.pdf (accessed on 20 January 2021).
49. The University of Michigan Health System. Acute Low Back Pain. 2010. Available online: http://www.med.umich.edu/1info/FHP/practiceguides/back/back.pdf (accessed on 9 October 2020).
50. Treatment of Low Back Pain Work Group Department of Veterans Affairs DoD. VA/DoD clinical practice guideline for diagnosis and treatment of low back pain (version 2.0). *J. Gen. Intern. Med.* **2019**, *34*, 2620–2629. [CrossRef]
51. Doniselli, F.M.; Zanardo, M.; Manfrè, L.; Papini, G.D.E.; Rovira, A.; Sardanelli, F.; Sconfienza, L.C.; Arana, E. A critical appraisal of the quality of low back pain practice guide-lines using the AGREE II tool and comparison with previous evaluations: A EuroAIM initiative. *Eur. Spine J.* **2018**, *27*, 2781–2790. [CrossRef] [PubMed]
52. Schünemann, H.J.; Wiercioch, W.; Brozek, J.; Etxeandia-Ikobaltzeta, I.; Mustafa, R.A.; Manja, V.; Brignardello-Petersen, R.; Neumann, I.; Falavigna, M.; Alhazzani, W.; et al. GRADE Evidence to Decision (EtD) frameworks for adoption, adaptation, and de novo development of trustworthy recommendations: GRADE-ADOLOPMENT. *J. Clin. Epidemiol.* **2017**, *81*, 101–110. [CrossRef] [PubMed]
53. Harstall, C.; Taenzer, P.; Angus, D.K.; Moga, C.; Schuller, T.; Scott, N.A. Creating a multidisciplinary low back pain guideline: Anatomy of a guideline adaptation process. *J. Eval. Clin. Pract.* **2011**, *17*, 693–704. [CrossRef] [PubMed]
54. van der Windt, D.A.; Simons, E.; Riphagen, I.I.; Ammendolia, C.; Verhagen, A.P.; Laslett, M.; Devillé, W.; Deyo, R.A.; Bouter, L.M.; de Vet, H.C.; et al. Physical examination for lumbar radiculopathy due to disc herniation in patients with low-back pain. *Cochrane Database Syst. Rev.* **2010**, CD007431. [CrossRef] [PubMed]
55. Majlesi, J.; Togay, H.; Ünalan, H.; Toprak, S. The Sensitivity and Specificity of the Slump and the Straight Leg Raising Tests in Patients With Lumbar Disc Herniation. *JCR J. Clin. Rheumatol.* **2008**, *14*, 87–91. [CrossRef]
56. Ostelo, R.W. Physiotherapy management of sciatica. *J. Physiother.* **2020**, *66*, 83–88. [CrossRef]
57. Pinto, R.Z.; Verwoerd, A.J.H.; Koes, B.W. Which pain medications are effective for sciatica (radicular leg pain)? *BMJ* **2017**, *359*, j4248. [CrossRef] [PubMed]
58. Williams, C.M.; Maher, C.G.; Latimer, J.; McLachlan, A.; Hancock, M.J.; O Day, R.; Lin, C.-W.C. Efficacy of paracetamol for acute low-back pain: A double-blind, randomised controlled trial. *Lancet* **2014**, *384*, 1586–1596. [CrossRef]
59. Rasmussen-Barr, E.; Held, U.; Grooten, W.J.; Roelofs, P.D.; Koes, B.W.; Van Tulder, M.W.; Wertli, M. Non-steroidal anti-inflammatory drugs for sciatica. *Cochrane Database Syst. Rev.* **2016**, *10*, CD012382. [CrossRef] [PubMed]
60. Oliveira, C.B.; Maher, C.G.; Ferreira, M.L.; Hancock, M.J.; Oliveira, V.C.; McLachlan, A.J.; Koes, B.W.; Ferreira, P.H.; Cohen, S.P.; Pinto, R.Z. Epidural corticosteroid injections for lumbosacral radicular pain. *Cochrane Database Syst. Rev.* **2020**, CD013577. [CrossRef] [PubMed]
61. Shallwani, S.M.; King, J.; Thomas, R.; Thevenot, O.; De Angelis, G.; Aburub, A.S.; Brosseau, L. Methodological quality of clinical practice guidelines with physical ac-tivity recommendations for people diagnosed with cancer: A systematic critical appraisal using the AGREE II tool. *PLoS ONE* **2019**, *14*, e0214846. [CrossRef]
62. Smith, C.A.M.; Toupin-April, K.; Jutai, J.W.; Duffy, C.M.; Rahman, P.; Cavallo, S.; Brosseau, L. A Systematic Critical Appraisal of Clinical Practice Guidelines in Juvenile Idiopathic Arthritis Using the Appraisal of Guidelines for Research and Evaluation II (AGREE II) Instrument. *PLoS ONE* **2015**, *10*, e0137180. [CrossRef]
63. Brosseau, L.; Rahman, P.; Toupin-April, K.; Poitras, S.; King, J.; De Angelis, G.; Loew, L.; Casimiro, L.; Paterson, G.; McEwan, J. A systematic critical appraisal for non-pharmacological management of osteoarthritis using the appraisal of guidelines research and evalua-tion II instrument. *PLoS ONE* **2014**, *9*, e82986.

Journal of
Clinical Medicine

MDPI

Article

Differences in Knee Shape between ACL Injured and Non-Injured: A Matched Case-Control Study of 168 Patients

Koen S.R. van Kuijk [1,2,*,†], Vincent Eggerding [1,†], Max Reijman [1], Belle L. van Meer [3], Sita M.A. Bierma-Zeinstra [1,4], Ewoud van Arkel [5], Jan H. Waarsing [1] and Duncan E. Meuffels [1]

1 Department of Orthopedic Surgery, Erasmus MC University Medical Centre, 3015 CN Rotterdam, The Netherlands; v.eggerding@erasmusmc.nl (V.E.); m.reijman@erasmusmc.nl (M.R.); s.bierma-zeinstra@erasmusmc.nl (S.M.A.B.-Z.); e.waarsing@erasmusmc.nl (J.H.W.); d.meuffels@erasmusmc.nl (D.E.M.)
2 Department of Radiology, Albert Schweitzer Hospital, 3318 AT Dordrecht, The Netherlands
3 Department of Sports Medicine, St Antonius Hospital, 3543 AZ Utrecht, The Netherlands; bellevanmeer@hotmail.com
4 Department of General Practice, Erasmus MC University Medical Centre, 3015 CN Rotterdam, The Netherlands
5 Department of Orthopaedic Surgery, Haaglanden Medical Centre Haaglanden, 2512 VA The Hague, The Netherlands; e.van.arkel@haaglandenmc.nl
* Correspondence: koenvankuijk@live.nl
† Both Koen S.R. van Kuijk and Vincent Eggerding are first authors of this paper.

check for updates

Citation: van Kuijk, K.S.R.; Eggerding, V.; Reijman, M.; van Meer, B.L.; Bierma-Zeinstra, S.M.A.; van Arkel, E.; Waarsing, J.H.; Meuffels, D.E. Differences in Knee Shape between ACL Injured and Non-Injured: A Matched Case-Control Study of 168 Patients. *J. Clin. Med.* **2021**, *10*, 968. https://doi.org/10.3390/jcm10050968

Academic Editors: Zsolt Balogh and Jos Runhaar

Received: 12 January 2021
Accepted: 23 February 2021
Published: 2 March 2021

Abstract: Objective: Anterior cruciate ligament (ACL) injury prevention programs could be more effective if we could select patients at risk for sustaining an ACL rupture. The purpose of this study is to identify radiographic shape variants of the knee between patients with and patients without an ACL rupture. Methods: We compared the lateral and Rosenberg view X-rays of 168 prospectively followed patients with a ruptured ACL to a control group with intact ACLs, matched for gender, after knee trauma. We used statistical shape modeling software to examine knee shape and find differences in shape variants between both groups. Results: In the Rosenberg view X-rays, we found five shape variants to be significantly different between patients with an ACL rupture and patients with an intact ACL but with knee trauma. Overall, patients who had ruptured their ACL had smaller, flatter intercondylar notches, a lower lateral tibia plateau, a lower medial spike of the eminence, and a smaller tibial eminence compared to control patients. Conclusion: Patients with an ACL rupture have smaller intercondylar notches and smaller tibial eminences in comparison to patients with an intact ACL after knee trauma.

Keywords: ACL; femoral intercondylar notch; knee anatomy; ACL prevention; ACL risk factors

1. Introduction

A rupture of the anterior cruciate ligament (ACL) is a common, usually sports-related injury. The annual incidence varies between 5 and 8 per 10,000 persons in the Western population [1–4]. Rupture of the ACL has immediate consequences resulting in swelling of the knee and pain, but also long-term consequences, as there is an almost fourfold risk to progress to moderate or severe radiological osteoarthritis after ten years [5]. Furthermore, in the young population, ACL rupture has a direct impact on sport participation. It has been found, for instance, that after ACL reconstruction, 82% of the patients returned to sport participation; however, only 63% returned to their preinjury sport level [6,7]. Amongst young patients who return to their pre-trauma sports activity, the prevalence of a re-rupture of their ACL may be as high as 30% [8,9]. Additionally, reports show that around 7% of patients need revision ACL surgery and around 3.4% of patients have ACL reconstructions on the contralateral side [10].

This has led to a rise in interest in the mechanism of ACL rupture, in risk factors, prevention of ACL rupture and secondary ACL injury. Neuromuscular and proprioceptive prevention programs have been demonstrated to significantly reduce the prevalence of ACL ruptures in young athletes by approximately 50% [11–14]. However, these prevention programs can be more efficient if they focus on athletes who are at increased risk of sustaining an ACL rupture. Therefore, it is essential to understand the mechanisms that lead to ACL rupture and to identify individuals with an increased risk of ACL rupture.

There is a relationship between shape variants of the knee and the need for reconstruction of the ACL after rupture [15]. This has encouraged us to investigate the relationship between shape and rupture of the ACL more profoundly. Risk factors for ACL rupture can be categorized into anatomical, hormonal, neuro-mechanical, and environmental. In the present study, we focused on osseous anatomical risk factors; anatomical risk factors have previously been studied with a focus on selected aspects of the anatomical properties of the knee. Anatomical factors that have previously been reported to be related to the risk of ACL rupture are increased tibial slope, decreased femoral notch size, and smaller ACL size [16,17]. With the use of statistical shape modeling (SSM), a hypothesis generating-methodology that identifies independent shape variants, we can quantitatively describe the complete morphology of a bone or joint. SSM reproduces all variation in shape that is present in the studied population. With the use of SSM, we can identify new shape variants of the knee that are clinically relevant in relation to an ACL rupture. Furthermore, it enables us to objectively review shape variants that have been investigated before. Although not all clinicians will have a program such as SSM in use, the results of this study can be used in daily practice and can help doctors in selecting patients at greater risk for sustaining an ACL rupture.

SSM has been used earlier by our group to determine whether certain shape aspects are correlated to clinical outcomes after ACL rupture [18]. We found that operatively treated patients with good subjective outcomes had a smaller intercondylar notch and a lower width intercondylar eminence, as evaluated by The International Knee Documentation Committee (IKDC) questionnaire, compared to patients with worse outcome. Non-operatively treated patients with good subjective outcomes had a more pyramidal shaped intercondylar notch.

The purpose of this study is to find radiographic shape variants of the knee between patients with and patients without an ACL rupture, which can be used in daily practice to help select patients with a greater risk for sustaining an ACL rupture.

2. Experimental Section

2.1. Cases

We included patients with a ruptured ACL from two previous series: the KNALL [19] and the CAS-ACL study [20].

The KNee osteoArthritis anterior cruciate Ligament Lesion (KNALL) study is a prospective observational follow-up study of 154 patients with a recent ACL rupture, who were treated operatively or non-operatively. Patients were selected from January 2009 to November 2010, and there was a two-year follow-up period. Physical examination and MRI confirmed ACL rupture. Patients were included from three collaborating hospitals.

The CAS-ACL study is a double-blinded randomized controlled trial of 100 patients who underwent ACL reconstruction. In this study, computer-assisted ACL reconstruction was compared to conventional ACL reconstruction [20,21]. The inclusion period was from January 2007 to November 2009 with a two-year follow-up period. Of the 254 patients included in the two studies, 183 had both Rosenberg view and lateral view radiographs and were enrolled in the present study.

All patients (both ACL injured as healthy controls) included had a Kellgren and Lawrence grade 0–1 at presentation (no radiological signs of osteoarthritis). Our medical ethics committee (MEC-2006-223 and MEC-2008-068) approved both studies. For the use of

the data of control patients, the medical ethics committee ruled that no specific approval was required (MEC-2017-422).

2.2. Controls

The control group consisted of patients identified retrospectively from the hospital records. They had consulted a trauma or orthopedic surgeon because of a knee trauma (median of 3 months and a range of 1–60 months between trauma and X-ray) with confirmed intact ACL by MRI and/or arthroscopy. Hospital records from January 2003 until July 2013 were searched. Patients were selected for the control group if they had both standard lateral view and Rosenberg view radiographs at the time of the first consult, were practicing sports before the injury, and had a Kellgren and Lawrence grade 0–1 at presentation (no radiological signs of osteoarthritis). Control patients and cases were matched for gender. For age, our patients were matched with a control patient older in age. We chose older control subjects to ensure that the older controls were exposed to pivoting sports for a longer period. This way, they were more sufficiently at risk for sustaining an ACL rupture. Of all patients found in the database, 168 control patients were matched to 168 patients with a ruptured ACL. See Figure 1. We were unable to match all patients due to younger age, since we wanted to only include older control patients. Fifteen control patients were younger than the matches from the ACL ruptured group.

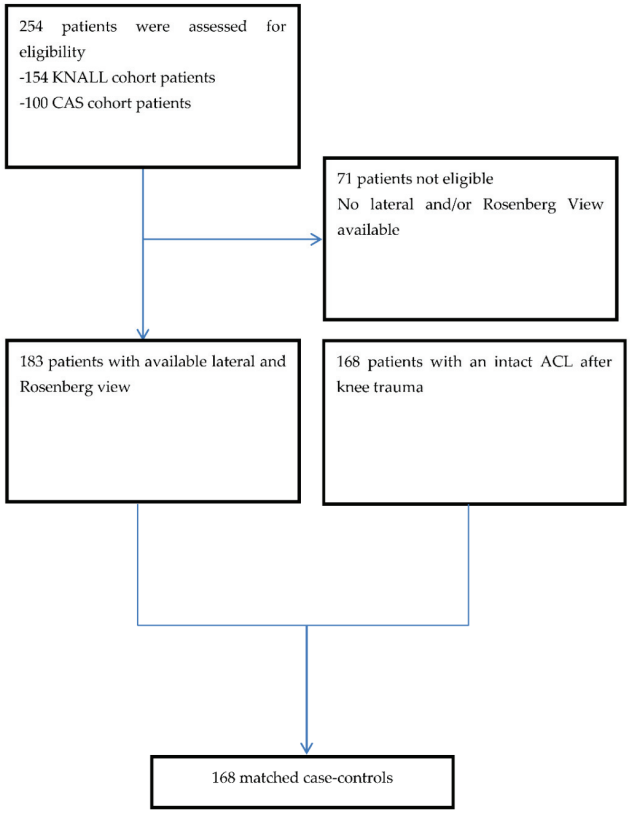

Figure 1. Flowchart of the selected patients included in the study. KNee osteoArthritis anterior cruciate Ligament Lesion (KNALL); Computer Assisted Surgery (CAS).

2.3. X-rays and Statistical Shape Modeling

We performed the radiological measurements using standard lateral view X-rays and Rosenberg view X-rays. The Rosenberg view is a weight-bearing postero-anterior radiograph taken at 45° flexion of the knee [22]. We chose to include the Rosenberg view X-rays because it gives a better view of the intercondylar notch and gives better insights in the shape of the femur.

With statistical shape modeling (SSM) [23], it is possible to quantify all shape aspects of the knee joint in the radiographs. This method is unique because it dissects nearly all possible shape variations into a limited number of objectively quantitative measures that each describe a certain shape variant. SSM has been used in studies of a possible association between knee shape variants and osteoarthritis [24]. In the radiographs, we outlined the distal femur, the proximal tibia, and fibula (ASM tool kit, Manchester University, Manchester, UK).

For the lateral view X-rays, the femur and tibia were outlined by 60 landmark points on the bones. For the Rosenberg view, 25 landmark points were necessary to completely outline the bones. Each point was placed in the same location in each image, as precisely as possible, to allow a comparison between shapes. For the exact placement of each landmark point, see the addendum. Statistical shape modeling transforms the set of points into a statistical shape model, which comprises a number of shape variants that together explain 95% of the variation in the shape of the individual knee of the study population. SSM represents relative variation in shape, independent of differences in the size of the joint. In this way, the method corrects errors caused by variation in magnification or in the size of the patient's knee.

Intraobserver reliability was established by randomly selecting 25 Rosenberg view X-rays of patients with a ruptured ACL and 25 Rosenberg view X-rays of patients with an intact ACL, which were outlined a second time after 2 weeks.

The description of which shape aspects a variant represents was determined at a consensus meeting. At this consensus meeting, an orthopedic surgeon, an expert on SSM, the first authors, and the principal investigator determined the different shape variants.

2.4. Statistical Analysis

We used logistic regression analysis to study the association between each shape variant and whether or not patients had a ruptured ACL. As the dependent variable, we used whether or not a patient had an ACL rupture (yes or no), and as independent variables, we selected the different variants. We applied Bonferroni correction for multiple testing. We investigated if there was a significant effect of the X-ray protocol on knee shape by comparing the shape models of the X-rays taken in the three participating hospitals. Furthermore, we analyzed if correction for age changed the outcomes. All statistical analyses were performed with IBM SPSS Statistics for Windows (Version 20.0. IBM Corp., Armonk, NY, USA).

3. Results

3.1. Patients

The study population consisted of two groups of 168 patients; each group consisted of 119 males and 49 females. The mean age of the 168 patients after ACL rupture was 31 (±standard deviation (SD) 7.4) years and of the control group 38 (±SD 12) years (Table 1). The diagnoses of the included control patients can be found in Table 1, including additional injuries of the ACL ruptured patients. The mean time between trauma and radiograph was 1.0 months for the ACL injured and 6.9 months for the control group.

Table 1. Baseline demographic variables.

	ACL Injured (n = 168)	Control Group (n = 168)
Age, year	31 ± 7.4	38 ± 12.0
BMI, kg/m^2	24.5 ± 3.4	24.7 ± 3.2
Female n (%)	49 (29.1)	49 (29.1)
Mean time in months between trauma and radiograph	1.0	6.9
Alternative/additional diagnosis, n (%)		
Medial Meniscus tear	10 (6)	57 (33.9)
Lateral meniscus tear	12 (7)	32 (19)
Cartilage lesion	60 (35)	15 (8.9)
Bone contusion	50 (30)	11 (6.5)
Collateral ligament lesion	0 (0)	7 (4.2)
No intra-articular lesions	0 (0)	46 (27.4)

Data are expressed as mean ± standard deviation or as n (%). BMI, body mass index. Anterior Cruciate Ligament (ACL).

3.2. SSM

SSM produced 30 variants for the Rosenberg view and 24 variants for the lateral view X-rays. After we applied Bonferroni correction for multiple testing, we considered a p-value of 0.0017 for the Rosenberg view (0.05/30 = 0.00167) and a p-value of 0.0021 for the lateral view (0.05/24 = 0.0021) as statistically significant.

In the Rosenberg view, five variants were significantly associated with rupture of the ACL (see Table 2). For the lateral view X-rays, none of the variants were statistically significantly associated with rupture of the ACL. For every increase of 1 SD, the OR is given, meaning that if a patient scores 1 SD in a specific variant, the given OR is the odds ratio for sustaining an ACL rupture compared to a patient who scores the mean (0 SD).

Table 2. Shape variants associated with ACL rupture.

	Odds Ratio	95% C.I.	Sig.
Variant 1	2.2	(1.7–2.8)	0.001
Variant 3	1.8	(1.4–2.3)	0.001
Variant 6	2.1	(1.6–2.7)	0.001
Variant 10	1.5	(1.2–1.8)	0.001
Variant 17	1.4	(1.1–1.8)	0.0015

We analyzed whether the protocols of the X-rays differed in the period of time of which the X-rays were taken. We did not find a significant difference between the three hospitals, nor did we find a significant difference in time. Correction for age did not alter the outcomes; therefore, we did not correct for age.

The intraobserver (ICC) was considered good to excellent with a range of 0.48–0.97 and 89% above 0.7.

3.3. Description of the Variants

Here, we present the description, defined at the consensus meeting, of the variants significantly associated with an ACL rupture. In Figure 2, we present the graphics of each variant. On the outside, the +2SD and -2SD variants are shown; in the middle, we present an overlay. Higher variants describe more subtle shape aspects, e.g., the variation in shape represented in shape variant 17 is much more subtle than the variation represented by shape variant 1.

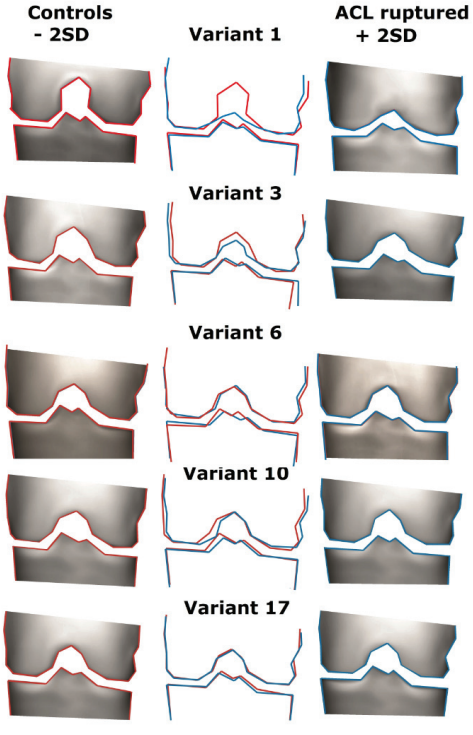

Figure 2. Graphic outcomes of statistic shape modeling: five variants that are significantly different for patients with intact and ruptured ACL. On the left and right sides are the two extremes (+/− 2.5 SD); in the middle is the overlay of both sides. SD = Standard.

Variant 1

Variant 1 describes a variation in the height of the intercondylar notch. Positive values represent a more flattened intercondylar notch. Patients with an ACL rupture had flatter intercondylar notches than control patients.

Variant 3

Variant 3 shows a variation in the width and height of the intercondylar notch. Positive values represent a smaller intercondylar notch. Patients with an ACL rupture had smaller intercondylar notches than control patients.

Variant 6

Variant 6 represents the size of the footprint of the ACL on the tibial eminence. Positive values represent a smaller, flatter tibial eminence. Patients with an ACL rupture had a smaller tibial eminence than control patients.

Variant 10

Variant 10 outlines the footprint of the ACL on the tibia, the width of the tibial eminence, and the width of the intercondylar notch. Positive values represent a smaller tibial eminence and a smaller intercondylar notch. Patients with ACL rupture had a smaller tibial eminence and a smaller intercondylar notch.

Variant 17

Variant 17 depicts a very subtle difference. Positive values represent a lower height of the lateral tibial plateau and the lower medial spike of the tibial eminence. Patients

with an ACL rupture had a lower lateral tibia plateau and a lower medial spike of the intercondylar eminence.

4. Discussion

The most important finding of the present study is that aspects of bony morphology in the Rosenberg view X-ray of the knee joint were different between patients with a ruptured ACL and a matched control group. Our findings indicate that a smaller, flatter intercondylar notch; a lower lateral tibia plateau; a lower medial spike of the eminence; and a smaller tibial eminence were more common in patients who ruptured their ACL compared to control patients. Lower body strength exercises (for example, Nordic hamstring, lunges, and heel-calf raise) are not performed by all (professional) athletes but have been proven to reduce the risk of ACL rupture [25]. If we can identify patients at higher risk for ACL injury, injury prevention programs might be even more effective, although this should be confirmed in a different study. Our results could, for example, be used during sports medical screening: most professional athletes already undergo X-rays of the knee in the medical screening process.

The results of our study are consistent with studies in the past, which have also found the notch width index and femoral notch size to be related to ACL rupture [26]. However, these previously conducted studies were primarily focused on anterior–posterior X-rays, while we used the Rosenberg view X-rays. The study of van Diek et al. [27] found no differences in morphology between patients with an ACL rupture and a control group in measurements with MRI. However, another MRI study performed by Whitney et al. [28] found a decreased femoral notch width to be related to ACL rupture. This was also a case-control study. A smaller femoral notch and smaller tibial eminence are related to smaller ACL size [29,30]. It is plausible that a smaller ACL could be less strong compared to a larger-sized ACL. The ACL is the main structure to prevent the bony relatively unstable lateral compartment from rotatory dislocation, i.e., rotation anterior of the tibia relative to the femur. The finding of a lower lateral tibia plateau in ACL patients could inspire the theory that these patients have even worse bony stability regarding the lateral compartment, which could be a risk factor for ACL injury.

In the lateral view X-rays, we did not find an association between shape variants and ACL rupture. Earlier, it has been demonstrated that the femoral condyle configuration [31] and the posterior tibial slope (PTS) [32–34] are related to increased stress on the ACL, but it is not known if this is connected to a higher risk of ACL rupture.

With the results of this study, we can identify individuals with certain shape variants of the knee, who are at greater risk for sustaining an ACL rupture. Because we did not use a predefined hypothesis, the found that risk factors are truly objective, whereas other researchers used a predefined hypothesis, potentially excluding numerous risk factors. These results can be used in daily practice, without the use of our program. Almost all clinicians can view the X-rays of their patients, and thus view the shape of the tibial eminence and the shape of the intercondylar notch.

Screening programs for professional athletes could focus on the intercondylar notch and tibial eminence as risk factors. With our results, screening programs could focus on the shape variants found and include patients with a higher risk of sustaining an ACL in their training programs, potentially making them more effective. In the past, research stated that the tibial slope could also be a determinant for sustaining an ACL rupture. With our hypothesis-free method, we did not find similar results. Excluding potential risk factors is also important because research should focus on risk factors that are more likely to be associated with sustaining ACL rupture

Although the odds ratios are relatively small, all the provided variants show a significant relationship to an ACL rupture, with odds ratios comparable to that of other studies investigating anatomical variants of the knee [16,27]. Further research could focus primarily on the shape variants found in this study and determine if these results can be

reproduced. Furthermore, prospective studies should be performed to see if, with these risk factors, the prevalence of ACL ruptures could be reduced.

We understand that in the current literature there are already studies using 3D reconstructions of MRIs of the knee. However, our goal was not to confirm previously conducted research but to objectively describe the shape variants of the knee that contribute to the risk for sustaining an ACL rupture. If we had used MRI, we would have had to make a prior hypothesis. Furthermore, 3D reconstructions of the knee are not used in daily practice by every clinician. With the use of X-rays, we are confident that more clinicians can use these results in their practice, without the use of complicated extra software.

A drawback of SSM is that the shape represented by each variant needs to be reviewed personally (which we did in the consensus meeting). SSM does not provide a measurable cut off point; this should be determined in follow-up studies.

When we examined the different variants in our consensus meeting, we viewed 3D, moving animations of the shape variants; in this animation, the differences are more clearly visible than in 2D images.

We did not perform a power analysis before conducting the study, because we did not know how many shape variants would be found beforehand. Therefore, we used the Bonferroni correction for multiple testing. Although we understand that these are not the same thing, we are confident the method provided is valid and reproducible.

The strength of our study is the use of a large study population of 336 patients, who all practiced sports. Among the advantages of SSM is that the programs scale all differences in the size of the joints, thus reducing the variation in magnification and reducing measurement errors. A limitation of our study is the use of older control patients. We used older control patients to ensure that they were sufficiently exposed to rotational trauma. Older patients potentially have more degenerative changes. To ensure that this did not affect our outcomes, we chose patients with K&L scores of <1 (no signs of osteoarthritis) for both the patients with an ACL rupture and the control patients. Another limitation could be the use of a hypothesis-free program, which makes it, for some clinicians, harder to use in daily practice. However, we are confident that the shape variants found are usable in daily practice. You do not need a program to assess the wideness of the intercondylar notch in a Rosenberg view X-ray. Clinicians and radiologists with some experience in knee X-rays can easily interpret the findings in our study.

We used Rosenberg and standard lateral view X-rays for our analyses. In 1997, Shelbourne et al. [35] advocated the use of Rosenberg view X-rays because of the standardized protocol. The advantage of the use of X-rays is that they are easily obtained, relatively cheap, have a low patient radiation dose, and thus are ideal for identifying risk factors for sustaining an ACL rupture in large groups of asymptomatic patients.

An interesting sequel of this research would be to compare the differences in bony morphology between patients with and without a re-rupture after ACL reconstruction. This could help the clinician in giving the patient individualized information on the risk of re-rupture.

5. Conclusions

This study indicates that a smaller, flatter intercondylar notch; a lower lateral tibia plateau; a lower medial spike of the eminence; and a smaller tibial eminence were more common in patients who ruptured their ACL compared to control patients.

Further research should focus on ways to implement these differences in bony morphology in prevention programs to prevent ACL rupture in an individual who is at greater risk for sustaining ACL rupture.

Author Contributions: Data curation, K.S.R.v.K., Vincent Eggerding, B.L.v.M., E.v.A. and D.E.M.; Formal analysis, M.R. and J.H.W.; Investigation, K.S.R.v.K. and B.L.v.M.; Methodology, K.S.R.v.K. and J.H.W.; Project administration, E.v.A. and D.E.M.; Software, K.S.R.v.K. and J.H.W.; Supervision, M.R., S.M.A.B.-Z. and D.E.M.; Validation, S.M.A.B.-Z.; Writing—original draft, K.S.R.v.K.; Writing—review

and editing, K.S.R.v.K., V.E., M.R., S.M.A.B.-Z. and D.E.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of the Erasmus medical Center. MEC-2006-223 (KNALL study); MEC-2008-068 (CAS-ACL study); MEC-2017-422 (Control patients).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to patients data privacy.

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

References

1. Frobell, R.B.; Lohmander, L.S.; Roos, H.P.; Lohmander, L.S. Acute rotational trauma to the knee: Poor agreement between clinical assessment and magnetic resonance imaging findings. *Scand. J. Med. Sci. Sports* **2006**, *17*, 109–114. [CrossRef]
2. Moses, B.; Orchard, J.; Orchard, J. Systematic review: Annual incidence of ACL injury and surgery in various populations. *Res. Sports Med.* **2012**, *20*, 157–179. [CrossRef]
3. Nordenvall, R.; Bahmanyar, S.; Adami, J.; Stenros, C.; Wredmark, T.; Felländer-Tsai, L. A Population-Based Nationwide Study of Cruciate Ligament Injury in Sweden, 2001–2009. *Am. J. Sports Med.* **2012**, *40*, 1808–1813. [CrossRef] [PubMed]
4. Lindanger, L.; Strand, T.; Mølster, A.O.; Solheim, E.; Inderhaug, E. Return to Play and Long-term Participation in Pivoting Sports After Anterior Cruciate Ligament Reconstruction. *Am. J. Sports Med.* **2019**, *47*, 3339–3346. [CrossRef]
5. Ajuied, A.; Wong, F.; Smith, C.; Norris, M.; Earnshaw, P.; Back, D.; Davies, A. Anterior cruciate ligament injury and radiologic progression of knee osteoarthritis: A systematic review and meta-analysis. *Am. J. Sports Med.* **2014**, *42*, 2242–2252. [CrossRef] [PubMed]
6. Ardern, C.L.; Taylor, N.F.; Feller, J.A.; Whitehead, T.; Webster, K.E. Psychological Responses Matter in Returning to Preinjury Level of Sport After Anterior Cruciate Ligament Reconstruction Surgery. *Am. J. Sports Med.* **2013**, *41*, 1549–1558. [CrossRef]
7. Niederer, D.; Engeroff, T.; Wilke, J.; Vogt, L.; Banzer, W. Return to play, performance, and career duration after anterior cruciate ligament rupture: A case-control study in the five biggest football nations in Europe. *Scand. J. Med. Sci. Sports* **2018**, *28*, 2226–2233. [CrossRef]
8. Di Stasi, S.; Myer, G.D.; Hewett, T.E. Neuromuscular training to target deficits associated with second anterior cruciate ligament injury. *J. Orthop. Sports Phys. Ther.* **2013**, *43*, 777–792. [CrossRef] [PubMed]
9. Chang, M.J.; Jeong, J.H.; Chang, C.B.; Kim, Y.J.; Seo, B.K.; Song, M.K.; Kang, T.; Kang, S.-B. Revision surgery for failed anterior cruciate ligament reconstruction with extension deficiency. *Scand. J. Med. Sci. Sports* **2018**, *28*, 2604–2610. [CrossRef] [PubMed]
10. Robertsson, O.; Lidgren, L.; Sundberg, M.; W-Dahl, A. *The Swedish Knee Ligament Registry*. Annual Report 2018. 2018. Available online: www.alcregister.nu (accessed on 28 February 2021).
11. Silvers, H.J.; Mandelbaum, B.R. Prevention of anterior cruciate ligament injury in the female athlete. *Br. J. Sports Med.* **2007**, *41* (Suppl. 1), i52–i59. [CrossRef]
12. Myklebust, G.; Engebretsen, L.; Braekken, I.H.; Skjolberg, A.; Olsen, O.E.; Bahr, R. Prevention of anterior cruciate ligament injuries in female team handball players: A prospective intervention study over three seasons. *Clin. J. Sport Med.* **2003**, *13*, 71–78. [CrossRef]
13. Arundale, A.J.H.; Silvers-Granelli, H.; Marmon, A.R.; Zarzycki, R.; Dix, C.; Snyder-Mackler, L. Changes in biomechanical knee injury risk factors across two collegiate soccer seasons using the 11+ prevention program. *Scand. J. Med. Sci. Sports* **2018**, *28*, 2592–2603. [CrossRef]
14. Fowler, P.J.; Messieh, S.S. Isolated posterior cruciate ligament injuries in athletes. *Am. J. Sports Med.* **1987**, *15*, 553–557. [CrossRef]
15. Eggerding, V.; Meuffels, D.E.; Bierma-Zeinstra, S.M.; Verhaar, J.A.; Reijman, M. Factors Related to the Need for Surgical Reconstruction After Anterior Cruciate Ligament Rupture: A Systematic Review of the Literature. *J. Orthop. Sports Phys. Ther.* **2014**, *45*, 37–44. [CrossRef]
16. Shen, L.; Jin, Z.-G.; Dong, Q.-R.; Li, L.-B. Anatomical Risk Factors of Anterior Cruciate Ligament Injury. *Chin. Med. J.* **2018**, *131*, 2960–2967. [CrossRef] [PubMed]
17. Wratten, C.J.; Tetsworth, K.; Hohmann, E. Three-Dimensional Femoral Notch Volume in Anterior Cruciate Ligament–Deficient Versus Anterior Cruciate Ligament–Intact Patients: A Matched Case-Control Study with Inter-gender Comparison. *Arthrosc. J. Arthrosc. Relat. Surg.* **2015**, *31*, 1117–1122. [CrossRef]
18. Eggerding, V.; Van Kuijk, K.S.R.; Van Meer, B.L.; Bierma-Zeinstra, S.M.A.; Van Arkel, E.R.A.; Reijman, M.; Waarsing, J.H.; Meuffels, D.E. Knee shape might predict clinical outcome after an anterior cruciate ligament rupture. *Bone Jt. J.* **2014**, *96*, 737–742. [CrossRef] [PubMed]

19. van Meer, B.L.; Oei, E.H.; Meuffels, D.E.; van Arkel, E.R.; Verhaar, J.A.; Bierma-Zeinstra, S.M.; Reijman, M. Degenerative Changes in the Knee 2 Years After Anterior Cruciate Ligament Rupture and Related Risk Factors: A Prospective Observational Follow-up Study. *Am. J. Sports Med.* **2016**, *44*, 1524–1533. [CrossRef]
20. Eggerding, V.; Reijman, M.; Scholten, R.J.; Verhaar, J.A.; Meuffels, D.E. Computer-assisted surgery for knee ligament reconstruction. *Cochrane Database Syst. Rev.* **2014**, *2014*, CD007601.
21. Meuffels, D.E.; Reijman, M.; Verhaar, J.A. Computer-assisted surgery is not more accurate or precise than conventional arthroscopic ACL reconstruction: A prospective randomized clinical trial. *J. Bone Jt. Surg. Am. Vol.* **2012**, *94*, 1538–1545. [CrossRef]
22. Årøen, A.; Verdonk, P. Posterior cruciate ligament, exploring the unknown. *Knee Surg. Sports Traumatol. Arthrosc.* **2013**, *21*, 996–997. [CrossRef]
23. Cootes, T.F.; Cooper, D.H.; Graham, J. Active shape models: Their training and application. *Comput. Vis. Image Underst.* **1995**, *61*, 38–59. [CrossRef]
24. Bollen, S. Epidemiology of knee injuries: Diagnosis and triage. *Br. J. Sports Med.* **2000**, *34*, 227–228. [CrossRef]
25. Alentorn-Geli, E.; Myer, G.D.; Silvers, H.J.; Samitier, G.; Romero, D.; Lázaro-Haro, C.; Cugat, R. Prevention of non-contact anterior cruciate ligament injuries in soccer players. Part 2: A review of prevention programs aimed to modify risk factors and to reduce injury rates. *Knee Surg. Sports Traumatol. Arthrosc.* **2009**, *17*, 859–879. [CrossRef]
26. Domzalski, M.; Grzelak, P.; Gabos, P. Risk factors for Anterior Cruciate Ligament injury in skeletally immature patients: Analysis of intercondylar notch width using Magnetic Resonance Imaging. *Int. Orthop.* **2010**, *34*, 703–707. [CrossRef]
27. van Diek, F.M.; Wolf, M.R.; Murawski, C.D.; van Eck, C.F.; Fu, F.H. Knee morphology and risk factors for developing an anterior cruciate ligament rupture: An MRI comparison between ACL-ruptured and non-injured knees. *Knee Surg. Sports Traumatol. Arthrosc.* **2013**, *22*, 987–994. [CrossRef] [PubMed]
28. Whitney, D.C.; Sturnick, D.R.; Vacek, P.M.; DeSarno, M.J.; Gardner-Morse, M.; Tourville, T.W.; Smith, H.C.; Slauterbeck, J.R.; Johnson, R.J.; Shultz, S.J.; et al. Relationship Between the Risk of Suffering a First-Time Noncontact ACL Injury and Geometry of the Femoral Notch and ACL: A Prospective Cohort Study With a Nested Case-Control Analysis. *Am. J. Sports Med.* **2014**, *42*, 1796–1805. [CrossRef] [PubMed]
29. Boden, B.P.; Dean, G.S.; Feagin, J.A., Jr.; Garrett, W.E., Jr. Mechanisms of anterior cruciate ligament injury. *Orthopedics* **2000**, *23*, 573–578. [CrossRef] [PubMed]
30. Dienst, M.; Schneider, G.; Altmeyer, K.; Voelkering, K.; Georg, T.; Kramann, B.; Kohn, D.M. Correlation of intercondylar notch cross sections to the ACL size: A high resolution MR tomographic in vivo analysis. *Arch. Orthop. Trauma Surg.* **2007**, *127*, 253–260. [CrossRef]
31. Fridén, T.; Jonsson, A.; Erlandsson, T.; Jonsson, K.; Lindstrand, A. Effect of femoral condyle configuration on disability after an anterior cruciate ligament rupture: 100 patients followed for 5 years. *Acta Orthop. Scand.* **1993**, *64*, 571–574. [CrossRef] [PubMed]
32. Sonnery-Cottet, B.; Archbold, P.; Cucurulo, T.; Fayard, J.-M.; Bortolletto, J.; Thaunat, M.; Prost, T.; Chambat, P. The influence of the tibial slope and the size of the intercondylar notch on rupture of the anterior cruciate ligament. *J. Bone Jt. Surg. Br. Vol.* **2011**, *93*, 1475–1478. [CrossRef] [PubMed]
33. Stijak, L.; Herzog, R.F.; Schai, P. Is there an influence of the tibial slope of the lateral condyle on the ACL lesion? *A case-control study. Knee Surg Sports Traumatol. Arthrosc.* **2008**, *16*, 112–117. [CrossRef] [PubMed]
34. Vyas, S.; Van Eck, C.F.; Vyas, N.; Fu, F.H.; Otsuka, N.Y. Increased medial tibial slope in teenage pediatric population with open physes and anterior cruciate ligament injuries. *Knee Surg. Sports Traumatol. Arthrosc.* **2011**, *19*, 372–377. [CrossRef] [PubMed]
35. Shelbourne, K.D.; Facibene, W.A.; Hunt, J.J. Radiographic and intraoperative intercondylar notch width measurements in men and women with unilateral and bilateral anterior cruciate ligament tears. *Knee Surg. Sports Traumatol. Arthrosc.* **1997**, *5*, 229–233. [CrossRef]

Journal of
Clinical Medicine

MDPI

Article

Estimating the Prevalence of Knee Pain and the Association between Illness Perception Profiles and Self-Management Strategies in the Frederiksberg Cohort of Elderly Individuals with Knee Pain: A Cross-Sectional Study

Elisabeth Ginnerup-Nielsen [1], Robin Christensen [1,2], Berit L Heitmann [1,3], Roy D. Altman [4], Lyn March [5], Anthony Woolf [6], Henning Bliddal [1] and Marius Henriksen [1,7,*]

1 The Parker Institute, Bispebjerg and Frederiksberg Hospital, 2000 Frederiksberg, Denmark; elisabeth.marie.ginnerup.laebo@regionh.dk (E.G.-N.); robin.christensen@regionh.dk (R.C.); berit.lilienthal.heitmann@regionh.dk (B.L.H.); henning.bliddal@regionh.dk (H.B.)
2 Research Unit of Rheumatology, Department of Clinical Research, University of Southern Denmark, University Hospital, 5000 Odense, Denmark
3 Department of Public Health, Section for General Practice, Copenhagen University, 1014 Copenhagen, Denmark
4 Division of Rheumatology and Immunology, University of California, Los Angeles, CA 90095, USA; journals@royaltman.com
5 Department of Rheumatology, Royal North Shore Hospital, Institute of Bone and Joint Research, University of Sydney, St Leonards, NSW 2065, Australia; lyn.march@sydney.edu.au
6 Bone and Joint Research Group, Royal Cornwall Hospital, Truro TR1 3HD, UK; anthony.woolf@btopenworld.com
7 Department of Physical and Occupational Therapy, Copenhagen University Hospital, Bispebjerg and Frederiksberg, 2400 Copenhagen, Denmark
* Correspondence: marius.henriksen@regionh.dk

check for updates

Citation: Ginnerup-Nielsen, E.; Christensen, R.; Heitmann, B.L; Altman, R.D.; March, L.; Woolf, A.; Bliddal, H.; Henriksen, M. Estimating the Prevalence of Knee Pain and the Association between Illness Perception Profiles and Self-Management Strategies in the Frederiksberg Cohort of Elderly Individuals with Knee Pain: A Cross-Sectional Study. *J. Clin. Med.* **2021**, *10*, 668. https://doi.org/10.3390/jcm10040668

Academic Editors: Jos Runhaar and Marienke van Middelkoop
Received: 27 December 2020
Accepted: 5 February 2021
Published: 9 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Abstract: Knee pain is an early sign of later incident radiographic knee osteoarthritis (OA). However, the prevalence of knee pain in the general population is unknown. Additionally, it is unknown how people with knee pain choose to self-manage the condition and if the perception of the illness affects these choices. In this study, 9086 citizens between 60–69 years old in the municipality of Frederiksberg, Copenhagen, Denmark, were surveyed, of which 4292 responded. The prevalence of knee pain was estimated, and associations between illness perceptions (brief illness perception questionnaire [B-IPQ]), self-management strategies, and knee symptoms were assessed. The prevalence of knee pain was 21.4% of which 40.5% reported to use no self-management strategies (non-users). These non-users perceived their knee pain as less threatening and reported less severe symptoms than users of self-management strategies. Further, we found that a more positive illness perception was associated with less severe knee symptoms. In conclusion, among Danes aged 60–69 years, the knee pain prevalence is 21.4%, of which 40.5% use no treatment and perceive the condition as non-threatening. These non-users with knee pain represent a subpopulation being at increased risk of developing knee OA later in life, and there is a potential preventive gain in identifying these persons.

Keywords: knee pain; knee osteoarthritis; early OA; illness perceptions; self-management strategies; cross-sectional study; survey

1. Introduction

The European prevalence of radiographically confirmed symptomatic knee osteoarthritis (OA) is around 4% [1], and the Global Burden of Disease 2010 study [1] ranked hip and knee osteoarthritis (OA) as one of the highest contributors of years lived with disability. With an ageing population, increasing body weight, and sedentary life style, all factors linked to risk of knee OA [2,3], society may head towards a potentially explosive

development in knee OA. The diagnosis of knee OA relies much on radiological changes in the knee [4]; however, intermittent knee pain by itself is a strong predictor of incident radiological knee OA [5–7], and studies have found that people often experience pain and decreased functional level before any radiological changes are evident [1,8]. By the time people are diagnosed with radiographic knee OA, they often experience a greater impact on their quality of life than patients with other chronic conditions such as cancer and heart diseases [9].

When faced with a benign but potentially life changing diagnosis of a chronic disease (such as knee OA), people (in general) develop an organized pattern of perceptions about their condition and how it impacts their lives [10]. These illness perceptions vary between individuals and have been shown to affect physical function. For example, a cohort study of patients with hand, spine, knee, or hip OA showed that patients with negative illness perceptions had significantly worse functional outcomes over six years compared to those with more positive perceptions [11]. Illness perceptions may also affect the way the individual chooses to handle the disease and seek health care assistance. A population-based study of elderly patients found that a notion that nothing can be done to treat arthritis, heart disease, or difficulty sleeping was associated with decreased utilization of preventative health services and decreased likelihood of being affiliated with a general practitioner [12]. Another study found that patients who initially believed that their symptoms would have serious consequences for them had a higher health care use over two years [13].

Various self-management strategies are being recommended for knee OA, and ESCEO and OARSI guidelines have advocated that structured exercise and weight loss should be core interventions in the treatment of OA along with pharmacological treatments [14–17]. Recent reviews recommend controlling obesity as an important aspect in the minimization of arthritic pain syndromes [18] and advocate the use of physical therapy [19] and obesity counselling to improve patient outcomes [20]. Furthermore, complementary and alternative medicine (CAM), such as nutritional supplements, herbal medicine, and acupuncture, are more or less accepted by authorities and have, for decades, been used consistently by patients for different health purposes including chronic joint pain [21–27]. It is, however, unknown how illness perceptions associate with the use of self-management strategies among people with knee pain.

Accordingly, there were two aims of this study: The primary aim was to quantify the prevalence of individuals with self-reported knee pain and knee OA (according to the NICE self-reported framework) in a representative sample of elderly individuals. Secondarily, we wanted to look for associations between self-management strategies and illness perception, intensity of knee pain and other knee symptoms, and health-related quality of life among elderly individuals with knee pain. Finally, we wanted to explore if different illness perception profiles exist and how these profiles differ in terms of self-management strategies, intensity of knee pain, and other knee symptoms, and health related quality of life among elderly individuals with knee pain.

We hypothesized that elderly individuals with knee pain, who use self-management strategies for their knee pain, have a more negative illness perception and have more severe symptoms than individuals who do not use self-management strategies. Further, respondents with a predominantly negative perception of their knee pain were hypothesized to report higher pain levels and lower quality of life and to more frequently use self-management strategies than people with more positive perceptions of their knee pain.

2. Materials and Methods

2.1. Design and Period

We performed a cross-sectional study as part of an ongoing prospective cohort study [28]. Study findings are reported according to Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines [29]. The prospective cohort study was pre-registered at clinicaltrials.gov (accessed on 20 December 2020) (NCT03472300) and

reviewed by the local health research ethics committee who deemed the study exempt from approval (j.no. 17024697). Data collection for this cross-sectional study was initiated on 6 September 2018, and ended 21 October 2018. No formal power calculation was performed due to the exploratory design, where we wanted to invite all the elderly citizens between 60 and 69 years of age. Thus, we anticipated a response rate of around 40 percent to include approximately 3500 participants to have approximately 1000 reporting knee pain [28].

Data was collected through a survey ("Frederiksbergundersøgelsen") sent to all citizens aged between 60 and 69 living in Frederiksberg Municipality, in Copenhagen, Denmark. The survey was sent through the public "Digital Post" system (electronic mailbox for letters from Danish authorities) administered by the platform "e-Boks" [30], linked to the individual's Personal Identification number—a national identification number, which is part of the personal information stored in the Civil Registration System. In the e-letter, an invitation to participate in the study was provided along with a link to an online survey that was managed using REDCap electronic data capture tools hosted at the capital region of Denmark. REDCap (Research Electronic Data Capture) is a secure, web-based software platform designed to support data capture for research studies, providing (1) an intuitive interface for validated data capture, (2) audit trails for tracking data manipulation and export procedures, (3) automated export procedures for seamless data downloads to common statistical packages, and (4) procedures for data integration and interoperability with external sources. (REDCap) [31,32]. The letter included information about the study as well as information about the rights to withdraw from the study at any time. Furthermore, the importance of responding to the questionnaire regardless of whether knee pain was present or not was emphasized.

2.2. Participants

Denmark is a country with a solid base for electronic communication. Most of the citizens have daily access to the internet, and Danes have a strong history of compliance with surveys. Frederiksberg Municipality was chosen for this study, being a very stable community with inhabitants rarely relocating. Inclusion criteria for the parent cohort study consisted of being between 60–69 years, living in the Municipality of Frederiksberg, being able to read and understand Danish, and having access to Digital Post. The age of 60–69 was chosen, as the incidence of knee OA increases with age, and an increasing prevalence of knee pain related disablement is encountered [1,33]. In this study, any knee pain during the last month when either sitting still or moving was an inclusion criterion as well. We had no formalized exclusion criteria.

2.3. Variables and Outcome Measures

The full survey description is available in the published protocol [28]. The questionnaires were adaptive based on how each respondent answered. It included a maximum of 189 questions for people reporting knee pain based on the initial triage question: "Have you experienced any pain from your knee/knees during the last month (both at work and rest)?"

The rest of the questions concerned use of conventional products and treatments and complementary and alternative medicines (CAMs) for knee pain/other reasons than knee pain, or general health. It furthermore concerned earlier knee injuries/surgeries, illness perceptions (related to knee pain), health related quality of life, musculoskeletal health, fitness and physical function, health beliefs, and attitudes concerning use of CAMs and physical activity, lifestyle, and demographics. The duration of knee pain was assessed based on predefined answers (0–6 months, 6–12 months, 1–2 years, 2–5 years, 5–10 years, and more than 10 years).

Furthermore, the following questionnaires were used:

2.3.1. The Brief Illness Perception Questionnaire

We used the Danish version of The Brief Illness Perception Questionnaire (B-IPQ). B-IPQ is a generic nine-item questionnaire developed to rapidly assess the cognitive and emotional representations in a variety of illnesses. B-IPQ is a short version of the 84-item revised illness perception questionnaire (IPQ-R) [34].

The first eight items are scored on a 1–10 numeric rating scale with descriptors (none or extreme) at either end with 1 being no perceived threat in items 1, 2, 5, 6, and 8 (e.g., no symptoms/no consequences/no concern) and highest perceived threat (e.g., no illness control/no effect of treatment/no illness understanding) in items 3, 4, and 7 [35].

B-IPQ assesses perceptions on the following five dimensions: Identity, Cause, Timeline, Consequences, and Cure-Control. [36]. Five of the items assess cognitive illness representations: consequences (Item 1: "How much does your illness affect your life?"), timeline (Item 2: "How long do you think your illness will continue?"), personal control (Item 3: "How much control do you feel you have over your illness?"), treatment control (Item 4: "How much do you think your treatment can help your illness?"), and identity (Item 5: "How much do you experience symptoms from your illness?"). Two of the items assess emotional representations: concern (Item 6: "How concerned are you about your illness?") and emotions (Item 8: "How much does you illness affect you emotionally? e.g., does it make you angry, scared, upset or depressed?"). One item assesses illness comprehensibility (Item 7: "How well do you feel you understand your illness?"). Item 9 is a free text field in which the respondent can formulate their beliefs about their condition (cause). We decided not to use this field in our study, as the data were inconsistent.

In all item questions, we replaced the word "illness" with "knee pain" as recommended when applying the B-IPQ to specific conditions [36]. Given that our respondents are not necessarily receiving any treatments and would therefore not be able to give a reasonable answer, the wording in item 4 was changed from "how much do you think your treatment can help your knee pain?" to "how much do you think treatment can help your knee pain?". This change in wording of the item can affect the generalizability of the question to some extent, as the revised question assesses treatment expectations rather than control among participants not using treatments for knee pain.

The B-IPQ scores have shown good test–retest reliability and adequate concurrent, discriminative, and predictive validity amongst patient samples with musculoskeletal disorders and other chronic disorders [36,37]. Based on other studies [38–40] and the fact that each item represents one component (of many) found to underlie the cognitive representation of illness [41], we explored each B-IPQ item score and not the overall score.

2.3.2. Knee Injury and Osteoarthritis Outcome Score

The Knee Injury and Osteoarthritis Outcome Score (KOOS) is developed as an instrument to assess the patient's opinion about their knee and associated problems. It is patient reported and can be used to assess groups and to monitor individuals. The KOOS consists of 42 items covering five domains, namely, Pain (nine items), Symptoms (seven items), Activities of Daily Living (ADL) (17 items), Sports and Recreation (five items), and knee related QoL (four items). The KOOS adopts a five-point Likert scale scoring system (ranging from 0 (least severe) to 4 (most severe) [42].

A normalized score is calculated for each domain with 100 indicating no symptoms and functional impairment and 0 indicating extreme symptoms and functional impairment. KOOS has been validated for short- and long-term follow-up studies of knee injury [43] and is considered reliable and responsive for assessment of knee complaints in a comparative review of knee-specific outcome [44]. Minimal clinically important differences (MCIDs) for the KOOS is suggested to be 10–17 points [45].

2.3.3. Current Knee Pain and Self-Reported Knee Osteoarthritis

Current knee pain, defined as "the level of pain in your knee today" was assessed with a 100 mm visual analogue scale (VAS) with anchors 0 = "no pain" and 100 = "worst imaginable pain".

Self-reported knee OA was assessed corresponding to the definition made by The National institute for Health and Care Excellence (NICE) in 2014 [46]:

- Age over 45
- Activity related knee pain and
- Morning joint-related stiffness that lasts no longer than 30 min

2.3.4. Health Related Quality of Life

To assess health related quality of life, we applied the EuroQoL five dimensions (EQ-5D) questionnaire. EQ-5D is a standardized measure of health status that provides a simple, generic measure of health. It is applicable to a wide range of health conditions and is ideally suited for use in surveys [47]. The EQ-5D consists of a descriptive system comprising five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Standardized answer options are given (three Likert boxes) and each question is assigned a score from 1 to 3. From the responses, an EQ-5D index is calculated ranging from −0.624 (worst) to 1.000 (best) [48]. EQ-5D is simple, taking only a few minutes to complete. A Danish version of the EQ-5D is available, and a Danish valuation set for reference is available [49].

2.4. Self-Management Strategies

We used the term self-management strategy as an overall definition of different treatment types (CAMs, conventional products, and conventional treatments) where the patient has an active role in choosing and administering the treatment.

2.5. Use of CAMs

We defined complementary and alternative medicines as either "alternative treatments", "dietary supplements", or "vitamins/minerals" taken specifically for knee pain.

- Vitamins and minerals were defined as any supplement containing only vitamins or minerals, taken by the respondent in order to relieve knee pain or promote health. Seventeen predefined choices for the use of vitamins or minerals were given (Supplementary file 1).
- Dietary supplements/herbal medicines (in this manuscript referred to as "dietary supplements") were defined as any supplement not being a vitamin/mineral. Twenty-five predefined choices for the use of dietary supplements were given including fish oil, rosehip, ginger, glucosamine, probiotics, and medical cannabinoids (Supplementary file 1).
- Non-medical treatments were defined as an (active) treatment (normally) not being delivered by a medical doctor or another authorized health professional. Sixteen predefined choices for alternative treatments were provided, including: acupuncture, acupressure, cranio-sacral therapy, hypnosis, and kinesiology (Supplementary file 1).

In the survey, we asked if the respondents used any of these CAMs regularly specifically for their knee pain. Regular use of vitamins/minerals and dietary supplements was defined as "daily or almost daily use", while regular use of non-medical treatments was defined as "use within last year".

2.6. Use of Conventional Products and Treatments

"Conventional products" used specifically for knee pain were defined as over the counter (acetaminophen, codeine, NSAID) and prescription pain medicines (NSAID, opioids). Regular use of these drugs was defined as "daily or almost daily use". We considered prescription pain medicine as part of self-management, as regular administration of this

medication is, in contrast to surgery, very much relying on each patient. "Conventional treatments" were defined as physiotherapy, chiropractic, or weight loss. Use of these treatments was defined as use of physiotherapy/chiropractic "within last year" specifically for the knee pain and having "ever" tried to lose weight specifically to relieve knee pain.

We categorized participants reporting use of any type of treatment for their knee pain as being "users" (of self-management strategies), while respondents not using any type of treatment were categorized as "non-users". In this process, we chose to combine conventional, proven effective treatments and CAMs as being all "self-management strategies", as our goal was to explore how people choose to manage their knee pain regardless of what is recommended by the authorities.

3. Statistical Analyses

Incidence rate (IR) and prevalence rate (PR) are the two most important measures to assess the disease risk and occurrence in epidemiological studies like the Frederiksberg Cohort Study. The incidence rate takes the number of newly identified self-reported knee pain cases divided by people at risk in a defined period of time, while the definition of prevalence is comparable to the proportion of individuals with self-reported knee pain in the population at baseline. These two measures serve different purposes for the Frederiksberg Cohort Study. The PR shows how widespread the condition is, and thus provides us with information on the burden of disease in comparison to the estimated global burden of disease [50]. The calculation of the prevalence of knee pain and knee OA in the Frederiksberg Cohort study is presented in Supplementary File 2.

In order to evaluate the "true prevalence" of self-reported knee pain in this group of elderly citizens, a series of analyses were done based on what we considered the intention-to-survey population. Nonresponse in sample surveys was handled by replacing each missing value with multiple imputations [8,51]: multiple imputation (MI) was used to account for participants who were invited to join the survey but did not report the primary outcome: self-reported knee pain (yes/no). Five imputations as well as the original dataset were performed, and results from these six datasets were combined using Rubin's Rules [51]. Based on these new "complete datasets, with no missing data" we applied standard complete-data methods to analyze the multiply-imputed sets using frequentist ideas (results combined using Rubin's Rules [51], as well as applying simple bootstrap resampling techniques to estimate the empirical range from minimum to maximum for the observed sample [52]).

The exploration of the associations between type and extent of self-management strategies, illness perception (B-IPQ), intensity of knee pain and other symptoms, and health related quality of life was done in several steps. We initially compared the two groups we had created ("users" and "non-users") on demographics and questionnaire data. As data did not follow a normal distribution, Kruskal–Wallis' non-parametric tests for independent samples were conducted to compare user/non-user groups and explore the association between the number of treatment types used and the B-IPQ, KOOS, Current VAS Pain, and EQ-5D.

To investigate if different illness perception profiles existed and how these profiles differed in terms of self-management strategies and health related outcomes measures, we performed a cluster analysis (CA) based on the B-IPQ-item scores and related these clusters to the distribution of users/non-users, knee pain intensity (KOOS pain and current VAS), other symptoms (KOOS), and health-related quality of life (EQ-5D). Based on Frostholm et al. [53] and recommendations in the literature [54], we applied a two-step procedure: (1) a hierarchical analysis (Ward's method) using squared Euclidean distance to determine the optimal number of clusters based on a reformed agglomeration schedule. An inconsistent decrease in the coefficient score is used to indicate that the clusters at this point are distinct and therefore the cluster process should be stopped one step earlier [53]. We then used the centroids from the hierarchical CA as a starting point in a *K*-means CA with a predefined number of clusters to validate the results from the hierarchical analysis.

Differences between clusters in age, knee pain duration, BMI, health related outcomes measures, and pain levels were assessed using Kruskal–Wallis test for non-parametric independent samples. The distributions of males/females and users between the clusters were assessed using Chi^2 test. As the B-IPQ only contains one type of scale (scores 1–10), no standardization was required.

All analyses were performed in SPSS (version 3.3). All P-values and 95% confidence intervals were two-sided. We did not apply explicit adjustments for multiplicity; rather, we interpreted the findings from the multiple tests performed, considering the serious risk of making a false discovery (i.e., a false-positive inference). The statistical tests were reported with P-values for standard hypothesis tests, and any claim of statistical significance was only intended for exploratory purposes (with a statistical α level of 0.05).

4. Results

Among a total of 9086 citizens in the Frederiksberg Community aged between 60–69 years, 882 did not have access to e-Boks. Thus, 8204 were invited to respond to the survey. At the end of the six-week data collection period, 4292 (52.3%) had initiated the questionnaire. Among these, 1758 (40.9%) reported knee pain and 570 (13.3%) reported self-reported knee OA corresponding to NICE self-reported framework.

From our bootstrap resampling technique, we have an empirical interval around the proportion having self-reported knee pain ranging from 19.9% to 22.7%. From this sample, it is fair to assume that the prevalence of knee OA is between 6.4% and 7.4% in this group of individuals in the age 60 to 69 years of age (Supplementary File 2). However, since this is probably a low-level guestimate, we also applied a conservative multiple imputation technique replacing the missing data (i.e., based on a tipping point analysis strategy); from these repeated datasets (combined using Rubin's rule) a conservative estimate on individuals with self-reported knee pain could be as high as 54.2% (rather than 21.4%). As a consequence, the prevalence of knee OA in this sample might be as high as 17.6%.

The analyses on illness perceptions, functional level, current VAS knee pain, and quality of life are based on respondents who reported knee pain, and had responded to all B-IPQ items and at least one of the following measures: KOOS, current VAS, or EQ-5D (n = 1552, 34.9%) (Flowchart Figure 1).

Among these 1552 respondents reporting knee pain, 64% were women. The duration of knee pain was evenly distributed across less than one year to more than 10 years.

A summary of the demographics and questionnaire data is presented in Table 1.

Table 1. Characteristics of participants reporting knee pain and being users or non-users of self-management strategies.

	Users	Non-Users	Difference	p *
			Median (95% CI)	
N (%)	923 (59.5)	629 (40.5)	N/A	
Demographics				
Women, N (%)	612 (66.3)	381 (60.6)	N/A	0.021
age years, median (IQR)	64 (62–67)	64 (62–67)	N/A	0.35
BMI (kg/m^2), median (IQR)	26.6 (23.6–30.5)	24.8 (22.6–27.9)	−1.8 (−2.27 to −1.33)	<0.0001
Knee OA ¶	374 (40.5)	125 (19.9)	N/A	<0.0001
KOOS. 0–100 score (median, IQR)				
KOOS symptoms	57.1 (50.0–64.3)	64.3 (57.1–67.9)	7.2 (5.9 to 8.5)	<0.0001
KOOS QOL	56.2 (43.8–68.8)	68.8 (62.5–81.2)	12.6 (11 to 14.2)	<0.0001
KOOS pain	75.0 (61.1–86.1)	86.1(77.8–91.7)	11.1 (9.7 to 12.5)	<0.0001
KOOS function	80.9 (66.2–91.2)	91.2 (83.8–97.1)	10.3 (8.8 to 11.8)	<0.0001
KOOS Sports and recreation	45 (25–70)	70.0 (50–90)	25.0 (21.7 to 28.3)	<0.0001

Table 1. *Cont.*

	Users	Non-Users	Difference	p *
			Median (95% CI)	
EQ-5D Index (median, IQR)	0.776 (0.723–0.824)	0.824 (0.776–1.000)	0.048 (0.039 to 0.057)	<0.0001
Current VAS pain, 0–100 mm (median, IQR)	25 (10.75–50)	12 (3–25)	−13 (−15 to −11)	<0.0001
Brief-IPQ, 1–10 score (median, IQR)				
consequences B-IPQ 1	4.0 (2–6)	2.0 (2–3)	−2.0 (−2.3 to −1.7)	<0.0001
timeline B-IPQ 2	10.0 (6–10)	8.0 (5–10)	−2.0 (−3.1 to −0.9)	<0.0001
personal control B-IPQ 3	6.0 (4–8)	7.0 (5–9)	1.0 (0.6 to 1.5)	<0.0001
treatment control B-IPQ 4	7.0 (5–9)	5.0 (2–7.5)	−2.0 (−2.4 to −1.6)	<0.0001
identity B-IPQ 5	5.0 (3–7)	3.0 (2–5)	−2.0 (−2.3 to −1.7)	<0.0001
concern B-IPQ 6	5.0 (3–7)	3.0 (2–5)	−2.0 (−2.4 to −1.6)	<0.0001
coherence B-IPQ 7	8.0 (5–9)	8.0 (5–9)	0	0.275
emotional representation B-IPQ 8	2.0 (1–5)	2.0 (1–2)	0	<0.0001

Values are median (IQR) unless otherwise stated. * Statistical significance accepted at $p < 0.05$. ¶ Corresponding to NICE framework for self-reported knee OA. KOOS: Knee Injury and Osteoarthritis Outcome Score, higher scores denote higher functional level. Brief-IPQ (B-IPQ): Brief Illness Perception Questionnaire, higher scores on items 1, 2, 5, 6, and 8 denote a more threatening view of the illness, while higher scores on item 3, 4, and 7 denote a less threatening view of the illness. EQ5D: EuroQol-5 Domain, higher scores denote better quality of life.

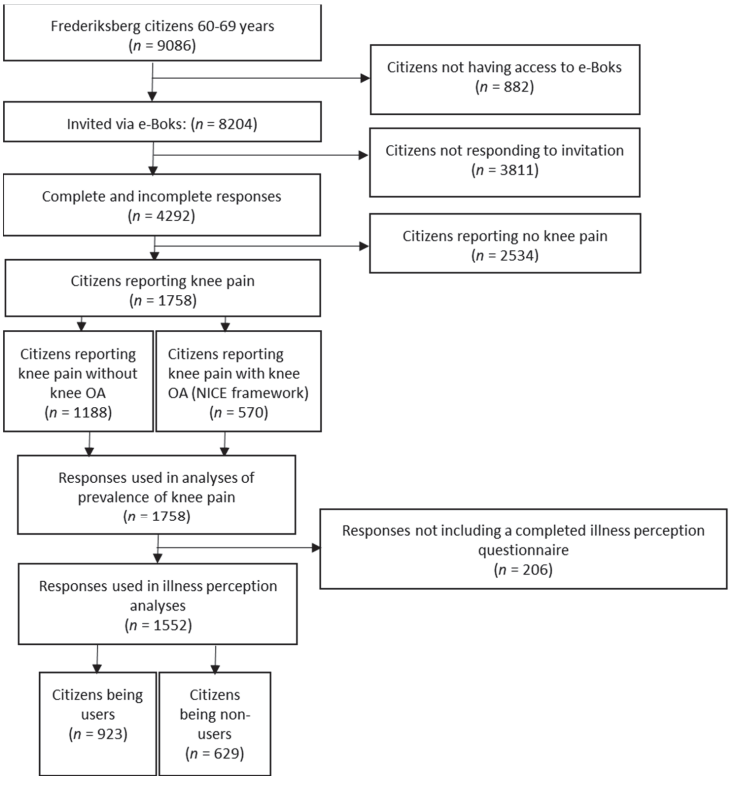

Figure 1. Flow chart response flow.

4.1. Self-Management Strategies for Knee Pain

Of the 1552 respondents, 923 (59.5%) reported use of any kind of treatment or supplement specifically for their knee pain. Hence 59.5% of the respondents were categorized as "users" (of self-management strategies), the rest (n = 629; 40.5%) as "non-users". Of the 923 users, 398 (43%) reported to use only one type of treatment, while only nine respondents (1%) reported use of more than five treatment types. Of the 923 users, 374 (40.5%) reported self-reported knee OA.

The proportions of women among the users and non-users were 66.3% and 60.6%, respectively, with statistically significant difference (p = 0.021).

4.2. Brief Illness Perception Scores in Users and Non-Users

We found differences between users and non-user on all B-IPQ items except for items 7 and 8. On items 1, 2, 4, 5, and 6, users scored higher than non-users (Table 1). On item 3, non-users scored higher than users (Table 1). Altogether, this suggests that non-users perceived their illness as less threatening than users.

The highest average scores were seen on item 2 (Timeline) "How long do you think your knee pain will continue?" with a median score of 8 (IQR 5-10) for non-users and 10 (IQR 6-10) for users. Accordingly, both groups expected their pain to continue for very long with users considering the condition to last longer than non-users.

The lowest average scores concerned the B-IPQ item 8 (Emotional representation)—"How much does your knee pain affect you emotionally?" with a median of 2 (IQR 1-2) for non-users and 2 (IQR 1-5) for users, indicating that both users and non-users were only a little emotionally affected by their knee pain. (Table 1).

On item 3: "How much control do you feel you have over your knee pain?", non-users had higher median scores than users (difference: 1 (95% CI: 0.55 to 1.45)), suggesting that they felt they had more control over their pain than users. On item 4, "How much do you think treatment can help your knee pain?", users had a higher median score than non-users (difference: 2 (95% CI: 1.57 to 2.43)) suggesting that they found treatment more likely to help them than non-users did (Table 1).

4.3. Health-Related Outcome Measures

Non-users scored significantly lower (less symptoms) than users on all KOOS subscales with the largest differences observed in the KOOS Sports and Recreation and KOOS QOL subscales (median difference respectively: 25 (95% CI: 21.0 to 28.34) and 12.6 (95% CI: 11 to14.2)) and the smallest difference on the KOOS symptoms subscale (median difference: 7.2 (95% CI: 5.9 to 8.5)). Likewise, EQ-5D scores were also significantly higher among non-users than users (median difference: 0.048 (95% CI: 0.039 to 0.057)), while the median Current VAS Pain was 13 points lower (95% CI: −15 to −11) among non-users compared to users (Table 1).

There were statistically significant associations ($p < 0.0001$) between the number of treatments used for knee pain and B-IPQ scores (except for item 7) (Figure 2A–H), KOOS scores (Figure 3A–E), VAS pain scores (Figure 4A), and EQ-5D scores (Figure 4B), suggesting that with more treatments used the illness perception was more negative and the symptoms and quality of life were worse.

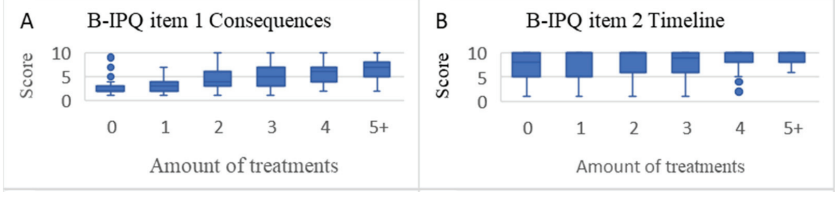

Figure 2. *Cont.*

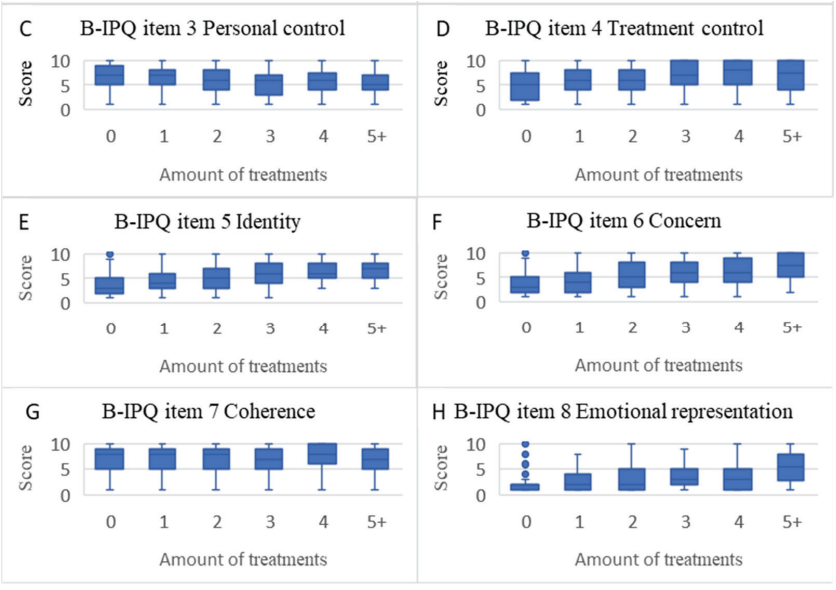

Figure 2. Boxplots of association between amount of treatments used and Brief Illness Perception Questionnaire scores (median, upper, and lower quartiles and outliers). (**A**): B-IPQ item 1 consequences. (**B**): B-IPQ item 2 timeline (**C**): B-IPQ item 3 personal control. (**D**): B-IPQ item 4 treatment control. (**E**): B-IPQ item 5 identity. (**F**): B-IPQ item 6 concern. (**G**): B-IPQ item 7 coherence. (**H**): B-IPQ item 8 emotional representation.

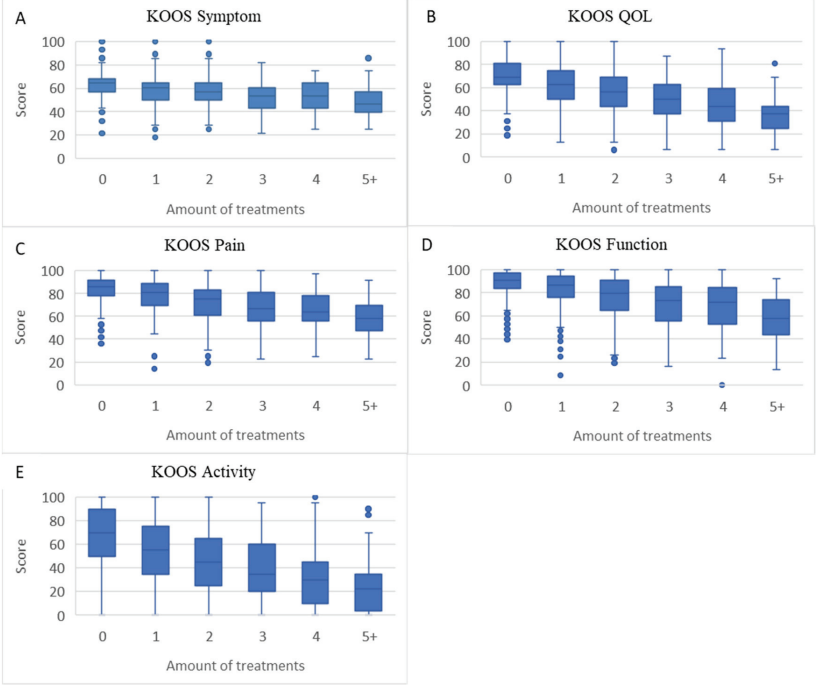

Figure 3. Boxplots of association between amount of treatments used and KOOS scores (median, upper, and lower quartiles and outliers). (**A**): KOOS Symptom, (**B**): KOOS QOL. (**C**): KOOS Pain. (**D**): KOOS Function, (**E**): KOOS Activity.

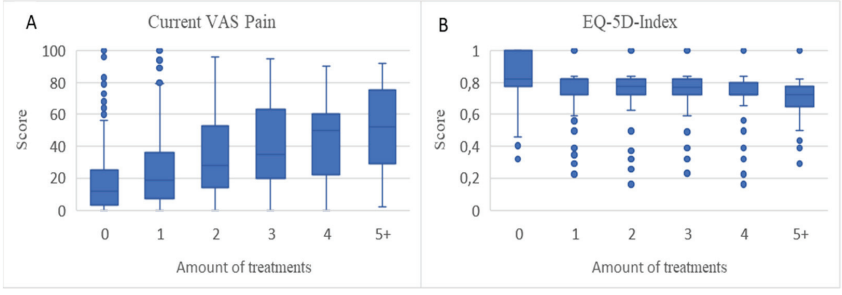

Figure 4. Boxplots of association between amount of treatments used and current VAS pain scores and EQ-5D scores (median, upper, and lower quartiles and outliers). (**A**): Current VAS Pain. (**B**): EQ-5D-Index.

4.4. Cluster Analysis of Brief-IPQ Scores

Based on the change between coefficients in the agglomeration schedule from the hierarchical cluster analysis, we found a solution of two clusters to be optimal. In the next step, we used the centroids of the clusters in a K-means cluster analysis with a pre-set number of two clusters.

We found statistically significant differences between all B-IPQ item scores except item 7 as well as between all health-related outcome measures and Current VAS Pain in the two-cluster solution (Table 2).

Table 2. Health outcome scores between clusters of knee pain perception.

	Cluster 1 "Concerned Optimists" (*n* = 642)	**Cluster 2 "Unconcerned Confident" (*n* = 910)**	**Difference**	
	Median (IQR)	**Median (IQR)**	**Median (95% CI)**	***p* ***
Demographics				
Women (N, %)	416 (64.8)	577 (63.4)	N/A	0.57
Age (median, IQR)	64 (62–67)	64 (62–67)	N/A	0.476
BMI (median, IQR)	26.8 (23.8–30.7)	25.2 (22.8–28.36)	1.6 (1.1 to 2.1)	<0.0001
Knee OA ¶	321 (50)	178 (19.6)	N/A	<0.0001
KOOS. 0–100 score (median, IQR)				
KOOS Symptoms	53.6 (46.4–60.7)	64.3 (57.1–67.9)	−10.7 (−12.1 to −9.3)	<0.0001
KOOS Qol	50 (37.5–56.2)	75 (62.5–81.2)	−25 (−25.7 to −24.3)	<0.0001
KOOS Pain	66.7 (55.6–77.8)	86.1 (80.6–91.7)	−19.4 (−20.0 to −18.8)	<0.0001
KOOS Function	72.1 (57.4–83,8)	92.6 (85.30–97.1)	−20.5 (−21.3 to −19.8)	<0.0001
KOOS Sports and recreation	35 (20–51.25)	70 (55–85)	−35 (−38 to −32)	<0.0001
EQ5D Index, median (IQR)	0.756 (0.723–0.824)	0.824 (0.818–1.000)	−0.068 (−0.074 to −0.062)	<0.0001
Current VAS Pain, median (IQR)	40 (20–62)	11 (3–24)	29 (26.4 to 31.6)	<0.0001

Table 2. *Cont.*

	Cluster 1 "Concerned Optimists" (n = 642)	Cluster 2 "Unconcerned Confident" (n = 910)	Difference	
	Median (IQR)	**Median (IQR)**	**Median (95% CI)**	p *
User types median (IQR)				
Non-users (N, %)	161 (25.1)	468 (51.4)	N/A	<0.0001
(only) CAM users	106 (10.6)	167 (18.4)	N/A	<0.0001
(only) Pharmacological treatment users [#]	28 (4.4)	10 (1.1)	N/A	
(only)Non-pharmacological treatment users [¤]	90 (14)	131 (14.4)	N/A	
Two or more treatment types [§]	257 (40)	134 (14.7)	N/A	
Brief-IPQ, 1–10 score (median, IQR)				
consequences B-IPQ 1	5 (4–7)	2 (2–3)	3 (2.9 to 3.1)	<0.0001
timeline B-IPQ 2	10 (8–10)	8 (4–10)	2 (1.7 to 2.3)	<0.0001
personal control B-IPQ 3	5 (3–7)	8 (5–9)	−3 (−3.4 to −2.6)	<0.0001
treatment control B-IPQ 4	8 (6–10)	5 (2–7)	3 (2.7 to 3.3)	<0.0001
identity B-IPQ 5	6 (5–7)	3 (2–4)	3 (2.9 to 3.1)	<0.0001
concern B-IPQ 6	7 (5–8)	2 (2–3)	5 (4.7 to 5.3)	<0.0001
coherence B-IPQ 7	7 (5–9)	8 (5–9)	−1 (−2.8 to 0.8)	0.275
emotional representation B-IPQ 8	4 (2–6)	1 (1–2)	3 (2.9 to 3.1)	<0.0001

Values are median (IQR), unless otherwise stated. * Statistical significance accepted at $p < 0.05$. [¶] Corresponding to NICE framework for self-reported knee OA. KOOS: Knee Injury and Osteoarthritis Outcome Score, higher scores denote higher functional level. Brief-IPQ (B-IPQ): Brief Illness Perception Questionnaire, higher scores on items 1, 2, 5, 6, and 8 denote a more threatening view of the illness, while higher scores on item 3, 4, and 7 denote a less threatening view of the illness. EQ5D: EuroQol-5 Domain, higher scores denote better quality of life. CAM: Conventional and alternative medicine. [#] Use of over the counter medication and prescription drugs. [¤] Use of physiotherapy, chiropractic, or weight loss. [§] Use of two or more treatment types (CAMS, non-pharmacological or pharmacological treatments).

The population in Cluster 1 had higher current VAS pain, higher BMI, and lower scores on EQ-5D and KOOS than cluster 2. Thus, cluster 1 was more affected on all health-related outcomes measures and perceived their pain as more "threatening" than cluster 2. Additionally, 481 (25.1%) of respondents in cluster 1 were characterized as non-users (Table 2).

Cluster 1 included 642 respondents. It was characterized by higher B-IPQ scores than cluster 2 on all items except for items 3 and 7 (Table 2). Based on the B-IPQ pattern, this cluster was named "concerned optimists".

Cluster 2 included 910 respondents. Respondents in cluster 2 had lower B-IPQ scores than cluster 1 on all items except for items 3 and 7 (Table 2). Based on the B-IPQ pattern this cluster was named "unconcerned confident".

Both clusters had similar scores on item 7 (coherence), suggesting that the two clusters had the same understanding of their knee pain (Table 2). The two clusters were similar with respect to age (median 64 in both clusters). The proportions of women in the two clusters were 64.8% in cluster 1 and 63.4% in cluster 2 ($p = 0.57$), while the proportions of self-reported knee OA were 50% in cluster 1 and 19.6% in cluster 2 ($p < 0.0001$).

5. Discussion

In this cross-sectional study, based on 4292 elderly individuals from the Frederiksberg Cohort, 1758 reported knee pain with 570 reporting symptoms corresponding to knee OA. The prevalence of knee pain and knee OA in the whole Frederiksberg Cohort (8204) was

estimated to be, respectively, 21.4% and 6.9%, which is a little lower than found in other studies [1,33,55].

To our knowledge, this is the first study to document associations between illness perceptions and management patterns of knee pain in a large sample of individuals between 60 and 69 years. We found that users of self-management strategies for knee pain are characterized by more negative illness perceptions, worse symptoms, and lower health related quality of life than non-users. Further, we identified a group of individuals that, even though they reported having knee pain and being functionally affected did not use any self-management strategy.

When compared to populations with manifest radiological knee OA included in clinical studies, users reported slightly less pain and better functional level (KOOS), while part of the non-users had a significantly better functional level [44,56]. Knee pain levels among non-users were comparable to other studies of patients with early OA [6,57], suggesting that the non-users in this study are at increased risk of developing radiographic knee OA.

Even if functional levels were relatively high among respondents, it is evident that levels are still markedly lower than in an age-matched population without knee issues [58] and while non-users, not surprisingly, report less symptoms than users, it is striking that some non-users report significant symptoms, with 19.9% actually reporting symptoms corresponding to self-reported knee OA. On the other hand, the non-users' median EQ-5D score is comparable with the Danish population norms [59], suggesting that their knee pain did not impact their overall quality of life.

Our results are similar to those found in a review of studies concerning the association between illness perception and different functional and psychological health outcomes measures among knee or hip arthroplasty populations [60], where higher scores on the B-IPQ items concerning consequences, identity, coherence, and emotional representation were predictive of worse knee function, pain interference with walking, and anxiety and depression. Another study of 1204 Irish citizens with chronic pain [61] found higher scores on timeline and concern items (items no. 2 and 6) to be predictive of pain-related disability.

We also found that illness perception, knee symptoms, and health related quality of life were associated with amount of treatments used. The more threatening respondents perceived their knee pain and the more physically impaired they were, the more treatments they reported to use. Similar results were found in a study by Hill et al. [39] concerning 2113 older people with musculoskeletal hand disorders, where participants who considered their disorder to have the most negative effects on their lives were more likely to consult a physician, take medication, or both.

On the other hand, a study by Bedson et al. [62] found that around 50% of individuals aged over 50 with disabling knee pain did not consult for it. Recent onset and severity of pain was associated with more use of general practice.

Together, these results indicate that experiencing knee pain is not necessarily considered a serious health threat to people, which could explain why many people do not consult general practice. As resent research [5–7] has shown intermittent knee pain to be a predictor of radiographic knee OA, it seems relevant to inform the general population about the potential gain of reacting with knee OA preventing measures to even mild knee pain. This would also be in line with earlier studies [63–65], emphasizing the effect of using population-based approaches to control determinants of incidence instead of individually targeted approaches to patients being at high risk or already having manifest disease.

Our cluster analysis revealed two clusters based on illness perception traits. Cluster 1 was characterized by worse illness perceptions, more pain, lower quality of life, and significantly worse health outcome measures than cluster 2. Cluster 2 had more positive illness perceptions but lower treatment control scores and better health outcome measures. Cluster 1 had a markedly higher proportion of users than cluster 2 (74.9% vs. 48.6%). The differences between clusters in especially KOOS subscales are highly clinically relevant and cannot only be explained by the distribution of users in the groups. The difference

supports the view that illness perceptions and knee pain related functional level are closely related, and that a considerable part of the respondents use no self-management strategy despite being affected by their knee pain (25.1% in cluster 1).

Other studies have used cluster analyses based on illness perceptions to identify subgroups and relate these to management of a disease. Lowe et al. [66] examined the difference between use of unscheduled health care services in UK among three clusters based on the B-IPQ on patients with both chronic and non-chronic disorders, while Frostholm et al. [53] identified three clusters based on IPQ-R and related these to the use of primary care in Denmark. Both studies found that clusters where patients perceived their pain as having most consequences for their lives had more primary care visits. Riviera et al. [67] identified three clusters based on 187 patients with heart failure, chronic obstructive pulmonary disease (COPD), and chronic kidney disease. The cluster perceiving the illness to have few consequences and a non-fluctuating pattern had the fewest hospitalizations while both the cluster perceiving many consequences to their illness and the cluster perceiving a high disease control and understanding of their illness had many hospitalizations. These studies support our results of more negative perceptions being related to higher use of health care services, but also indicate that more negative perceptions of an illness might not be consequently related to use of more health-related services. As health services operate differently in each country, and as we did not measure hospitalizations, these results are not directly comparable to our results. Nevertheless, the results indicate that different illness perception profiles exist and are related to illness management.

This study has some limitations. First, as this is a cross-sectional study, we cannot conclude on the causality of the relationship between illness perception traits, management patterns, and health related outcome measures we have found. Our study population was sampled from a relatively wealthy area in Copenhagen, which may limit generalizability. Further, we defined "management of knee pain" as the use of different predefined treatment strategies based on recommendations and also the most common CAMs, which may not be the best way to describe an individual's way of managing knee pain. However, we chose to include CAMs as a treatment for knee pain regardless of recommendations, as many people use them and believe in their effect, and as the effect of many types of CAM has not yet been investigated, we considered it relevant to include these.

As CAMs are often broadly defined depending on country or region [21,22,24–26,68,69], generalization related to CAMs in general should be done with caution.

The strengths of this study include a large sample size and a good response rate. Data was collected in an electronic survey questionnaire with branching, ensuring that all responses were presented and collected correctly. Another strength is that the main findings related to the association between illness perception and use (or non-use) of self-management strategies were documented using two analytical approaches.

6. Conclusions

In conclusion, we found the prevalence of knee pain in the Frederiksberg Cohort of elderly between 60–69 years to be 21.4%, while the prevalence of self-reported knee OA was 6.9%. We further found that elderly individuals with knee pain that use treatments and supplements as management strategies seem to be more concerned and have more severe symptoms than individuals that use no treatments or supplements. Additionally, we documented that worse illness perception was associated with higher degrees of knee symptoms, lower quality of life, and a tendency to use more treatments for knee pain. A large proportion of the elderly are not using any treatments even though they are affected by their knee pain, with 19.9% reporting knee symptoms corresponding to self-reported knee OA. These non-users with knee pain represent a subpopulation with an increased risk of developing radiographic knee OA, and there may be a potential preventive gain in identifying these persons that are not yet patients but report symptoms of early OA.

Supplementary Materials: The following are available online at https://www.mdpi.com/2077-0383/10/4/668/s1: Supplementary file 1: Predefined choices, Supplementary file 2: Prevalence of knee OA.

Author Contributions: Conceptualization: E.G.-N., R.C., B.L.H., R.D.A., L.M., A.W., H.B., and M.H.; methodology: E.G.-N., R.C., B.L.H., R.D.A., L.M., A.W., H.B., and M.H.; formal analysis: E.G.-N. and R.C.; investigation: E.G.-N., H.B., and M.H.; resources: H.B.; data curation: E.G.-N.; writing—original draft preparation: E.G.-N.; writing—review and editing: all authors. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no specific funding, but the Cohort "Frederiksbergundersøgelsen" was funded by Ellen Mørchs Foundation (grant number: 324962019). The Parker Institute, Bispebjerg, and Frederiksberg Hospital are supported by a core grant from the Oak Foundation (OCAY-18-774-OFIL).

Institutional Review Board Statement: According to the local health research ethics committee, this study did not require ethical approval. (j.no. 17024697).

Informed Consent Statement: All subjects gave informed consent when responding to the survey.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to Danish and European law.

Acknowledgments: The authors would like to thank Julie Midtgaard Klausen, Ida Kær Thorsen, and Sine Rossen for their assistance with this project.

Conflicts of Interest: E.G.-N.: Declares no conflicts of interest. R.C.: Declares no conflicts of interest. B.L.H.: Declares no conflicts of interest. R.D.A.: Declares no conflicts of interest. L.M.: Declares no conflicts of interest. A.W.: Declares no conflicts of interest. H.B.: Declares no conflicts of interest. M.H.: Consultancies for The Thuasne Group. 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. Cross, M.; Smith, E.; Hoy, D.; Nolte, S.; Ackerman, I.; Fransen, M.; Bridgett, L.; Williams, S.; Guillemin, F.; Hill, C.L.; et al. The global burden of hip and knee osteoarthritis: Estimates from the global burden of disease 2010 study. *Ann. Rheum. Dis.* **2014**, *73*, 1323–1330. [CrossRef] [PubMed]
2. Apold, H.; Meyer, H.E.; Nordsletten, L.; Furnes, O.; Baste, V.; Flugsrud, G.B. Risk factors for knee replacement due to primary osteoarthritis, a population based, prospective cohort study of 315,495 individuals. *BMC Musculoskelet. Disord.* **2014**, *15*, 217. [CrossRef] [PubMed]
3. Järvholm, B.; Lewold, S.; Malchau, H.; Vingård, E. Age, bodyweight, smoking habits and the risk of severe osteoarthritis in the hip and knee in men. *Eur. J. Epidemiol.* **2005**, *20*, 537–542. [CrossRef]
4. Sharma, L.; Kapoor, D.; Issa, S. Epidemiology of osteoarthritis: An update. *Curr. Opin. Rheumatol.* **2006**, *18*, 147–156. [CrossRef] [PubMed]
5. Neogi, T.; Felson, D.; Niu, J.; Nevitt, M.; Lewis, C.E.; Aliabadi, P.; Sack, B.; Torner, J.; Bradley, L.; Zhang, Y. Association between radiographic features of knee osteoarthritis and pain: Results from two cohort studies. *BMJ* **2009**, *339*. [CrossRef] [PubMed]
6. Peat, G.; Thomas, E.; Duncan, R.; Wood, L. Is a "false-positive" clinical diagnosis of knee osteoarthritis just the early diagnosis of pre–radiographic disease? *Arthritis Care Res.* **2010**, *62*, 1502–1506. [CrossRef] [PubMed]
7. Thorstensson, C.; Andersson, M.; Jönsson, H.; Saxne, T.; Petersson, I. Natural course of knee osteoarthritis in middle-aged subjects with knee pain: 12-year follow-up using clinical and radiographic criteria. *Ann. Rheum. Dis.* **2009**, *68*, 1890–1893. [CrossRef]
8. Peat, G.; McCarney, R.; Croft, P. Knee pain and osteoarthritis in older adults: A review of community burden and current use of primary health care. *Ann. Rheum. Dis.* **2001**, *60*, 91–97. [CrossRef] [PubMed]
9. Wu, M.; Brazier, J.E.; Kearns, B.; Relton, C.; Smith, C.; Cooper, C.L. Examining the impact of 11 long-standing health conditions on health-related quality of life using the EQ-5D in a general population sample. *Eur. J. Health Econ.* **2015**, *16*, 141–151. [CrossRef] [PubMed]
10. Lau-Walker, M. Predicting self-efficacy using illness perception components: A patient survey. *Br. J. Health Psychol.* **2006**, *11*, 643–661. [CrossRef] [PubMed]
11. Kaptein, A.A.; Bijsterbosch, J.; Scharloo, M.; Hampson, S.E.; Kroon, H.M.; Kloppenburg, M. Using the common sense model of illness perceptions to examine osteoarthritis change: A 6-year longitudinal study. *Health Psychol.* **2010**, *29*, 56. [CrossRef]
12. Goodwin, J.S.; Black, S.A.; Satish, S. Aging versus disease: The opinions of older black, Hispanic, and non-Hispanic white Americans about the causes and treatment of common medical conditions. *J. Am. Geriatr. Soc.* **1999**, *47*, 973–979. [CrossRef]

13. Frostholm, L.; Fink, P.; Christensen, K.S.; Toft, T.; Oernboel, E.; Olesen, F.; Weinman, J. The patients' illness perceptions and the use of primary health care. *Psychosom. Med.* **2005**, *67*, 997–1005. [CrossRef] [PubMed]
14. Arden, N.K.; Perry, T.A.; Bannuru, R.R.; Bruyère, O.; Cooper, C.; Haugen, I.K.; Hochberg, M.C.; McAlindon, T.E.; Mobasheri, A.; Reginster, J.-Y. Non-surgical management of knee osteoarthritis: Comparison of ESCEO and OARSI 2019 guidelines. *Nat. Rev. Rheumatol.* **2020**, 1–8. [CrossRef]
15. Zhang, W.; Moskowitz, R.; Nuki, G.; Abramson, S.; Altman, R.; Arden, N.; Bierma-Zeinstra, S.; Brandt, K.; Croft, P.; Doherty, M. OARSI recommendations for the management of hip and knee osteoarthritis, part I: Critical appraisal of existing treatment guidelines and systematic review of current research evidence. *Osteoarthr. Cartil.* **2007**, *15*, 981–1000. [CrossRef] [PubMed]
16. Zhang, W.; Moskowitz, R.; Nuki, G.; Abramson, S.; Altman, R.; Arden, N.; Bierma-Zeinstra, S.; Brandt, K.; Croft, P.; Doherty, M. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. *Osteoarthr. Cartil.* **2008**, *16*, 137–162. [CrossRef]
17. Zhang, W.; Nuki, G.; Moskowitz, R.; Abramson, S.; Altman, R.; Arden, N.; Bierma-Zeinstra, S.; Brandt, K.; Croft, P.; Doherty, M. OARSI recommendations for the management of hip and knee osteoarthritis: Part III: Changes in evidence following systematic cumulative update of research published through January 2009. *Osteoarthr. Cartil.* **2010**, *18*, 476–499. [CrossRef]
18. Farha, N.; Spencer, A.; McGervey, M. Outpatient Evaluation of Knee Pain. *Med Clin. N. Am.* **2020**, *105*, 117–136. [CrossRef] [PubMed]
19. Dantas, L.O.; De Fátima Salvini, T.; McAlindon, T.E. Knee osteoarthritis: Key treatments and implications for physical therapy. *Braz. J. Phys. Ther.* **2020**. [CrossRef] [PubMed]
20. Inam, S.H.A.; Riaz, B.; Jamil, H.; Rafique, D.; Siddiqi, U.A.; Iqbal, M.; Sherwani, N.Z.F.; Khan, W. Do Patients with Osteoarthritis get Weight Loss Counseling? *Cureus* **2020**, *12*. [CrossRef] [PubMed]
21. Mbizo, J.; Okafor, A.; Sutton, M.A.; Burkhart, E.N.; Stone, L.M. Complementary and Alternative Medicine Use by Normal Weight, Overweight, and Obese Patients with Arthritis or Other Musculoskeletal Diseases. *J. Altern. Complement. Med.* **2016**, *22*, 227–236. [CrossRef] [PubMed]
22. Al-Windi, A. Predictors of herbal medicine use in a Swedish health practice. *Pharmacoepidemiol. Drug Saf.* **2004**, *13*, 489–496. [CrossRef] [PubMed]
23. Biltoft-Jensen, A.A.; Hindborg, H.; Christensen, T.; Saxholt, E.; Bredsdorff, L.; Knudsen, V.K.; Ravn-Haren, G.; Fødevareinstituttet, D. De fleste får nok vitaminer og mineraler fra kosten alene. *E-artikel DTU Fødevareinstitutet* **2016**, *2016*, 1–13.
24. Clarke, T.C.; Black, L.I.; Stussman, B.J.; Barnes, P.M.; Nahin, R.L. Trends in the use of complementary health approaches among adults: United States, 2002–2012. *Natl. Health Stat. Rep.* **2015**, *79*, 1.
25. Djuv, A.; Nilsen, O.G.; Steinsbekk, A. The co-use of conventional drugs and herbs among patients in Norwegian general practice: A cross-sectional study. *BMC Complement. and Altern. Med.* **2013**, *13*, 295. [CrossRef]
26. Phang, J.K.; Kwan, Y.H.; Hendra, G.; Tan, V.I.C.; Thumboo, J.; Østbye, T.; Fong, W. Complementary and alternative medicine for rheumatic diseases: A systematic review of randomized controlled trials. *Complement. Ther. Med.* **2018**. [CrossRef] [PubMed]
27. Schnabel, K.; Binting, S.; Witt, C.M.; Teut, M. Use of complementary and alternative medicine by older adults—A cross-sectional survey. *BMC Geriatr.* **2014**, *14*, 38. [CrossRef]
28. Ginnerup-Nielsen, E.M.; Henriksen, M.; Christensen, R.; Heitmann, B.L.; Altman, R.; March, L.; Woolf, A.; Karlsen, H.; Bliddal, H. Prevalence of self-reported knee symptoms and management strategies among elderly individuals from Frederiksberg municipality: Protocol for a prospective and pragmatic Danish cohort study. *BMJ Open* **2019**, *9*, e028087. [CrossRef]
29. Vandenbroucke, J.P.; Von Elm, E.; Altman, D.G.; Gotzsche, P.C.; Mulrow, C.D.; Pocock, S.J.; Poole, C.; Schlesselman, J.J.; Egger, M.; Initiative, S. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration. *Ann. Intern. Med.* **2007**, *147*, W163–W194. [CrossRef]
30. e-Boks. Available online: https://www.e-boks.com/danmark/en (accessed on 15 December 2020).
31. Harris, P.A.; Taylor, R.; Minor, B.L.; Elliott, V.; Fernandez, M.; O'Neal, L.; McLeod, L.; Delacqua, G.; Delacqua, F.; Kirby, J. The REDCap consortium: Building an international community of software platform partners. *J. Biomed. Inform.* **2019**, *95*, 103208. [CrossRef]
32. Harris, P.A.; Taylor, R.; Thielke, R.; Payne, J.; Gonzalez, N.; Conde, J.G. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inform.* **2009**, *42*, 377–381. [CrossRef] [PubMed]
33. Kiadaliri, A.A.; Lohmander, L.S.; Moradi-Lakeh, M.; Petersson, I.F.; Englund, M. High and rising burden of hip and knee osteoarthritis in the Nordic region, 1990–2015: Findings from the Global Burden of Disease Study 2015. *Acta Orthop.* **2018**, *89*, 177–183. [CrossRef]
34. Moss-Morris, R.; Weinman, J.; Petrie, K.; Horne, R.; Cameron, L.; Buick, D. The revised illness perception questionnaire (IPQ-R). *Psychol. Health* **2002**, *17*, 1–16. [CrossRef]
35. Broadbent, E.; Petrie, K.J.; Main, J.; Weinman, J. The Brief Illness Perception Questionnaire. Available online: http://ipq.h.uib.no/pdf/B-IPQ-Danish.pdf (accessed on 25 May 2020).
36. Broadbent, E.; Petrie, K.J.; Main, J.; Weinman, J. The brief illness perception questionnaire. *J Psychosom. Res.* **2006**, *60*, 631–637. [CrossRef]
37. Broadbent, E.; Wilkes, C.; Koschwanez, H.; Weinman, J.; Norton, S.; Petrie, K.J. A systematic review and meta-analysis of the Brief Illness Perception Questionnaire. *Psychol. Health* **2015**, *30*, 1361–1385. [CrossRef]

38. Costa, E.C.; Vale, S.; Sobral, M.; Graca Pereira, M. Illness perceptions are the main predictors of depression and anxiety symptoms in patients with chronic pain. *Psychol. Health Med.* **2016**, *21*, 483–495. [CrossRef] [PubMed]
39. Hill, S.; Dziedzic, K.; Thomas, E.; Baker, S.; Croft, P. The illness perceptions associated with health and behavioural outcomes in people with musculoskeletal hand problems: Findings from the North Staffordshire Osteoarthritis Project (NorStOP). *Rheumatology* **2007**. [CrossRef] [PubMed]
40. Wilski, M.; Tasiemski, T. Illness perception, treatment beliefs, self-esteem, and self-efficacy as correlates of self-management in multiple sclerosis. *Acta Neurol. Scand.* **2016**, *133*, 338–345. [CrossRef]
41. Weinman, J.; Petrie, K.J.; Moss-Morris, R.; Horne, R. The illness perception questionnaire: A new method for assessing the cognitive representation of illness. *Psychol. Health* **1996**, *11*, 431–445. [CrossRef]
42. Roos, E.M.; Roos, H.P.; Lohmander, L.S.; Ekdahl, C.; Beynnon, B.D. Knee Injury and Osteoarthritis Outcome Score (KOOS)—Development of a self-administered outcome measure. *J. Orthop. Sports Phys. Ther.* **1998**, *28*, 88–96. [CrossRef] [PubMed]
43. Roos, E.M.; Toksvig-Larsen, S. Knee injury and Osteoarthritis Outcome Score (KOOS)–Validation and comparison to the WOMAC in total knee replacement. *Health Qual. Life Outcomes* **2003**, *1*, 17. [CrossRef] [PubMed]
44. Collins, N.; Prinsen, C.; Christensen, R.; Bartels, E.; Terwee, C.; Roos, E. Knee Injury and Osteoarthritis Outcome Score (KOOS): Systematic review and meta-analysis of measurement properties. *Osteoarthr. Cartil.* **2016**, *24*, 1317–1329. [CrossRef] [PubMed]
45. Monticone, M.; Ferrante, S.; Salvaderi, S.; Motta, L.; Cerri, C. Responsiveness and minimal important changes for the Knee Injury and Osteoarthritis Outcome Score in subjects undergoing rehabilitation after total knee arthroplasty. *Am. J. Phys. Med. Rehabil.* **2013**, *92*, 864–870. [CrossRef]
46. National Institute for Health and Clinical Excellence. *Osteoarthritis: National Clinical Guideline for Care and Management in Adults*; National Institute for Health and Clinical Excellence: London, UK, 2014.
47. Rabin, R.; Charro, F.D. EQ-5D: A measure of health status from the EuroQol Group. *Ann. Med.* **2001**, *33*, 337–343. [CrossRef] [PubMed]
48. Lauritsen, J. Danske Normtal for Euroqol-5d. Available online: http://uag.dk/simpelfunktion/pdf/eq5dknorm.pdf (accessed on 21 September 2020).
49. Wittrup-Jensen, K.U.; Lauridsen, J.; Gudex, C.; Pedersen, K.M. Generation of a Danish TTO value set for EQ-5D health states. *Scand. J. Public Health* **2009**, *37*, 459–466. [CrossRef] [PubMed]
50. Vos, T.; Lim, S.S.; Abbafati, C.; Abbas, K.M.; Abbasi, M.; Abbasifard, M.; Abbasi-Kangevari, M.; Abbastabar, H.; Abd-Allah, F.; Abdelalim, A. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: A systematic analysis for the Global Burden of Disease Study 2019. *Lancet* **2020**, *396*, 1204–1222. [CrossRef]
51. Rubin, D.B. *Multiple Imputation for Nonresponse in Surveys*; John Wiley & Sons: Hoboken, NJ, USA, 2004; Chapter 1; Volume 81.
52. Bland, J.M.; Altman, D.G. Statistics notes: Bootstrap resampling methods. *BMJ* **2015**, *350*. [CrossRef] [PubMed]
53. Frostholm, L.; Hornemann, C.; Ørnbøl, E.; Fink, P.; Mehlsen, M. Using Illness Perceptions to Cluster Chronic Pain Patients. *Clin. J. Pain* **2018**, *34*, 991–999. [CrossRef]
54. Clatworthy, J.; Hankins, M.; Buick, D.; Weinman, J.; Horne, R. Cluster analysis in illness perception research: A Monte Carlo study to identify the most appropriate method. *Psychol. Health* **2007**, *22*, 123–142. [CrossRef]
55. Postler, A.; Ramos, A.L.; Goronzy, J.; Günther, K.-P.; Lange, T.; Schmitt, J.; Zink, A.; Hoffmann, F. Prevalence and treatment of hip and knee osteoarthritis in people aged 60 years or older in Germany: An analysis based on health insurance claims data. *Clin. Interv. Aging* **2018**, *13*, 2339. [CrossRef]
56. Henriksen, M.; Christensen, R.; Klokker, L.; Bartholdy, C.; Bandak, E.; Ellegaard, K.; Boesen, M.P.; Riis, R.G.C.; Bartels, E.M.; Bliddal, H. Evaluation of the benefit of corticosteroid injection before exercise therapy in patients with osteoarthritis of the knee: A randomized clinical trial. *JAMA Intern. Med.* **2015**, *175*, 923–930. [CrossRef]
57. Hayati, M.; Yazdi, Z.; Abbasi, M. Comparison of non-steroidal anti-inflammatory drugs and knee kinesio taping in early osteoarthritis pain: A randomized controlled trial. *J. Bodyw. Mov. Ther.* **2019**, *23*, 666–670. [CrossRef]
58. Marot, V.; Murgier, J.; Carrozzo, A.; Reina, N.; Monaco, E.; Chiron, P.; Berard, E.; Cavaignac, E. Determination of normal KOOS and WOMAC values in a healthy population. *Knee Surg. Sports Traumatol. Arthrosc.* **2019**, *27*, 541–548. [CrossRef]
59. Sørensen, J.; Davidsen, M.; Gudex, C.; Pedersen, K.M.; Brønnum-Hansen, H. Danish EQ-5D population norms. *Scand. J. Public Health* **2009**, *37*, 467–474. [CrossRef]
60. Sawyer, A.T.; Harris, S.L.; Koenig, H.G. Illness perception and high readmission health outcomes. *Health Psychol. Open* **2019**, *6*, 2055102919844504. [CrossRef]
61. Raftery, M.N.; Sarma, K.; Murphy, A.W.; De la Harpe, D.; Normand, C.; McGuire, B.E. Chronic pain in the Republic of Ireland—community prevalence, psychosocial profile and predictors of pain-related disability: Results from the Prevalence, Impact and Cost of Chronic Pain (PRIME) study, part 1. *Pain* **2011**, *152*, 1096–1103. [CrossRef]
62. Bedson, J.; Mottram, S.; Thomas, E.; Peat, G. Knee pain and osteoarthritis in the general population: What influences patients to consult? *Fam. Pract.* **2007**, *24*, 443–453. [CrossRef]
63. Capewell, S.; Dowrick, C. Healthful diet and physical activity for cardiovascular disease prevention in adults without known risk factors: Is behavioral counselling necessary? *JAMA Intern. Med.* **2017**, *177*, 1254–1255. [CrossRef]
64. Rose, G. Sick individuals and sick populations. *Int. J. Epidemiol.* **2001**, *30*, 427–432. [CrossRef]
65. Sacks, G.; Swinburn, B.; Lawrence, M. Obesity Policy Action framework and analysis grids for a comprehensive policy approach to reducing obesity. *Obes. Rev.* **2009**, *10*, 76–86. [CrossRef]

66. Lowe, R.; Porter, A.; Snooks, H.; Button, L.; Evans, B.A. The association between illness representation profiles and use of unscheduled urgent and emergency health care services. *Br. J. Health Psychol.* **2011**, *16*, 862–879. [CrossRef] [PubMed]
67. Rivera, E.; Corte, C.; Steffen, A.; DeVon, H.A.; Collins, E.G.; McCabe, P.J. Illness representation and self-care ability in older adults with chronic disease. *Geriatrics* **2018**, *3*, 45. [CrossRef] [PubMed]
68. Brown, M.; Dean, S.; Hay-Smith, E.J.C.; Taylor, W.; Baxter, G.D. Musculoskeletal pain and treatment choice: An exploration of illness perceptions and choices of conventional or complementary therapies. *Disabil. Rehabil.* **2010**, *32*, 1645–1657. [CrossRef] [PubMed]
69. Nielsen, M.W.; Hansen, E.H.; Rasmussen, N.K. Use of natural medicines in the Danish population: A national cross-sectional survey. *Ann. Pharmacother.* **2005**, *39*, 1534–1538. [CrossRef]

Journal of
Clinical Medicine

MDPI

Article

The Added Value of Radiographs in Diagnosing Knee Osteoarthritis Is Similar for General Practitioners and Secondary Care Physicians; Data from the CHECK Early Osteoarthritis Cohort

Qiuke Wang [1], Jos Runhaar [1,*], Margreet Kloppenburg [2], Maarten Boers [3], Johannes W. J. Bijlsma [4], Sita M. A. Bierma-Zeinstra [1,5] and The CREDO Expert Group

1 Department of General Practice, Erasmus MC University Medical Center Rotterdam, 3015 GD Rotterdam, The Netherlands; q.wang@erasmusmc.nl (Q.W.); s.bierma-zeinstra@erasmusmc.nl (S.M.A.B.-Z.)
2 Department of Rheumatology, Leiden University Medical Center, 2333 ZA Leiden, The Netherlands; G.Kloppenburg@lumc.nl
3 Department of Epidemiology & Data Science, Amsterdam Rheumatology & Immunology Center, Amsterdam UMC, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands; eds@amsterdamumc.nl
4 Department of Rheumatology and Clinical Immunology, University Medical Centre Utrecht, 3584 CX Utrecht, The Netherlands; j.w.j.bijlsma@umcutrecht.nl
5 Department of Orthopaedics, Erasmus MC University Medical Center Rotterdam, 3015 GD Rotterdam, The Netherlands
* Correspondence: j.runhaar@erasmusmc.nl; Tel.: +31(0)-10-7044192; Fax: +31(0)-10-7044766

Received: 28 August 2020; Accepted: 19 October 2020; Published: 21 October 2020

Abstract: Objective: The purpose of this study was to evaluate the added value of radiographs for diagnosing knee osteoarthritis (KOA) by general practitioners (GPs) and secondary care physicians (SPs). Methods: Seventeen GPs and nineteen SPs were recruited to evaluate 1185 knees from the CHECK cohort (presenters with knee pain in primary care) for the presence of clinically relevant osteoarthritis (OA) during follow-up. Experts were required to make diagnoses independently, first based on clinical data only and then on clinical plus radiographic data, and to provide certainty scores (ranging from 1 to 100, where 1 was "certainly no OA" and 100 was "certainly OA"). Next, experts held consensus meetings to agree on the final diagnosis. With the final diagnosis as gold standard, diagnostic indicators were calculated (sensitivity, specificity, positive/negative predictive value, accuracy and positive/negative likelihood ratio) for all knees, as well as for clinically "certain" and "uncertain" knees, respectively. Student paired *t*-tests compared certainty scores. Results: Most diagnoses of GPs (86%) and SPs (82%) were "consistent" after assessment of radiographic data. Diagnostic indicators improved similarly for GPs and SPs after evaluating the radiographic data, but only improved relevantly in clinically "uncertain" knees. Radiographs added some certainty to "consistent" OA knees (GP 69 vs. 72, $p < 0.001$; SP 70 vs. 77, $p < 0.001$), but not to the consistent no OA knees (GP 21 vs. 22, $p = 0.16$; SP 20 vs. 21, $p = 0.04$). Conclusions: The added value of radiographs is similar for GP and SP, in terms of diagnostic accuracy and certainty. Radiographs appear to be redundant when clinicians are certain of their clinical diagnosis.

Keywords: knee osteoarthritis; radiography; general practitioner; secondary care physician; diagnosis

1. Introduction

In routine clinical practice, the diagnosis of knee osteoarthritis (KOA) is usually made based on the clinician's expertise, and radiographs are frequently used to confirm clinical suspicion of KOA [1,2].

However, there are insufficient data on the necessity and the potential role of radiographs in the diagnostic process.

The European League Against Rheumatism Recommendations (EULAR) reported that three symptoms (knee pain, morning stiffness less than 30 min and functional limitation) combined with three clinical signs (crepitus, restricted range of motion and bone enlargement) could predict 99% radiographic KOA [3]. Similarly, recent studies showed that clinical manifestations, such as knee pain, crepitus, joint line tenderness, bony swelling and pain on flexion/extension could be used for identifying radiographic KOA [4–6]. Current recommendations advise not to use imaging in patients with typical OA presentations, but these were mainly based on expert opinion [7,8].

As a common and chronic disease [8–10], KOA is usually diagnosed both by general practitioners (GPs) and secondary care physicians (SPs). The added diagnostic value of radiographs can be different between the two kinds of clinicians given their different clinical expertise. However, there is no scientific literature on this aspect.

In this study, we recruited both GPs and SPs with osteoarthritis (OA) expertise to assess clinical vignettes taken from the CHECK cohort study (a longitudinal cohort study of primary care patients with knee complaints suggestive for early stage KOA, followed for 10 years) of potential KOA patients and to provide diagnoses based on either clinical data alone, or clinical combined with radiographic data. The aim of this study was to evaluate the added value of radiographs above clinical findings in diagnosing KOA and to see whether this differed between GPs and SPs.

2. Methods

2.1. Clinical Experts

The protocol has been approved by the Ethical Committee of UMC Utrecht (protocol number 02/017-E). We recruited experts who fulfilled one of the following criteria in this study: (i) had a degree in general practice, orthopedics, rheumatology or sports medicine for 2 or more years; (ii) were in-training for a degree in these specialties combined with a PhD in OA research.

We tested experts' characteristics by querying them on the number of OA patients treated per week, experience in OA treatment (years), and their perception on the importance of radiographs in making the diagnosis.

2.2. Clinical and Radiographic Data

For the present study, we included all patients from the CHECK cohort [11–13]. The CHECK cohort recruited patients between October 2002 and September 2005, and all patients were supposed to be followed for 10 years. Patients whose medical records and radiographs were available from 5 up to 10-year follow-up were included in this study.

Clinical data, obtained at a 5, 8 and 10-year follow up, consisted of demographics (including sex, age, BMI (body mass index), racial background, marital status, menopausal status, educational level, chronic diseases, occupation, smoking status and alcohol usage), physical examinations (presence of knee pain, morning stiffness in knee, knee warmth, bony tenderness, crepitus, knee pain on extension and flexion, range of motion (extension and flexion)), WOMAC (Western Ontario and McMaster Universities Osteoarthritis Index) subscales of pain, function, and stiffness, NRS (numeric rating scale) pain scores and incidence of other diseases (quadriceps tendinitis, intra-articular fracture, Baker's cyst, ligament or meniscus damage, osteochondritis dissecans, plica syndrome and septic arthritis).

Radiographic data consisted of scores from centralized reading by trained readers evaluating standard weight bearing (posterior-anterior fixed flexion view) knee radiographic films at 5, 8, and 10 years follow up (for details see [12]). The scores included information on tibial attrition, femoral/tibial sclerosis, joint space narrowing, femoral/tibial osteophytes and Kellgren and Lawrence grades. Both posterior-anterior fixed flexion and lateral films were also made available to the experts.

Table S1 summarizes all clinical and radiographic data presented to the experts. All data were stored and presented in special software (built in-house) for optimal presentation. The software recorded actual access to the radiographic films.

2.3. Obtaining Diagnoses

Before starting the diagnostic process, all experts received written information and completed two example patients to get familiarized with the procedures and software. We obtained expert diagnosis between June 2018 and January 2019.

Experts were divided into pairs; each pair consisted of one GP and one SP, where possible. The diagnostic process is presented in Figure 1. Each pair assessed the same subset of knees (40–50 patients). First, longitudinal clinical data of each patient were presented. Each expert evaluated these independently and, for each knee, chose between "yes, clinically relevant OA has developed" and "no, clinically relevant OA has not developed". In addition, the experts had to provide their certainty on a 1 to 100 scale (integers with unit of 1), where 1 was "certainly no OA" and 100 was "certainly OA". If a knee was diagnosed as "OA", the certainty score had to be between 51 and 100, with a higher score expressing greater certainty; if the knee was diagnosed as "no OA", certainty score had to be between 1 and 49, with a lower score expressing greater certainty. Next, access to longitudinal radiographic data and films was activated. Experts were asked the same questions and had to provide new certainty scores. At this stage, experts had read-only access to the clinical data and their corresponding diagnoses.

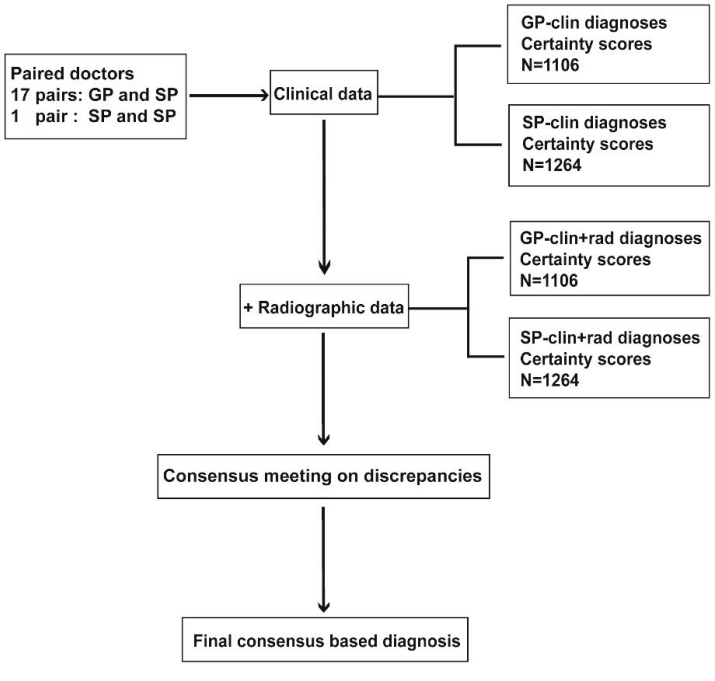

Figure 1. Diagnostic process. OA, osteoarthritis; GP, general practitioner; SP, secondary care physician; Clin diagnosis, diagnoses based on clinical data only; clin + rad, diagnoses based on clinical and radiographic data.

After individually finishing all these evaluations, knees assigned certainty scores >30 and <70 were defined as "uncertain", the remainder as "certain". Where the two experts agreed (yes/no OA, regardless

of certainty) the diagnosis automatically became final. Each pair held a consensus meeting to re-assess the knees where the individual diagnoses were discrepant, except if both experts were "uncertain". At that meeting the expert pair evaluated both clinical and radiographic data of the discrepancies, as done when evaluating these individually, and made a final diagnosis together. Knees where no consensus could be reached and those for which the experts disagreed after the individual scoring, but both were "uncertain", were all labeled as "consensus based uncertain".

2.4. Statistics

Categorical variables were presented as counts and percentages and normally distributed continuous data as mean ± standard deviation. Experts characteristics were compared by Mann–Whitney U test or Wilcoxon W test, where appropriate. The number of consistent and amended diagnoses after assessment of the radiographic data were presented for GP and SP, split for no OA and OA diagnoses obtained when evaluating the clinical data only. Chi-square tests were used for comparing diagnoses before and after viewing radiographic data in GP/SP. We calculated sensitivity, specificity, positive/negative predictive value (PPV/NPV), accuracy, positive/negative likelihood ratio (LR+/–) and their 95% CIs (confidence intervals) for the GP and SP diagnosis separately, with the consensus based final diagnosis as gold standard. Next, we split all knees into clinically "certain" (individual certainty scores ≤30 or ≥70, based on clinical data only) and clinically "uncertain" (individual certainty scores >30 and <70, based on clinical data only) and the same diagnostic indicators were calculated within both groups. The primary objective of the present study was to assess the clinically relevant value of radiographs. With the consideration that statistical tests could be too sensitive to detect minor differences in such a large sample, we did not apply statistical analysis for comparing the above described diagnostic indicators. As no comparable results have been reported before, outcomes were deemed exploratory.

For the analysis of certainty scores, the knees were divided into four subgroups: "consistent OA" (the clinical diagnosis of OA was retained after viewing radiographic data), "amended to no OA" (clinical diagnosis OA amended to no OA after viewing radiographic data); and likewise, "consistent no OA", and "amended to OA". Paired *t*-tests assessed whether diagnostic certainty was improved with radiographic information, in "consistent OA" and "consistent no OA" knees. To assure robustness of the results, a sensitivity analysis compared certainty scores between left knees only, with a paired *t*-test.

Analysis was performed with SPSS version 25.0 (IBM, Chicago, IL, USA); the significance level was 0.05 using a 2-sided p value for all tests.

3. Results

3.1. Experts and Patients

A total of 36 experts were recruited, 17 GPs and 19 SPs (8 orthopedists, 9 rheumatologists and 2 sports physicians). Seventeen pairs of one GP and one SP were formed, and the remaining pair included one orthopedist and one rheumatologist.

Expert characteristics are shown in Table 1. Among all the characteristics, only the perceived importance of radiography was significantly different between GP and SP ($p < 0.001$).

Table 1. Expert characteristics.

	General Practitioner (N = 17)	Orthopedist (N = 8)	Rheumatologist (N = 9)	Sports Physician (N = 2)	*p* Values [&]
Importance radiograph *, median (range)	2 (1–4)	4 (4–4)	3 (2–4)	3 (3–3)	<0.001
Number of OA patients treated per week, mean (SD)	5 (3)	53 (25)	6 (7)	3 (2)	0.06
Years of experience, mean (SD)	12 (9)	14 (5)	17 (9)	20 (14)	0.10

OA, osteoarthritis; * Perceived importance of radiography for making the diagnosis of knee osteoarthritis: 1, not important; 2, of minor importance; 3, somewhat important; 4, very important; [&] Comparing characteristics between GP and SP.

The study included 761 patients with 1185 symptomatic knees; 79% female, mean (SD) age 56 (5) years. The pairs formed by one GP and one SP evaluated 1106 knees, and the pair consisted of two SP did for 79 knees. During the diagnostic process, GP viewed the actual films in 45% of the knees and SP did this in 69% of the knees.

3.2. Diagnostic Abilities

3.2.1. General Practitioners

GPs diagnosed 42% of knees as OA based on the clinical data only and 44% after viewing radiographic data. In total, 86% of diagnoses were consistent after viewing radiographic data; 6% OA knees were amended to no OA, and 8% no OA knees to OA (Figure 2). Of the 14% amended diagnoses, 8% were deemed correct, compared to the final diagnosis (Table 2). In general, the changes in diagnoses were statistically significant ($p < 0.001$) and all diagnostic indicators were somewhat improved after viewing radiographic data (Table 3).

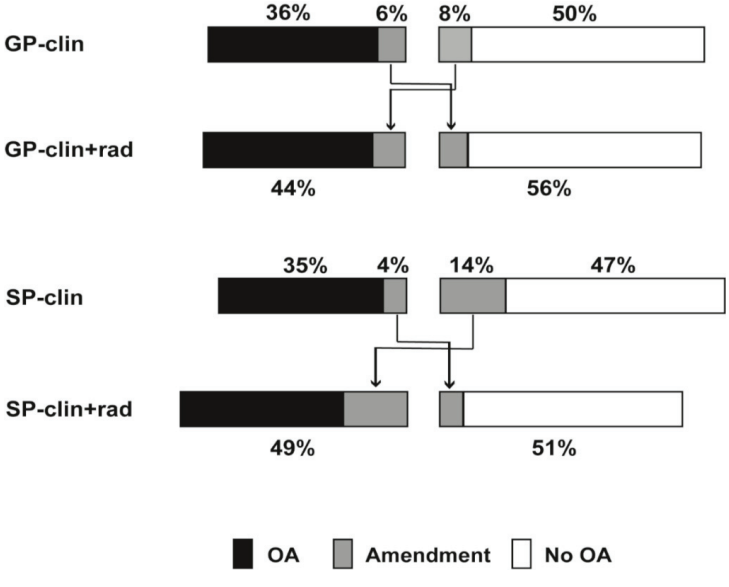

Figure 2. General practitioner (GP) and secondary care physician (SP) diagnoses based on clinical data (clin), clinical combined with radiographic data (clin + rad); OA, osteoarthriti.

GPs were uncertain about 41% of their clinical diagnoses. They were much more likely to amend uncertain diagnoses than their certain diagnoses (23% uncertain vs. 7% of certain diagnoses).

Furthermore, the rate of correct amendments in clinically uncertain diagnoses was 2 times higher than that in certain diagnoses (Table 4). Without radiographic assessment, the GP diagnostic ability was good in certain diagnoses (PPV = 0.82 (95% CI, 0.76–0.87), NPV = 0.71 (95%CI, 0.67–0.75)), but poor in uncertain diagnoses (PPV = 0.59 (95% CI, 0.53–0.65), NPV = 0.57 (95%CI, 0.43–0.58)). Diagnostic abilities were much improved by radiographs in uncertain diagnoses (Table 5). In clinically certain diagnoses, slight improvements were found in PPV, NPV and LR− (PPV: 0.82 to 0.84; NPV: 0.71 to 0.77; LR−: 0.30 to 0.18), LR+ values were consistently greater than 10 (15.39 to 31.69); in clinically uncertain diagnoses, all indicators were much improved (PPV: 0.59 to 0.73; NPV: 0.51 to 0.61; LR+: 2.14 to 4.89; LR−: 0.30 to 0.04).

3.2.2. Secondary Care Physicians

SPs diagnosed 39% of knees as OA based on clinical data only and 49% after viewing radiographic data. In total, 82% of diagnoses were consistent after viewing radiographic data; 4% OA knees were amended to no OA and 14% no OA knees to OA (Figure 2). Of the 18% amended diagnoses, 9% were deemed correct as compared with the final diagnosis (Table 2). In general, the changes in diagnoses were statistically significant ($p < 0.001$) and all diagnostic indicators were somewhat improved after viewing radiographic data (Table 3).

SPs were uncertain about 36% of their clinical diagnoses. They were much more likely to amend uncertain diagnoses than certain diagnoses (27% uncertain vs. 14% of certain diagnoses). Furthermore, the rate of correct amendments in clinically uncertain diagnoses was 3 times higher than that in certain diagnoses (Table 4). Without radiographic assessment, the SP diagnostic ability was good in certain diagnoses (PPV = 0.78 (95%CI, 0.72–0.83), NPV = 0.74 (95%CI, 0.70–0.77)), but poor in uncertain diagnoses (PPV = 0.57 (95%CI, 0.50–0.63), NPV = 0.42 (95%CI, 0.35–0.49)). Diagnostic abilities were much improved by radiographs in uncertain diagnoses (Table 5). In clinically certain diagnoses, some fluctuations were found in PPV, NPV and LR− (PPV: 0.78 to 0.71; NPV: 0.74 to 0.86; LR−: 0.28 to 0.11), LR+ values were consistently greater than 10 (14.98 to 21.02); in clinically uncertain diagnoses, all indicators were much improved (PPV: 0.57 to 0.66; NPV: 0.42 to 0.68; LR+: 1.88 to 4.81; LR−: 0.45 to 0.05).

3.3. Diagnostic Certainties

Diagnostic certainty scores are presented in Figure 3. For GP, diagnostic certainty improved somewhat in the "consistent OA" knees (69 ± 12 vs. 72 ± 14, $p < 0.001$), while no significant improvement was found in "consistent no OA" knees (21 ± 13 vs. 22 ± 14, $p = 0.16$). Sensitivity analysis showed very similar results (see Table S2).

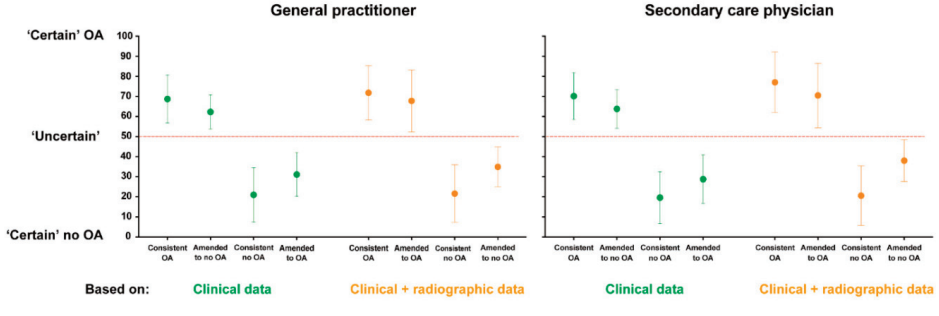

Figure 3. Certainty scores of four subgroups before and after viewing radiographic data (means and standard deviations).

Table 2. Expert diagnoses and certainty scores.

	OA %	Certainty Scores of OA Knees Mean (SD)	No OA %	Certainty Scores of No-OA Knees Mean (SD)	Consistent Diagnoses * %	Amended Diagnoses %	Correctly Amended # %
GP-clin, N = 1106	42 &	68 (12)	58 &	22 (14)	86	14	8
GP-clin + rad, N = 1106	44 &	71 (14)	56 &	23 (15)			
SP-clin, N = 1264	39 $	69 (12)	61 $	22 (13)	82	18	9
SP-clin + rad, N = 1264	49 $	75 (16)	51 $	22 (15)			

OA, osteoarthritis; GP-clin, general practitioners' diagnoses based on clinical data only; GP-clin + rad, general practitioners' diagnoses based on clinical and radiographic data; SP-clin, secondary care physicians' diagnoses based on clinical data only; SP-clin + rad, secondary care physicians' diagnoses based on clinical and radiographic data; * the clinical diagnosis of OA/no OA was retained after viewing radiographic data; # diagnoses correctly amended after viewing radiographs, compared to final diagnosis; & p <0.001, chi-square test for comparing the diagnoses before and after viewing radiographs in GP; $ p <0.001, chi-square test for comparing the diagnoses before and after viewing radiographs in SP.

Table 3. Comparison of initial and consensus-based final diagnosis.

		Final Diagnosis			Sensitivity (95%CI)	Specificity (95%CI)	PPV (95%CI)	NPV (95%CI)	Accuracy (95%CI)	LR+ (95%CI)	LR− (95%CI)
		OA *	No A *	Uncertain *,†							
GP-clin (N = 1106)	OA	29	6	7	0.76 (0.72–0.80)	0.86 (0.82–0.89)	0.68 (0.64–0.73)	0.66 (0.62–0.69)	0.67 (0.63–0.71)	5.23 (4.24–6.63)	0.28 (0.24–0.33)
	No OA	9	38	11							
GP-clin + rad (N = 1106)	OA	34	3	6	0.89 (0.86–0.92)	0.92 (0.90–0.95)	0.78 (0.74–0.82)	0.72 (0.68–0.75)	0.75 (0.71–0.78)	11.75 (8.60–16.04)	0.12 (0.09–0.16)
	No OA	4	41	12							
SP-clin (N = 1264)	OA	26	6	7	0.73 (0.69–0.77)	0.88 (0.84–0.90)	0.67 (0.62–0.71)	0.66 (0.62–0.69)	0.66 (0.62–0.70)	5.86 (4.68–7.32)	0.31 (0.26–0.36)
	No OA	10	40	11							
SP-clin + rad (N = 1264)	OA	33	4	12	0.93 (0.90–0.95)	0.92 (0.90–0.94)	0.69 (0.65–0.72)	0.82 (0.78–0.85)	0.75 (0.71–0.78)	11.84 (8.94–15.70)	0.08 (0.06–0.11)
	No OA	3	42	6							

OA, osteoarthritis; GP-clin, general practitioners' diagnoses based on clinical data only; GP-clin + rad, general practitioners' diagnoses based on clinical and radiographic data; SP-clin, secondary care physicians' diagnoses based on clinical data only; SP-clin + rad, secondary care physicians' diagnoses based on clinical and radiographic data; PPV, positive predictive value; NPV, negative predictive value; LR+, positive likelihood ratio; LR−, negative likelihood ratio. The diagnostic indicators were calculated by treating the final diagnosis as reference standard and using all knees, including OA, no OA and uncertain diagnoses, in the calculations; * Percentages (%) of patients of each subgroup; ‡ Knees diagnosed as "consensus based uncertain".

Table 4. Final diagnoses and rate of amendment split by certainty of initial clinical diagnosis (% of patients of each subgroup).

		Clin+ Rad Clin	OA	No OA	Consistent Diagnoses *	Amended Diagnoses	Correctly Amended #
Clinically 'certain' knees	GP (N = 658)	OA	28	2	93	7	5
		no OA	5	65			
	SP (N = 811)	OA	27	2	86	14	5
		no OA	12	59			
Clinically 'uncertain' knees	GP (N = 448)	OA	48	13	77	23	13
		no OA	10	29			
	SP (N = 453)	OA	49	9	73	27	16
		no OA	18	24			

OA, osteoarthritis; Clin: knees diagnosed as OA or no OA based on clinical data only; Clin + rad: knees diagnosed as OA or no OA based on clinical and radiographic data; * the clinical diagnosis of OA/no OA was retained after viewing radiographic data; # diagnoses correctly amended after viewing radiographs, compared to the final diagnosis.

Table 5. Comparing experts' clinically certain/uncertain diagnoses to consensus based final diagnosis.

			Final Diagnosis			Sensitivity (95%CI)	Specificity (95%CI)	PPV (95%CI)	NPV (95%CI)	Accuracy (95%CI)	LR+ (95%CI)	LR− (95%CI)
			OA *	No OA *	Uncertain *,†							
Clinically 'certain' knees	GP-clin (N = 658)	OA	24	2	3	0.71 (0.65–0.77)	0.95 (0.92–0.97)	0.82 (0.76–0.87)	0.71 (0.67–0.75)	0.74 (0.71–0.81)	15.39 (9.47–25.02)	0.30 (0.25–0.37)
		No OA	10	50	11							
	GP-clin + rad (N = 658)	OA	28	1	4	0.82 (0.77–0.87)	0.97 (0.95–0.99)	0.84 (0.78–0.88)	0.77 (0.72–0.80)	0.79 (0.75–0.84)	31.69 (16.58–60.56)	0.18 (0.14–0.24)
		No OA	6	51	10							
	SP-clin (N = 811)	OA	23	3	4	0.74 (0.68–0.79)	0.95 (0.93–0.97)	0.78 (0.72–0.83)	0.74 (0.70–0.77)	0.75 (0.71–0.80)	14.98 (9.90–22.67)	0.28 (0.22–0.34)
		No OA	8	52	10							
	SP-clin + rad (N = 811)	OA	27	2	9	0.90 (0.85–0.93)	0.96 (0.93–0.97)	0.71 (0.66–0.76)	0.86 (0.82–0.89)	0.80 (0.74–0.84)	21.02 (13.51–32.71)	0.11 (0.08–0.16)
		No OA	3	53	6							
Clinically 'uncertain' knees	GP-clin (N = 448)	OA	36	11	13	0.81 (0.75–0.86)	0.62 (0.53–0.70)	0.59 (0.53–0.65)	0.51 (0.43–0.58)	0.56 (0.48–0.62)	2.14 (1.72–2.67)	0.30 (0.22–0.40)
		No OA	9	20	11							
	GP-clin + rad (N = 448)	OA	43	6	9	0.97 (0.93–0.98)	0.80 (0.73–0.86)	0.73 (0.67–0.78)	0.61 (0.54–0.68)	0.68 (0.60–0.73)	4.89 (3.51–6.83)	0.04 (0.02–0.09)
		No OA	2	25	15							
	SP-clin (N = 453)	OA	33	11	14	0.72 (0.66–0.78)	0.62 (0.53–0.70)	0.57 (0.50–0.63)	0.42 (0.35–0.49)	0.51 (0.42–0.56)	1.88 (1.49–2.37)	0.45 (0.36–0.57)
		No OA	13	18	11							
	SP-clin + rad (N = 453)	OA	44	6	17	0.96 (0.92–0.98)	0.80 (0.72–0.86)	0.66 (0.60–0.71)	0.68 (0.60–0.76)	0.67 (0.60–0.73)	4.81 (3.40–6.79)	0.05 (0.02–0.10)
		No OA	2	23	8							

OA, osteoarthritis; GP-clin, general practitioners' diagnoses based on clinical data only; GP-clin + rad, general practitioners' diagnoses based on clinical and radiographic data; SP-clin, secondary care physicians' diagnoses based on clinical data only; SP-clin + rad, secondary care physicians' diagnoses based on clinical and radiographic data; PPV, positive predictive value; NPV, negative predictive value; LR+, positive likelihood ratio; LR−, negative likelihood ratio. The diagnostic indicators were calculated by treating the final diagnosis as reference standard and using all knees, including OA, no OA and uncertain diagnoses, in the calculations; * Percentages (%) of patients of each subgroup; † Knees diagnosed as "consensus based uncertain".

For SPs, diagnostics certainty improved somewhat in the "consistent OA'" knees (70 ± 12 vs. 77 ± 15, $p < 0.001$). Diagnostic certainty of "consistent no OA" was minimally but significantly altered after viewing radiographic data (SP 20 ± 13 vs. 21 ± 15, $p = 0.04$). Sensitivity analysis showed very similar results (see Table S2).

4. Discussion

In this study, we showed that radiographs added only to the diagnostic ability of both GPs and SPs in clinically "uncertain" diagnoses. Overall, diagnostic ability, diagnostic certainty and the added value of radiographs were very similar for GP and SP.

Both GP and SP amended some of their diagnoses after viewing the radiographic data, but the majority of diagnoses remained the same. As a time-consuming, costly and potentially radiation-hazardous examination, radiographs seem to be redundant in most cases suspicious for knee OA. The diagnostic abilities of GP and SP, without access to radiographic data, were already comparable to findings in other chronic musculoskeletal diseases, such as lumbar spinal stenosis [14] and lumbar disc herniation [15]. Therefore, for clinically "certain" knees, diagnostic abilities based on clinical data only should be considered as good enough, in contrast to clinically "uncertain" knees. Our results support expert recommendations and the results of previous studies [3,4,6–8], where diagnoses based on clinical findings were found to be reliable and where radiographs were deemed unnecessary for diagnosing typical KOA.

On the other hand, after viewing radiographic data, diagnostic indicators for both GP and SP were much improved in clinically "uncertain" knees. Likelihood ratios, calculated by using sensitivity and specificity, can directly reflect the ability of diagnosing OA/no OA [16,17]. LR+ is deemed clinically meaningful if greater than 10 and LR− when lower than 0.1. In this study, radiographs helped to improve LR+ from 2 to 5 in "uncertain" knees. According to a previous literature report [17], it indicates the probability of correct diagnosis for OA knees was increased by 15%. LR- was improved from 0.4 to 0.05. The probability of a correct diagnosis of no OA knees was increased by 25%. Hence, we believe the improvements in "uncertain" knees are clinically meaningful, and radiographs could be considered in these cases.

Both GPs and SPs seemed to be more certain of their radiographically confirmed OA ("consistent OA") diagnoses. However, as other joint diseases were excluded from the CHECK cohort at baseline [11,13] and the incidence of these diseases during follow-up was quite low (3%), all the abnormalities presented in radiographic data would direct experts to an OA diagnosis, rather than to other conditions. In other words, our results could be inflated compared to real practice. Furthermore, because this is the first study of its kind, it remains unclear whether the certainty improvements are clinically relevant. On the other hand, our results did not support the strategy of using radiographs for improving certainty of no OA diagnoses. On average, the experts were already "fairly" certain (certainty scores < 30) about clinically no OA diagnoses and neither GPs nor SPs became more certain after viewing radiographic data in consistent no OA knees.

In this study, we provided standardized radiographic scores to the experts, which should be helpful to diminish the bias of image reading skill differences between different experts. Even though actual films were also available if required, not all films were viewed by experts. SP seemed to be more interested in the actual films than GP in this study. This aligns with their characteristics and also can be explained by their differences in image interpretation skills, which is correlated with image exposure in daily clinical work [18,19].

Since the major aim of this study was to evaluate the added value of radiographs above clinical findings in diagnosing KOA, we did not perform specific statistical analysis on the diagnostic results between GPs and SPs. Generally, SPs amended slightly more diagnoses than GPs after viewing radiographic data (18% vs. 14%), which could be explained by the expert characteristics as SPs place more emphasis on radiographs, but the rate of correct amendment was similar (9% vs. 8%). Furthermore, there was no obvious difference among diagnostic indicators between GPs and SPs

either before or after viewing radiographic data. Similar results were also found in certainty scores. Therefore, we believe the added value of radiographs should be considered as similar for GPs and SPs.

This study has limitations. First, there is likely some incorporation bias when we compare the GPs' and SPs' diagnoses to the consensus-based final diagnoses, because the individual diagnoses which both experts agreed on were incorporated into final diagnosis automatically [20]. That means the absolute values of diagnostic indicators in these comparisons are potentially overestimated. Decary et al. reported that the amount of overestimation of sensitivity and specificity caused by incorporation bias depended on the true specificity of the test method [21]. It was impossible to quantify the overestimation in the current study, due to the lack of the true specificity of expert diagnoses. A second potential concern, 424 patients with bilateral knee complaints were included in this study. Two knees from the same patient shared the same demographic data and WOMAC scales. In principle, it is inappropriate to view them as fully independent observations. However, our sensitivity analysis limited to left knee data only yielded results similar to the main analysis, suggesting this is not a problem in our dataset. Third, standard radiographic scores (i.e., Kellgren and Lawrence grade) as well as actual films were provided to experts in this study, which differs from the scenario of routine clinical work. Even so, most clinical diagnoses remained same after viewing both the scores and films. This, to some degree, supports our conclusion that radiographs seem to be redundant in most cases. Fourth, we did not provide the skyline view to the experts, so some patellofemoral joint OA might be missed. However, we believe any influence on our conclusions is limited because the prevalence of patellofemoral joint OA in the CHECK cohort was quite low (4.6%) [22], and its presence would also have been suggested by the lateral radiograph findings as well as clinical history and physical examination, e.g., knee crepitus [23]. Fifth, the process of obtaining final diagnosis could have been influenced by authority, since SPs likely have more authority than GPs. In this case, diagnostic indicators of SPs would be higher than those of GPs. As the diagnostic indicators were quite similar between GPs and SPs, we believe this is not a big issue.

In conclusion, radiography could be of importance in cases where the clinical diagnosis of KOA is uncertain. Radiographs helped to improve the certainty of OA diagnoses, but the clinical relevance of this improvement is unclear. Overall, all results were similar for GPs and SPs.

Supplementary Materials: The following are available online at http://www.mdpi.com/2077-0383/9/10/3374/s1. Table S1: Clinical and radiographic data; Table S2: Expert diagnoses and certainty scores (left knees only).

Author Contributions: Q.W.: Conception and design, analysis and interpretation of the data, drafting of the article, final approval of the article. J.R.: Conception and design, analysis and interpretation of the data, critical revision of the article for important intellectual content, final approval of the article, collection and assembly of data, obtaining of funding. M.K.: Conception and design, interpretation of the data, critical revision of the article for important intellectual content, final approval of the article. M.B.: Conception and design, statistical expertise, table and figure formatting, interpretation of the data, critical revision of the article for important intellectual content, final approval of the article. J.W.J.B.: Conception and design, interpretation of the data, critical revision of the article for important intellectual content, final approval of the article. S.M.A.B.-Z.: Conception and design, interpretation of the data, critical revision of the article for important intellectual content, final approval of the article, obtaining of funding, administrative support. The CREDO experts group: Conception and design, collection and assembly of data, final approval of the article. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Dutch Arthritis Society (Project ID 15-1-301); Qiuke Wang was financed by China Scholarship Council (CSC) (grant number: 201906230308).

Acknowledgments: We would like to acknowledge the CREDO experts group (N.E. Aerts-Lankhorst, R. Agricola, A.N. Bastick, R.D.W. van Bentveld, P.J. van den Berg, J. Bijsterbosch, A. de Boer, M. Boers, A.M. Bohnen, A.E.R.C.H. Boonen, P.K. Bos, T.A.E.J. Boymans, H.P. Breedveldt-Boer, R.W. Brouwer, J.W. Colaris, J. Damen, G. Elshout, P.J. Emans, W.T.M. Enthoven, E.J.M. Frölke, R. Glijsteen, H.J.C. van der Heide, A.M. Huisman, R.D. van Ingen, M.L. Jacobs, R.P.A. Janssen, P.M. Kevenaar, M. Kloppenburg, M.A. van Koningsbrugge, P. Krastman, N.O. Kuchuk, M.L.A. Landsmeer, W.F. Lems, H.M.J. van der Linden, R. van Linschoten, E.A.M. Mahler, B.L. van Meer, D.E. Meuffels, W.H. Noort-van der Laan, J.M. van Ochten, J. van Oldenrijk, G.H.J. Pols, T.M. Piscaer, J.B.M. Rijkels-Otters, N. Riyazi, J.M. Schellingerhout, H.J. Schers, B.W.V. Schouten, G.F. Snijders, W.E. van Spil, S.A.G. Stitzinger, J.J. Tolk, Y.D.M. van Trier, M. Vis, V.M.I Voorbrood, B.C. de Vos, and A. de Vries) for evaluating the medical files and for their feedback on the manuscript.

Conflicts of Interest: Q.W., J.R., M.K., J.W.J.B., S.M.A.B.-Z. report grant from the Dutch Arthritis Association. MK reports fee for consultancy (Abbvie, Pfizer, Levicept, GlaxoSmithKline, Merck-Serono, Kiniksa, Flexion), CHDR and local investigator of industry-driven trial (Abbvie), a grant from Pfizer, royalty for Wolters Kluwer (UptoDate), Springer Verlag (Reumatologie en klinische immunologie). S.M.A.B.-Z. reports consulting for Phizer, as an invited speaker in Osteoarthritis Research Society International (OARSI).

References

1. Brand, C.A.; Harrison, C.; Tropea, J.; Hinman, R.S.; Britt, H.; Bennell, K. Management of osteoarthritis in general practice in australia. *Arthritis Care Res.* **2014**, *66*, 551–558. [CrossRef] [PubMed]
2. Demehri, S.; Guermazi, A.; Kwoh, C.K. Diagnosis and longitudinal assessment of osteoarthritis: Review of available imaging techniques. *Rheum. Dis. Clin. N. Am.* **2016**, *42*, 607–620. [CrossRef] [PubMed]
3. Zhang, W.; Doherty, M.; Peat, G.; Bierma-Zeinstra, M.A.; Arden, N.K.; Bresnihan, B.; Herrero-Beaumont, G.; Kirschner, S.; Leeb, B.F.; Lohmander, L.S.; et al. Eular evidence-based recommendations for the diagnosis of knee osteoarthritis. *Ann. Rheum. Dis.* **2010**, *69*, 483–489. [CrossRef]
4. Parsons, C.; Fuggle, N.R.; Edwards, M.H.; Goulston, L.; Litwic, A.E.; Jagannath, D.; van der Pas, S.; Cooper, C.; Dennison, E.M.; Group, E.R. Concordance between clinical and radiographic evaluations of knee osteoarthritis. *Aging Clin. Exp. Res.* **2018**, *30*, 17–25. [CrossRef]
5. Decary, S.; Ouellet, P.; Vendittoli, P.A.; Roy, J.S.; Desmeules, F. Diagnostic validity of physical examination tests for common knee disorders: An overview of systematic reviews and meta-analysis. *Phys. Ther. Sport* **2017**, *23*, 143–155. [CrossRef] [PubMed]
6. Decary, S.; Feldman, D.; Fremont, P.; Pelletier, J.P.; Martel-Pelletier, J.; Fallaha, M.; Pelletier, B.; Belzile, S.; Sylvestre, M.P.; Vendittoli, P.A.; et al. Initial derivation of diagnostic clusters combining history elements and physical examination tests for symptomatic knee osteoarthritis. *Musculoskelet. Care* **2018**, *16*, 370–379. [CrossRef] [PubMed]
7. Sakellariou, G.; Conaghan, P.G.; Zhang, W.; Bijlsma, J.W.J.; Boyesen, P.; D'Agostino, M.A.; Doherty, M.; Fodor, D.; Kloppenburg, M.; Miese, F.; et al. Eular recommendations for the use of imaging in the clinical management of peripheral joint osteoarthritis. *Ann. Rheum. Dis.* **2017**, *76*, 1484–1494. [CrossRef]
8. Hunter, D.J.; Bierma-Zeinstra, S.M.A. Osteoarthr. *Lancet* **2019**, *393*, 1745–1759. [CrossRef]
9. Bierma-Zeinstra, S.; van Middelkoop, M.; Runhaar, J.; Schiphof, D. Nonpharmacological and nonsurgical approaches in OA. *Best Pract. Res. Clin. Rheumatol.* **2020**, 101564. [CrossRef]
10. Kellgren, J.H.; Lawrence, J.S. Radiological assessment of osteo-arthrosis. *Ann. Rheum. Dis.* **1957**, *16*, 494–502. [CrossRef]
11. Schiphof, D.; Runhaar, J.; Waarsing, J.H.; van Spil, W.E.; van Middelkoop, M.; Bierma-Zeinstra, S.M.A. The clinical and radiographic course of early knee and hip osteoarthritis over 10 years in check (cohort hip and cohort knee). *Osteoarthr. Cartil.* **2019**, *27*, 1491–1500. [CrossRef]
12. Damen, J.; Schiphof, D.; Wolde, S.T.; Cats, H.A.; Bierma-Zeinstra, S.M.; Oei, E.H. Inter-observer reliability for radiographic assessment of early osteoarthritis features: The check (cohort hip and cohort knee) study. *Osteoarthr. Cartil.* **2014**, *22*, 969–974. [CrossRef] [PubMed]
13. Wesseling, J.; Boers, M.; Viergever, M.A.; Hilberdink, W.K.; Lafeber, F.P.; Dekker, J.; Bijlsma, J.W. Cohort profile: Cohort hip and cohort knee (check) study. *Int. J. Epidemiol.* **2016**, *45*, 36–44. [CrossRef]
14. Cook, C.J.; Cook, C.E.; Reiman, M.P.; Joshi, A.B.; Richardson, W.; Garcia, A.N. Systematic review of diagnostic accuracy of patient history, clinical findings, and physical tests in the diagnosis of lumbar spinal stenosis. *Eur. Spine J.* **2019**. [CrossRef]
15. Petersen, T.; Laslett, M.; Juhl, C. Clinical classification in low back pain: Best-evidence diagnostic rules based on systematic reviews. *BMC Musculoskelet. Disord.* **2017**, *18*, 188. [CrossRef]
16. Thornbury, J.R.; Fryback, D.G.; Edwards, W. Likelihood ratios as a measure of the diagnostic usefulness of excretory urogram information. *Radiology* **1975**, *114*, 561–565. [CrossRef]
17. McGee, S. Simplifying likelihood ratios. *J. Gen. Intern. Med.* **2002**, *17*, 646–649. [CrossRef]
18. Esserman, L.; Cowley, H.; Eberle, C.; Kirkpatrick, A.; Chang, S.; Berbaum, K.; Gale, A. Improving the accuracy of mammography: Volume and outcome relationships. *J. Natl. Cancer Inst.* **2002**, *94*, 369–375. [CrossRef]
19. Taruc-Uy, R.L.; Lynch, S.A. Diagnosis and treatment of osteoarthritis. *Prim. Care* **2013**, *40*, 821–836. [CrossRef]
20. Worster, A.; Carpenter, C. Incorporation bias in studies of diagnostic tests: How to avoid being biased about bias. *Can. J. Emerg. Med.* **2008**, *10*, 174–175. [CrossRef]

21. Karch, A.; Koch, A.; Zapf, A.; Zerr, I.; Karch, A. Partial verification bias and incorporation bias affected accuracy estimates of diagnostic studies for biomarkers that were part of an existing composite gold standard. *J. Clin. Epidemiol.* **2016**, *78*, 73–82. [CrossRef]
22. Lankhorst, N.E.; Damen, J.; Oei, E.H.; Verhaar, J.A.N.; Kloppenburg, M.; Bierma-Zeinstra, S.M.A.; van Middelkoop, M. Incidence, prevalence, natural course and prognosis of patellofemoral osteoarthritis: The cohort hip and cohort knee study. *Osteoarthr. Cartil.* **2017**, *25*, 647–653. [CrossRef]
23. Van Middelkoop, M.; Bennell, K.L.; Callaghan, M.J.; Collins, N.J.; Conaghan, P.G.; Crossley, K.M.; Eijkenboom, J.; van der Heijden, R.A.; Hinman, R.S.; Hunter, D.J.; et al. International patellofemoral osteoarthritis consortium: Consensus statement on the diagnosis, burden, outcome measures, prognosis, risk factors and treatment. *Semin. Arthritis Rheum.* **2018**, *47*, 666–675. [CrossRef]

Journal of
Clinical Medicine

MDPI

Article

Generalized Joint Hypermobility and Injuries: A Prospective Cohort Study of 185 Pre-Professional Contemporary Dancers

Rogier M. van Rijn [1,2,*] and Janine H. Stubbe [1,2,3,4]

[1] Codarts Rotterdam, University of the Arts, 3012 CC Rotterdam, The Netherlands; jhstubbe@codarts.nl
[2] PErforming Artist and Athlete Research Lab (PEARL), 3012 CC Rotterdam, The Netherlands
[3] Erasmus MC University Medical Center, Department of General Practice, 3015 GD Rotterdam, The Netherlands
[4] Rotterdam Arts and Science Lab (RASL), 3012 CC Rotterdam, The Netherlands
* Correspondence: rmvanrijn@codarts.nl; Tel.: +31-6-25173419

Abstract: Generalized joint hypermobility (GJH) has been mentioned as one of the factors associated with dance injuries, but the findings are inconclusive. This study aims to investigate whether GJH, based on different Beighton score cut-off points, is a potential risk factor for injuries in pre-professional dancers. Four cohorts of first-year pre-professional dancers ($N = 185$), mean age 19.1 ± 1.3 years, were screened on musculoskeletal functioning at the start of their academic year. The Beighton score was used to measure GJH. During the academic year, the dancers completed monthly questionnaires about their physical and mental health. Based on the Oslo Sports Trauma Research Centre Questionnaire on Health Problems (OSTRC), three injury definitions were used (i.e., all complaints, substantial injury, and time-loss injury). To examine potential risk factors for injuries, univariate and multivariate regression models were applied. The response rate of monthly completed questionnaires was 90%. The overall mean (SD) Beighton score was 2.8. The 1-year injury incidence proportion was 67.6% ($n = 125$), 43.2% ($n = 80$), and 54.6% ($n = 101$) for all complaint injuries, substantial injuries, and time-loss injuries, respectively. The multivariate analyses showed a significant association between a previous long lasting injury in the past year and the three injury definitions ($p < 0.05$). Pre-professional contemporary dancers are at high risk for injuries and hypermobility. However, these two variables are not associated with each other. Health professionals should take injury history into account when assessing dance students, because this variable is associated with increased injury risk.

Keywords: injury; risk factor; dance; hypermobility; physical examination

Citation: van Rijn, R.M.; Stubbe, J.H. Generalized Joint Hypermobility and Injuries: A Prospective Cohort Study of 185 Pre-Professional Contemporary Dancers. *J. Clin. Med.* **2021**, *10*, 1007. https://doi.org/10.3390/jcm10051007

Academic Editor: Jos Runhaar

Received: 14 January 2021
Accepted: 23 February 2021
Published: 2 March 2021

1. Introduction

At an early age, dancers participate in long hours of training, classes, and performances to improve their aesthetic, technical, and athletic skills. Artistic, physical, and psychological demands are being pushed when the performance levels of dancers increase [1]. These demands make young dancers prone to injuries, especially during periods of maturation and development [2–4]. Injury incident rates within pre-professional and/or contemporary dancers ranged from 0.77 to 4.71 per 1000 h of exposure [3,5]. Insight into the risk factors of dance injuries can give direction to the development of preventive strategies [6,7]. However, a systematic review on risk factors in pre-professional dancers concluded that the level of evidence is scarce [3]. High-quality prospective studies are needed to unravel the multifactorial association between risk factors and dance injuries.

Increased joint flexibility, also known as generalized joint hypermobility (GJH), has been mentioned as one of the factors associated with injuries. GJH is characterized by excessive range of motion in multiple joints that exceeds normal limits, and is assessed by the Beighton score [8,9]. Prevalence of GJH in dance populations ranges between 2% and 44% [10]. This huge variation in prevalences is due to the use of different cut-off points to measure GJH. Although pre-professional dancers with increased joint flexibility

are often marked as dancers being "full of potential", GJH in dancers might have serious disadvantages [11,12]. GJH can be viewed as a sign of vulnerability in terms of physical fitness, psychological distress, and more musculoskeletal complaints (i.e., pain, fatigue) in dancers [13,14]. GJH has been associated with increased risk of (ligament) injuries, recurrent dislocations, knee and ankle effusions, and possible premature arthritis [9,10,15,16]. On the other hand, studies found that GJH was not predictive of injury or total days injured [11,17]. These inconclusive findings might be explained by the lack of high-quality studies [3]. Studies with sufficient power, prospective designs, and uniform injury definitions are needed [3]. Therefore, the objective of this study was to investigate whether GJH, based on different Beighton score cut-off points, is a potential risk factor for injuries in pre-professional dancers.

2. Methods

2.1. Participants and Procedures

Four cohorts of first-year pre-professional dancers (N = 185) of Codarts Rotterdam (University of the Arts) were screened on musculoskeletal functioning at the start of the their academic year in 2016, 2017, 2018, and 2019. During the first month of the academic year, baseline characteristics were recorded including age (years), sex (male/female), educational program (Bachelor Dance or Dance teacher), and injury history (yes/no). Injury history was defined as "any physical complaint resulting in a fulltime loss of dance activities (participation in class, rehearsal, performance practice, and so on) for at least one week beyond the day of onset in the past year" [18,19]. During the academic year, all pre-professional dancers were asked to complete monthly questionnaires about their physical and mental health by using the performing artist and athlete health monitor (PAHM). All pre-professional dancers were informed about the purpose and procedure of this study and provided written informed consent. The study was approved by the Medical Ethics Committee (MEC-2019-0163) of the Erasmus University Medical Center Rotterdam, The Netherlands.

2.2. Generalized Joint Hypermobility

Generalized joint hypermobility (GJH) testing was performed during the musculoskeletal screening at the start of the academic year and was done by the physical therapist, specialized in dance medicine, who works at the Performing Arts Health Centre of the university. The Beighton score was used to measure GJH and consists of nine tests with one point allocated for each positive finding: (1) passive dorsiflexion of the fifth metacarpophalangeal joints >90° (one point per side), (2) passive apposition of the thumb to the flexor aspects of the forearm (one point per side), (3) hyperextension of the elbows beyond 10° (one point per side), (4) hyperextension of the knee beyond 10° (one point per side), and (5) forward flexion of the trunk with knees fully extended so that the palms of the hands rest flat on the floor (one point) [8]. The Beighton score was calculated by summing the scores of the five tests, resulting in total scores ranging from 0 (hypomobile) to 9 (hypermobile).

2.3. Injury Registration

The Oslo Sport Trauma Research Center (OSTRC) questionnaire is part of the monthly questionnaire and consists of four key questions about the consequences of health problems on participation, training volume, and performance as well as the degree to which the student perceived symptoms. All items ranged from 0 (no problem, no reduction, no effect, or no symptoms) to 25 (cannot participate at all or severe symptoms) [20]. Questions 1 and 4 were scored on a four-point scale (0–8–17–25), while questions 2 and 3 were scored on a five-point scale (0–6–13–19–25). The severity of a health problem was calculated by the sum score of the four questions (scale 0–100) according to the method proposed by Clarsen et al. [20]. If the severity score was higher than zero, a health problem was registered and the student was asked whether the health problem was an injury, mental problem, or other problem. For injuries, the pre-professional dancer was automatically

directed to an injury registration form based on an international consensus statement on injury surveillance methodology for football to collect further details (e.g., location, history, and acute or overuse onset) [21]. If a pre-professional dancer reported the same injury as the most severe health problem in two or more consecutive questionnaires, this was counted as one 'unique' case of a (fluctuating) problem [5,22].

2.4. Injury Definitions

Three definitions of dance related injuries were utilized:

1. All complaints injury: any physical complaint resulting in a severity score higher than zero on the OSTRC questionnaire irrespective of the need for medical attention or time loss from dance activities [23].
2. Substantial injury: any physical complaint resulting in a severity score higher than zero on the OSTRC questionnaire irrespective of the need for medical attention and resulting in problems leading to moderate or severe reductions (value $\geq$13 on question 2 or 3 of the OSTRC) in training volume or moderate or severe reductions in performance or complete inability to participate in dance [23].
3. Time-loss injury: any physical complaint resulting in a severity score higher than zero on the OSTRC questionnaire and resulting in a dancer not being able to complete a class, rehearsal, or performance or a subsequent class, rehearsal, or performance for one or more days beyond the day of onset [24].

2.5. Statistical Analysis

All statistical procedures were performed using SPSS (SPSS, V25.0, IBM, Chicago, IL, USA, 2017) and the statistical significance level was set at an alpha level of <0.05. Descriptive statistics were used to describe the baseline characteristics of all participants using mean values and standard deviation (SD) or number and percentages (%).

The 1-year incidence proportion (IP) of all complaints injuries, substantial injuries, and time-loss injuries was calculated by dividing the number of students who reported at least one injury during the academic year by the number of respondents. The severity of injuries was calculated by the mean (SD) number of full days that a student completely/partly missed their dance activities because of their injury. The characteristics (i.e., location) of injuries were expressed in percentages for the total injuries.

To examine potential risk factors for injuries, univariate and multivariate regression models were applied. Potential risk factors included age (years), sex (female), educational program (Bachelor Dance vs. Bachelor Dance), injury history in the previous year, and generalized joint hypermobility. Analyses were performed in four ways using a different measure of generalized joint hypermobility each time, namely, (1) Beighton score (mean), (2) Beighton cut-off point <3 (not hypermobile), 4–6 (hypermobile), >6 (extreme hypermobile) [25]; (3) Beighton cut-off point $\geq$4 (hypermobile) [26]; and (4) Beighton cut-off point $\geq$6 (hypermobile) [26]. First, univariate associations between the potential risk factors and the dichotomized outcome, injured during follow-up (yes/no), were assessed using the three injury definitions. Second, multivariate regression modeling was performed including all potential risk factors and the outcome of interest, resulting in four models (for each measure of GJH) per injury definition. The results of the regression analyses were expressed in odds ratios (ORs) with corresponding 95% confidence interval (95% CI). As logistic regression analyses do not show the proportion of variance that was explained by the model, the Nagelkerke R^2 was used to express this variance [27]. The Nagelkerke R^2 varies between -1 and $+1$. A positive value indicates that, as the value of the risk factor increases, so does the likelihood of incurring an injury. A negative value implies that, as the value of the risk factor increases, the likelihood of the outcome occurring decreases. If a risk factor has a small value of R^2, it contributes only a small amount to the model [27].

3. Results

3.1. Participants

All first year students of the Bachelor Dance and Bachelor Dance teacher programs agreed to participate and were consequently included in the present study, resulting in a participation rate of 100%. The cohort comprised 185 students (68.6% females) and the mean age was 19.1 (1.3) years (Table 1). In total, 1665 questionnaires were sent to the students and 1499 were completed, resulting in a response rate of 90.0%.

Table 1. Baseline characteristics of the total study population and separated for the Bachelor Dance program and the Bachelor Dance teacher program.

	Total Population (*N* = 185)	Dance (*n* = 117)	Dance Teacher (*n* = 68)
Sex (female)	127 (68.6%)	63 (53.8%)	64 (94.1%)
Age (years)	19.1 ± 1.3	18.8 ± 0.9	19.4 ± 1.9
Previous long-lasting injury (yes) #	60 (33.3%)	29 (25.9%)	31 (45.6%)
Generalized joint hypermobility *			
Beighton score	2.8 ± 2.3	3.2 ± 2.5	2.2 ± 2.0
Not hypermobile (0–3)	129 (70.9%)	76 (65.5%)	53 (80.3%)
Hypermobile (4–6)	25 (13.7%)	17 (14.7%)	8 (12.1%)
Extreme hypermobile (7–9)	28 (15.4%)	23 (19.8%)	5 (7.6%)
Not hypermobile (0–3)	129 (70.9%)	76 (65.5%)	53 (80.3%)
Hypermobile (4–9)	53 (29.1%)	40 (34.5%)	13 (19.7%)
Not hypermobile (0–5)	154 (84.6%)	93 (80.2%)	61 (92.4%)
Hypermobile (6–9)	28 (15.4%)	23 (19.8%)	5 (7.6%)

Data are presented as mean (SD) or number (%). # missing data of *n* = 5 (5 dance), * missing data of *n* = 3 (1 Dance, 2 Dance teacher).

3.2. Generalized Joint Hypermobility

The overall mean (SD) Beighton score for the whole cohort was 2.8 (2.3), with students from the Bachelor Dance program scoring significantly higher than students of the Bachelor Dance teacher program, at 3.2 (2.5) versus 2.2 (2.0), respectively (Table 1). Using a Beighton score cut-off point of 0–3 (not hypermobile), 4–6 (hypermobile), and 7–9 (extreme hypermobile), a total of 25 (13.7%) students were hypermobile and 28 students (15.4%) were extreme hypermobile. No significant difference was found between both Bachelor programs using this cut-off points. Applying the criterion of a Beighton score $\geq$4/9 resulted in 53 (29.1%) students having GJH. Of these, 40 students were of the Bachelor Dance program versus 13 of the Bachelor Dance teacher program, resulting in a significant difference between both educational programs. Applying the criterion of a Beighton score $\geq$6/9 resulted in 28 (15.4%) students having GJH. There was a significant difference between the two Bachelor programs in the proportion (19.8% vs. 7.6%) meeting hypermobility at this cut-off point.

3.3. Injuries

The injury incidence proportion for one academic year was 67.6% (*n* = 125), 43.2% (*n* = 80), and 54.6% (*n* = 101) for all complaint injuries, substantial injuries, and time-loss injuries, respectively. A total of 285 all complaints injuries were reported by 125 students. The mean number of full days a student completely/partly missed dance activities as a result of an all complaint injury was 9.7 ± 19.8 days. A total of 188 substantial injuries were reported by 80 students. The mean number of full days a student completely/partly missed dance activities as a result of an substantial injury was 13.6 ± 23.7 days. A total of 244 time-loss injuries were reported by 101 students. The mean number of full days a student completely/partly missed dance activities as a result of an time-loss injury was

12.0 ± 21.4 days. For all three injury definitions, the most reported injury locations were knee, ankle, lower back, lower leg (front side), and the foot/toe (Figure 1).

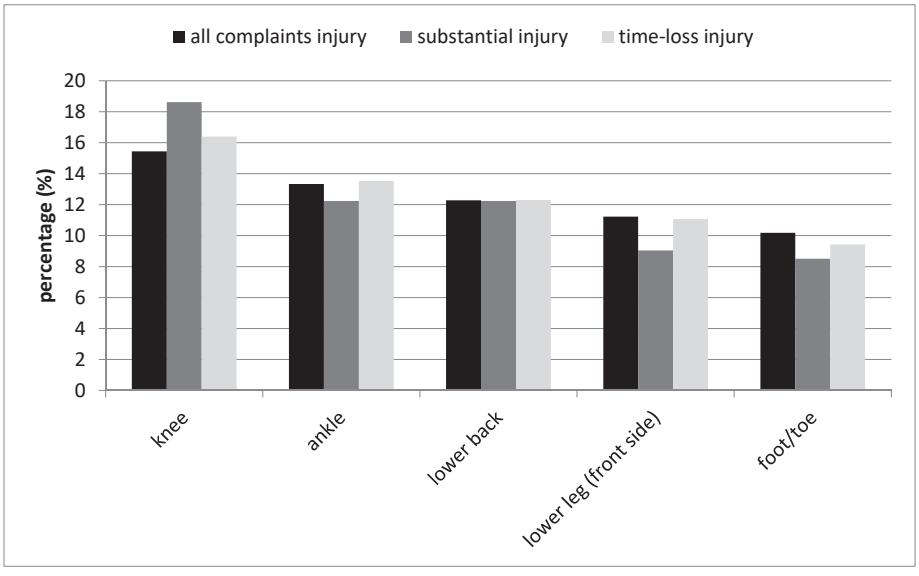

Figure 1. Top five most reported injury locations according to the three injury definitions.

3.4. Risk Factors

Univariate analyses showed a significant association between a previous long-lasting injury in the past year and all complaints injuries (OR 2.58; 95% CI 1.24–5.35), substantial injuries (OR 3.22; 95% CI 1.69–6.13), and time-loss injuries (OR 3.09; 95% CI 1.59–6.02) during follow-up. Besides, univariate analyses showed that being a student enrolled in the Bachelor Dance teacher program is significantly associated with the occurrence of substantial injuries during follow-up (OR 1.87; 95% CI 1.02–3.42). None of the four operationalized definitions of GJH and none of the other tested variables were univariate associated with the outcomes of interest (Tables 2–4). The multivariate analyses showed a significant association between a previous long-lasting injury in the past year and all complaints injuries (range ORs 2.60 to 2.74), substantial injuries (range ORs 2.76 to 2.96), and time-loss injuries (range ORs 2.94 to 3.04) during follow-up. None of the four operationalized definitions of GJH and none of the other tested variables were associated with the outcomes of interest in the multivariate analyses (Tables 2–4). The models (for each measure of GJH) explained 5.7% to 6.0% of the variance when using the all complaint injury definition, 11.0% to 11.5% for substantial injuries, and 9.8% to 9.9% when using the time-loss injury definition.

Table 2. Univariate and multivariate models of potential risk factors for all complaints injuries.

	Non-Injured (n = 60)	Injured (n = 125)	Univariate Analysis OR (95% CI)	Multivariate Analysis OR (95% CI)			
Sex (female)	44 (73.3%)	83 (66.4%)	0.72 (0.36–1.42)	0.75 (0.33–1.70)	0.65 (0.29–1.49)	0.67 (0.30–1.53)	0.71 (0.32–1.56)
Age (years)	19.0 ± 1.15	19.1 ± 1.43	1.08 (0.85–1.39)	1.04 (0.80–1.35)	1.04 (0.80–1.35)	1.04 (0.80–1.35)	1.04 (0.80–1.35)
Educational Program							
Bachelor dance	37 (61.7%)	80 (64.0%)	Ref.	Ref.	Ref.	Ref.	Ref.
Bachelor dance teacher	23 (38.3%)	45 (36.0%)	0.91 (0.48–1.71)	0.84 (0.37–1.88)	0.93 (0.42–2.06)	0.94 (0.42–2.07)	0.88 (0.40–1.92)
Previous long-lasting injury (yes) #	12 (20.3%)	48 (39.7%)	**2.58 (1.24–5.35)**	**2.74 (1.26–5.94)**	**2.69 (1.24–5.82)**	**2.60 (1.21–5.59)**	**2.68 (1.24–5.81)**
Generalized joint hypermobility *							
Beighton score (0–9)	2.95 ± 2.23	2.78 ± 2.40	0.97 (0.85–1.11)	0.96 (0.82–1.12)			
Not hypermobile (0–3)	41 (71.9%)	88 (70.4%)	Ref.		Ref.		
Hypermobile (4–6)	7 (12.3%)	18 (14.4%)	1.20 (0.46–3.09)		1.42 (0.51–3.90)		
Extreme hypermobile (7–9)	9 (15.8%)	19 (15.2%)	0.98 (0.41–2.36)		0.94 (0.35–2.49)		
Not hypermobile (0–3)	41 (71.9%)	88 (70.4%)	Ref.			Ref.	
Hypermobile (4–9)	16 (28.1%)	37 (29.6%)	1.08 (0.54–2.16)			1.15 (0.53–2.49)	
Not hypermobile (0–5)	48 (84.2%)	106 (84.8%)	Ref.				Ref.
Hypermobile (6–9)	9 (15.8%)	19 (15.2%)	0.96 (0.40–2.27)				0.87 (0.34–2.27)
Nagelkerke R^2				0.059	0.060	0.057	0.057

Data are presented as mean (SD) or number (%). # missing data of n = 5 (5 students of Bachelor of Dance), * missing data of n = 3 (1 student of Bachelor of Dance, 2 students of Bachelor of Dance Teacher), Ref. = reference; OR = odds ratio; CI = confidence interval. Bold, significant associations.

Table 3. Univariate and multivariate models of potential risk factors for substantial injuries.

	Non-Injured (n = 105)	Injured (n = 80)	Univariate Analysis OR (95% CI)	Multivariate Analysis OR (95% CI)			
Sex (female)	69 (65.7%)	58 (72.5%)	1.38 (0.73–2.60)	1.14 (0.52–2.50)	0.97 (0.44–2.15)	1.01 (0.46–2.21)	1.04 (0.49–2.22)
Age (years)	19.04 ± 1.27	19.08 ± 1.44	1.02 (0.82–1.27)	0.95 (0.75−1.20)	0.95 (0.75–1.20)	0.95 (0.75–1.20)	0.95 (0.75–1.20)
Educational Program							
Bachelor dance	73 (69.5%)	44 (55.0%)	Ref.	Ref.	Ref.	Ref.	Ref.
Bachelor dance teacher	32 (30.5%)	36 (45.0%)	**1.87 (1.02–3.42)**	1.57 (0.73–3.39)	1.78 (0.83–3.82)	1.79 (0.84–3.83)	1.70 (0.81–3.58)
Previous long-lasting injury (yes) #	23 (22.3%)	37 (48.1%)	**3.22 (1.69–6.13)**	**2.96 (1.50–5.84)**	**2.85 (1.45–5.61)**	**2.76 (1.42–5.37)**	**2.84 (1.45–5.59)**
Generalized joint hypermobility *							
Beighton score (0–9)	2.92 ± 2.38	2.71 ± 2.31	0.96 (0.85–1.09)	0.94 (0.81–1.08)			
Not hypermobile (0–3)	73 (71.6%)	56 (70.0%)	Ref.		Ref.		
Hypermobile (4–6)	13 (12.7%)	12 (15.0%)	1.20 (0.51–2.84)		1.33 (0.51–3.43)		
Extreme hypermobile (7–9)	16 (15.7%)	12 (15.0%)	0.98 (0.43–2.23)		0.90 (0.35–2.27)		
Not hypermobile (0–3)	73 (71.6%)	56 (70.0%)	Ref.			Ref.	
Hypermobile (4–9)	29 (28.4%)	24 (30.0%)	1.08 (0.57–2.05)			1.08 (0.52–2.25)	
Not hypermobile (0–5)	86 (84.3%)	68 (85.0%)	Ref.				Ref.
Hypermobile (6–9)	16 (15.7%)	12 (15.0%)	0.95 (0.42–2.14)				0.84 (0.34–2.10)
Nagelkerke R^2				0.115	0.113	0.110	0.110

Data are presented as mean (SD) or number (%). # missing data of $n = 5$ (5 students of Bachelor of Dance), * missing data of $n = 3$ (1 student of Bachelor of Dance, 2 students of Bachelor of Dance Teacher), Ref. = reference; OR = odds ratio; CI = confidence interval. Bold, significant associations.

Table 4. Univariate and multivariate models of potential risk factors for time-loss injuries.

	Non-Injured (n = 84)	Injured (n = 101)	Univariate Analysis OR (95% CI)	Multivariate Analysis OR (95% CI)			
Sex (female)	60 (71.4%)	67 (66.3%)	0.79 (0.42–1.48)	0.65 (0.30–1.39)	0.60 (0.28–1.31)	0.60 (0.28–1.31)	0.62 (0.29–1.30)
Age (years)	18.94 ± 1.10	19.15 ± 1.51	1.13 (0.90–1.42)	1.05 (0.82–1.35)	1.05 (0.82–1.35)	1.05 (0.82–1.35)	1.05 (0.82–1.35)
Educational Program							
Bachelor dance	57 (67.9%)	60 (59.4%)	Ref.	Ref.	Ref.	Ref.	Ref.
Bachelor dance teacher	27 (32.1%)	41 (40.6%)	1.44 (0.79–2.65)	1.46 (0.68 −3.15)	1.57 (0.73–3.36)	1.57 (0.73–3.36)	1.54 (0.73–3.26)
Previous long-lasting injury (yes) #	17 (20.5%)	43 (44.3%)	**3.09 (1.59–6.02)**	**3.04 (1.50–6.15)**	**2.95 (1.46–5.94)**	**2.94 (1.47–5.89)**	**2.94 (1.46–5.93)**
Generalized joint hypermobility *							
Beighton score (0–9)	2.94 ± 2.29	2.74 ± 2.39	0.97 (0.85–1.09)	0.97 (0.84–1.13)			
Not hypermobile (0–3)	57 (70.4%)	72 (71.3%)	Ref.		Ref.		
Hypermobile (4–6)	12 (14.8%)	13 (12.9%)	0.86 (0.36–2.02)		1.10 (0.43–2.83)		
Extreme hypermobile (7–9)	12 (14.8%)	16 (15.8%)	1.06 (0.46–2.41)		1.09 (0.43–2.75)		
Not hypermobile (0–3)	57 (70.4%)	72 (71.3%)	Ref.			Ref.	
Hypermobile (4–9)	24 (29.6%)	29 (28.7%)	0.96 (0.50–1.82)			1.09 (0.53–2.27)	
Not hypermobile (0–5)	69 (85.2%)	85 (84.2%)	Ref.				Ref.
Hypermobile (6–9)	12 (14.8%)	16 (15.8%)	1.08 (0.48–2.44)				1.07 (0.43–2.64)
Nagelkerke R^2				0.099	0.098	0.098	0.098

Data are presented as mean (SD) or number (%). # missing data of n = 5 (5 students of Bachelor of Dance), * missing data of n = 3 (1 student of Bachelor of Dance, 2 students of Bachelor of Dance Teacher), Ref. = reference; OR = odds ratio; CI = confidence interval. Bold, significant associations.

4. Discussion

This study investigated whether GJH, measured with the Beighton score, is a risk factor for injuries in pre-professional contemporary dancers. We found an overall mean Beighton score of 2.8 ± 2.3. Using the traditional Beighton cut-off point of 4 or more resulted in a prevalence of GJH of 29.1% (n = 53). A more stringent cut-off point (≥6/9) led to a prevalence of 15.4% of the dance students having GJH.

The prevalence of GJH in the current study is within the range of 2% to 81%, as reported in previous studies [10,11,13,16,26,28]. The wide range of reported prevalences can be explained by the lack of consistency in the quality of the studies, different classification and methodology used to measure GJH, and the use of different injury definitions. Therefore, a comparison of the results with the existing literature remains difficult. In a cross sectional study, including 75 pre-professional dancers, a prevalence of GJH of 81% (Beighton cut-off point ≥4/9) and 53% (Beighton cut off point ≥6/9) was reported [26]. Armstrong and colleagues (2019) performed a prospective cohort study among eighty pre-professional [11]. They reported a mean Beighton score of 4.68 ± 1.81 and classified sixty-one dancers (74%) as hypermobile using a Beighton score of ≥4/9. A prospective cohort study by Bronner et al. (2018), in which 180 pre-professional dancers were analyzed, reported a mean Beighton score of 3.59 ± 2.08 [16]. Roussel et al. (2009) found a prevalence of GJH of 44% (Beighton cut-off point ≥4/9) in a population of thirty-two pre-professional dancers. Besides, they reported that 25% presented a Beighton score ranging from 4 to 6 (hypermobile), and 19% were excessively hypermobile (Beighton score 7–9) [29]. Although these studies used the same nine-point Beighton test, they reported a higher prevelance of GJH using the same cut-off points and higher mean Beighton scores compared with our study.

Furthermore, we found a 1-year incidence of all complaints injuries, substantial injuries, and time-loss injuries of 67.6%, 43.2%, and 54.6%, respectively. The results of the multivariate analysis showed that GJH, using one of the Beighton cut-off point variants, was not associated with the occurrence of injuries defined according to the three injury definitions. However, students with an injury history, defined as any physical complaint resulting in a fulltime loss of dance activities for at least 1 week beyond the day of onset in the past year, were more likely to sustain an all complaints injury, a substantial injury, or a time-loss injury.

Our results are in line with those of the prospective cohort study of Roussel et al. (2009) in which GJH was assessed using the Beighton test and injuries were defined as any musculoskeletal condition requiring time away from dancing [29]. They found no association between GJH and a higher injury risk. In addition, Armstrong and colleagues (2019) concluded that a total Beighton score (mean) was a weak predictor of injuries resulting in absence from dancing for one or more days (time-loss) [11]. In contrast, a prospective study of Bronner et al. found that low (0–2) and high (5–9) Beighton scores were significant predictors for medical attention injuries (injuries evaluated by a physical therapist) and time-loss injuries (subset of medical attention injury that involved one or more days of time-loss from dancing following the event) [16].

Strengths and Limitations of the Study

The first major strength of this study was the high participation rate, low drop-out rate, and high response rate to the monthly questionnaires. This might be explained by two reasons. The high response rate resulted in a large sample size (N = 185), enabling us to include multiple factors into the regression models without violation of the commonly used "rule of 10" [30].

The second strength is the use of a prospective cohort design. Several studies have investigated the association between GJH and injuries using a cross-sectional design [13,14]. This design is limited to data at one time point, and can only be applied to study associations. Prospective designs are typically ranked higher in the hierarchy of evidence than cross-sectional designs. In prospective designs, study samples are followed over

time, enabling us to investigate a clear sequential relationship (i.e., causality) between an exposure, which happens first, and an outcome, which happens after [31].

The third strength of our study is the use of a broad injury definition (all complaints, substantial injuries, and time-loss injuries). Previous research has shown that the amount of dance students' injuries and their duration vary depending on the injury definition [24]. Therefore, for every injury definition, we conducted a separate regression analysis. In addition, multiple measurements of GJH were included in the analyses each time. By taking into account broad injury definitions and multiple GJH cut-off points, we analyzed the association between GJH and injury risk to the full extent.

However, there are also some limitations. First of all, in our study, all injuries were self-reported. Most dance students lack medical expertise. Therefore, detailed diagnostic information on registered injuries was missing. This limited us to distinguish between diagnoses of different injury types (for example, acute versus overuse injuries).

Secondly, we asked the student to provide injury information on a monthly basis. This long recall period may have led to underestimation of injuries. Although it is recommended to register injuries on a weekly basis, the literature shows that the average prevalence of injuries was not affected by the sampling frequency up to a frequency of one sample every 4 weeks [20]. Therefore, we do not believe that our recall period of 4 weeks has influenced the reliability of the injury registration. Finally, in the current study, health problems were assessed with the OSTRC questionnaire [20]. Since data collection finished, an update of the OSTRC questionnaire has been published in which changes to the wording, structure, and logic were introduced [32]. We do not believe that using the updated version of the OSTRC questionnaire would have resulted in different findings, because Clarsen and colleagues showed that the number of substantial health problems captured with both versions of the OSTRC questionnaire largely overlap [32].

5. Conclusions and Implications for Practice

Healthcare professionals involved in the care for pre-professional dancers should be aware of the increased injury risk and high hypermobility scores in this specific population. The Beighton score has frequently been used in practice to assess GJH and its relation to injury risk in dancers [11,16,29]. The results of the present study show that this test is not conclusive enough to define which dancers are prone to injuries. This may not be a surprise, because the literature states that there is no screening test with adequate test properties to predict injuries [33]. Our findings suggest that dancers with a history of injuries are at higher risk for sustaining a new injury (defined as all complaint injuries, substantial injuries, and time-loss injuries). This is supported by several other publications [3,16]. Therefore, we believe that health professionals should incorporate a thorough examination of the injury history in the physical examination of pre-professional dancers. The explained variance of our models was low, therefore, further research is needed to gain more insight into risk factors in order to develop preventive strategies.

Author Contributions: Conceptualization, R.M.v.R. and J.H.S.; methodology, R.M.v.R. and J.H.S.; formal analysis, R.M.v.R. and J.H.S.; writing—original draft preparation, R.M.v.R. and J.H.S.; writing—review and editing, R.M.v.R. and J.H.S.; supervision, R.M.v.R. and J.H.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Association of Applied Sciences SIA (grant number: SPR.VG01.007).

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of the Erasmus University Medical Center Rotterdam, The Netherlands (MEC-2019-0163, 26 April 2019).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Conflicts of Interest: Rogier M van Rijn and Janine H Stubbe declare that they have no conflicts of interest.

References

1. Caine, D.; Goodwin, B.J.; Caine, C.G.; Bergeron, G. Epidemiological Review of Injury in Pre-Professional Ballet Dancers. *J. Dance Med. Sci.* **2015**, *19*, 140–148. [CrossRef] [PubMed]
2. Bowerman, E.A.; Whatman, C.; Harris, N.; Bradshaw, E. A review of the risk factors for lower extremity overuse injuries in young elite female ballet dancers. *J. Dance Med. Sci.* **2015**, *19*, 51–56. [CrossRef] [PubMed]
3. Kenny, S.J.; Whittaker, J.L.; Emery, C.A. Risk factors for musculoskeletal injury in preprofessional dancers: A systematic review. *Br. J. Sports Med.* **2016**, *50*, 997–1003. [CrossRef] [PubMed]
4. Lee, L.; Reid, D.; Cadwell, J.; Palmer, P. Injury Incidence, Dance Exposure and the Use of the Movement Competency Screen (Mcs) to Identify Variables Associated with Injury in Full-Time Pre-Professional Dancers. *Int. J. Sports Phys. Ther.* **2017**, *12*, 352–370.
5. van Winden, D.; Van Rijn, R.M.; Richardson, A.; Savelsbergh, G.J.P.; Oudejans, R.R.D.; Stubbe, J.H. Detailed injury epidemiology in contemporary dance: A 1-year prospective study of 134 students. *BMJ Open Sport Exerc. Med.* **2019**, *5*, e000453. [CrossRef] [PubMed]
6. Finch, C. A new framework for research leading to sports injury prevention. *J. Sci. Med. Sport* **2006**, *9*, 3–9, discussion 10. [CrossRef]
7. van Mechelen, W.; Hlobil, H.; Kemper, H.C. Incidence, severity, aetiology and prevention of sports injuries. A review of concepts. *Sports Med.* **1992**, *14*, 82–99. [CrossRef]
8. Beighton, P.; Solomon, L.; Soskolne, C.L. Articular mobility in an African population. *Ann. Rheum. Dis.* **1973**, *32*, 413–418. [CrossRef] [PubMed]
9. Grahame, R.; Jenkins, J.M. Joint hypermobility—Asset or liability? A study of joint mobility in ballet dancers. *Ann. Rheum. Dis.* **1972**, *31*, 109–111. [CrossRef] [PubMed]
10. Day, H.; Koutedakis, Y.; Wyon, M.A. Hypermobility and dance: A review. *Int. J. Sports Med.* **2011**, *32*, 485–489. [CrossRef]
11. Armstrong, R. The Beighton Score and Injury in Dancers: A Prospective Cohort Study. *J. Sport Rehabil.* **2019**, *29*, 563–571. [CrossRef] [PubMed]
12. McCormack, M. Teaching the Hypermobile Dancer. *IADMS Bull. Teach.* **2010**, *2*, 5–8.
13. Ruemper, A.; Watkins, K. Correlations Between General Joint Hypermobility and Joint Hypermobility Syndrome and Injury in Contemporary Dance Students. *J. Dance Med. Sci.* **2012**, *16*, 161–166. [PubMed]
14. Scheper, M.C.; de Vries, J.E.; de Vos, R.; Verbunt, J.; Nollet, F.; Engelbert, R.H. Generalized joint hypermobility in professional dancers: A sign of talent or vulnerability? *Rheumatology* **2013**, *52*, 651–658. [CrossRef] [PubMed]
15. Smith, T.O.; Davies, L.; de Medici, A.; Hakim, A.; Haddad, F.; Macgregor, A. Prevalence and profile of musculoskeletal injuries in ballet dancers: A systematic review and meta-analysis. *Phys. Ther. Sport* **2016**, *19*, 50–56. [CrossRef]
16. Bronner, S.; Bauer, N.G. Risk factors for musculoskeletal injury in elite pre-professional modern dancers: A prospective cohort prognostic study. *Phys. Ther. Sport* **2018**, *31*, 42–51. [CrossRef] [PubMed]
17. Biernacki, J.L.; Stracciolini, A.; Fraser, J.; Miceli, L.J.; Sugimoto, D. Risk Factors for Lower-Extremity Injuries in Female Ballet Dancers. *Clin. J. Sports Med.* **2018**. [CrossRef]
18. van Seters, C.; van Rijn, R.M.; van Middelkoop, M.; Stubbe, J.H. Risk Factors for Lower-Extremity Injuries Among Contemporary Dance Students. *Clin. J. Sport Med.* **2020**, *30*, 60–66. [CrossRef]
19. van Winden, D.; van Rijn, R.M.; Savelsbergh, G.J.P.; Oudejans, R.R.D.; Stubbe, J.H. Limited Coping Skills, Young Age, and High BMI Are Risk Factors for Injuries in Contemporary Dance: A 1-Year Prospective Study. *Front. Psychol.* **2020**, *11*, 1452. [CrossRef]
20. Clarsen, B.; Myklebust, G.; Bahr, R. Development and validation of a new method for the registration of overuse injuries in sports injury epidemiology: The Oslo Sports Trauma Research Centre (OSTRC) overuse injury questionnaire. *Br. J. Sports Med.* **2013**, *47*, 495–502. [CrossRef]
21. Fuller, C.W.; Ekstrand, J.; Junge, A.; Andersen, T.E.; Bahr, R.; Dvorak, J.; Hagglund, M.; McCrory, P.; Meeuwisse, W.H. Consensus statement on injury definitions and data collection procedures in studies of football (soccer) injuries. *Scand. J. Med. Sci. Sports* **2006**, *16*, 83–92. [CrossRef] [PubMed]
22. Stubbe, J.H.; Richardson, A.; van Rijn, R.M. Prospective cohort study on injuries and health problems among circus arts students. *BMJ Open Sport Exerc. Med.* **2018**, *4*, e000327. [CrossRef] [PubMed]
23. Clarsen, B.; Ronsen, O.; Myklebust, G.; Florenes, T.W.; Bahr, R. The Oslo Sports Trauma Research Center questionnaire on health problems: A new approach to prospective monitoring of illness and injury in elite athletes. *Br. J. Sports Med.* **2014**, *48*, 754–760. [CrossRef] [PubMed]
24. Kenny, S.J.; Palacios-Derflingher, L.; Whittaker, J.L.; Emery, C.A. The Influence of Injury Definition on Injury Burden in Preprofessional Ballet and Contemporary Dancers. *J. Orthop. Sports Phys. Ther.* **2018**, *48*, 185–193. [CrossRef]
25. Stewart, D.R.; Burden, S.B. Does generalised ligamentous laxity increase seasonal incidence of injuries in male first division club rugby players? *Br. J. Sports Med.* **2004**, *38*, 457–460. [CrossRef]
26. Chan, C.; Hopper, L.; Zhang, F.; Pacey, V.; Nicholson, L.L. The prevalence of generalized and syndromic hypermobility in elite Australian dancers. *Phys. Ther. Sport* **2018**, *32*, 15–21. [CrossRef]
27. Field, A. *Discovering Statistics Using SPSS*; Sage Publications: London, UK, 2009.

28. Longworth, B.; Fary, R.; Hopper, D. Prevalence and predictors of adolescent idiopathic scoliosis in adolescent ballet dancers. *Arch. Phys. Med. Rehabil.* **2014**, *95*, 1725–1730. [CrossRef]
29. Roussel, N.A.; Nijs, J.; Mottram, S.; Van Moorsel, A.; Truijen, S.; Stassijns, G. Altered lumbopelvic movement control but not generalized joint hypermobility is associated with increased injury in dancers. A prospective study. *Man. Ther.* **2009**, *14*, 630–635. [CrossRef]
30. Peduzzi, P.; Concato, J.; Kemper, E.; Holford, T.R.; Feinstein, A.R. A simulation study of the number of events per variable in logistic regression analysis. *J. Clin. Epidemiol.* **1996**, *49*, 1373–1379. [CrossRef]
31. Welsh, T.; Ambegaonkar, J.; Mainwaring, L. *Research in the Dance Sciences*; University Press of Florida: Gainesville, FL, USA, in press.
32. Clarsen, B.; Bahr, R.; Myklebust, G.; Andersson, S.H.; Docking, S.I.; Drew, M.; Finch, C.F.; Fortington, L.V.; Harøy, J.; Khan, K.M.; et al. Improved reporting of overuse injuries and health problems in sport: An update of the Oslo Sport Trauma Research Center questionnaires. *Br. J. Sports Med.* **2020**, *54*, 390–396. [CrossRef] [PubMed]
33. Bahr, R. Why screening tests to predict injury do not work-and probably never will . . . : A critical review. *Br. J. Sports Med.* **2016**, *50*, 776–780. [CrossRef]

Journal of
Clinical Medicine

MDPI

Article

Is Heel Height Associated with Pain Exacerbations in Hip Osteoarthritis Patients?—Results from a Case-Crossover Study

Kai Fu [1,2,*], Ben R. Metcalf [3], Kim L. Bennell [3], Yuqing Zhang [4], K. Douglas Gross [5], Kathryn Mills [6], Leticia A. Deveza [1], Sarah R. Robbins [1,2] and David J. Hunter [1,2]

1 Kolling Institute of Medical Research, Institute of Bone and Joint Research, University of Sydney, Sydney, NSW 2065, Australia; leticia.alle@sydney.edu.au (L.A.D.); sarah.robbins@sydney.edu.au (S.R.R.); david.hunter@sydney.edu.au (D.J.H.)
2 Department of Rheumatology, Royal North Shore Hospital and Northern Clinical School, University of Sydney, Sydney, NSW 2065, Australia
3 Centre for Health, Exercise and Sports Medicine, Department of Physiotherapy, University of Melbourne, Melbourne, VIC 3053, Australia; b.metcalf@unimelb.edu.au (B.R.M.); k.bennell@unimelb.edu.au (K.L.B.)
4 Division of Rheumatology, Allergy, and Immunology, Massachusetts General Hospital, Harvard School of Medicine, Boston, MA 02115, USA; yzhang108@mgh.harvard.edu
5 Department of Physical Therapy, MGH Institute of Health Professions, Boston, MA 02129, USA; dgross1@mghihp.edu
6 Department of Health Professions, Macquarie University, Sydney, NSW 2109, Australia; kathryn.mills@mq.edu.au
* Correspondence: kafu7690@uni.sydney.edu.au; Tel.: +61-468476117; Fax: +61-2-9463-1077

Received: 24 April 2020; Accepted: 10 June 2020; Published: 16 June 2020

Abstract: The etiology of osteoarthritis (OA) pain exacerbations is not well understood. The purpose of this study is to evaluate the association of heel height and duration of wearing shoes with higher heels with pain exacerbations in people with hip OA. Eligible participants with symptomatic hip OA were instructed to complete online questionnaires every 10 days over a 90-day follow-up period. They were required to complete the questionnaire whenever they were experiencing hip pain exacerbation. Of 252 participants recruited, 137 (54.4%) contributed both case and control period data, and were included in the analysis. Wearing shoes with a heel height ≥ 2.5 cm during the past 24 h was associated with lower odds of pain exacerbations (OR: 0.54, 95% CI: 0.30 to 0.99). A longer duration (>6 h) of wearing shoes with heel height ≥ 2.5 cm was also associated with a lower risk of hip pain exacerbations (p for linear trend = 0.003). Wearing shoes with heel height ≥ 2.5 cm and longer duration in the past 24 h may be protective against hip pain exacerbations in people with symptomatic hip OA. Given the observational study nature, it would be prudent for this to be replicated in an independent data set.

Keywords: osteoarthritis; hip; pain; footwear; case-crossover study

1. Introduction

Osteoarthritis (OA) of the hip is the second most common site of lower limb OA after the knee, with an estimated overall prevalence in the general adult population of 11%, and somewhat higher in older adults [1,2].

Pain is the ubiquitous symptom of OA and a major driver of clinical decision-making. While patients with OA often have persistent pain [3], many experience pain fluctuations or periods of exacerbation during which pain and functional impairment worsen [4]. In the early stages of OA, pain is often intermittent, becoming more frequent and severe as the disease progresses [5]. The etiology

of OA pain exacerbations is not well understood, despite recent progress in the understanding of chronic OA pain [6].

As disease-modifying treatment is not yet available for OA, pain-modifying treatments remain essential [7]. Key treatments include education, exercise, weight loss when needed and walking aids, as indicated [7]. Appropriate footwear may be a cost-effective and safe intervention for managing the painful symptoms of hip OA [8]. Though people regularly wear high-heeled shoes, it has been suggested that continued utilization may negatively influence lower extremity musculoskeletal health [9]. But the mechanism behind this has not been fully elucidated. There are many studies that have measured the biomechanical effects of shoes with higher heels in people with knee OA, but not hip OA, showing that wearing heels of increasing height could increase the medial tibiofemoral compartment and patellofemoral joint loading [9]. Furthermore, there is a strong association between medial knee load and increased pain during walking in patients with knee OA [10]. Nonetheless, these results cannot necessarily be generalized from knee OA to hip OA, given differences in risk factors, biomechanics and response to treatment [11].

To our knowledge, no study has examined the effect of footwear on pain exacerbations in people with hip OA. To better understand this, we conducted an Internet-based, case-crossover study to evaluate the associations of wearing shoes with higher heels and the duration of time spent wearing shoes with higher heels with hip pain exacerbations in people with symptomatic hip OA. Our hypothesis was that higher-heeled shoes, especially if worn for a longer duration, would be associated with a greater risk of pain exacerbations in people with hip OA.

2. Experimental Section

2.1. Study Design

An Internet-based, case-crossover study was designed to assess the relation of wearing shoes with different heel heights, and the duration of wearing such shoes to the risk of hip pain exacerbation in Australia from May 2015 to June 2017. This design has previously been described in hip and knee OA pain exacerbation studies [12–17]. Briefly, the case-crossover design uses each participant as their own control to assess the effects of transient exposures (risk factors) on episodic events (e.g., pain exacerbation) during a certain follow-up period (e.g., 90 days). The study was approved by the ethics committees of the University of Sydney (HREC 2014/801) and University of Melbourne (HREC 1443509), and all participants provided informed consent.

2.2. Study Population

To be eligible for the study, participants were required to: be ≥40 years old; self-report hip pain on most days (5–7 days/week or 20–30 days/month); have at least one hip meeting American College of Rheumatology radiological criterion for hip OA [18]; have a Kellgren and Lawrence grade of hip OA ≥ 2 [19]; have access to the Internet; and have a good understanding of spoken and written English. Participants were excluded if they had a history of total hip replacement in the index hip, a scheduled total hip replacement or a consultation with an orthopedic surgeon for consideration of total hip replacement of the symptomatic hip(s); or a history of inflammatory arthritis, osteonecrosis or Paget's disease affecting the hip.

2.3. Data Collection

An online screening survey tool was designed to recruit study participants across Australia. We advertised the study through social media (Facebook), patient advocacy websites (Arthritis Australia, etc.), local newspapers and flyers.

When a potential study candidate registered interest in participation, their contact details were emailed to a study coordinator (in Sydney or Melbourne, depending on their state). The study coordinators then contacted participants for further assessment and enrollment.

Meanwhile, the participants were also asked to post their most recent X-rays (within one year) to the study coordinators for eligibility criteria assessment. Radiographic grading of hip OA was read by two rheumatologists (L.A.D. and D.J.H.) in our study. The inter-rater reliability was assessed using 12 radiographs providing a weighted kappa of 0.6 for Kellgren and Lawrence grades. Eligible participants were enrolled and provided access to the study's website. Participants were followed for 90 days and asked to complete online questionnaires at baseline and every succeeding 10-day interval. The intensity of hip pain was assessed using an 11-point numeric rating scale (NRS; ranging from 0—"no pain" to 10—"worst pain possible"), which showed high test-retest reliability in osteoarthritis populations [20,21]. Participants were asked to indicate how bad their hip pain was at its mildest and worst in the baseline questionnaires. Consistent with previous studies [12–16], a pain exacerbation was operationally defined as an increase of ≥2 points in the participant's pain level compared with their mildest hip pain level reported in the baseline questionnaire. We used the mildest pain level at baseline as the comparator to facilitate the identification of the maximum number of pain exacerbation events [13]. When a participant considered that they were experiencing a pain exacerbation, they were asked to complete an online questionnaire, which automatically determined whether it met the operational definition of a pain exacerbation by comparing their current reported pain level with the baseline pain level. If so, they were then guided to complete the questionnaire. To avoid biased reporting, the amount of pain increase needed to be defined as a pain exacerbation was not disclosed to participants. The system sent reminder emails during every 10-day interval to remind the participants to log onto the study website when they were experiencing any pain exacerbations. A case period was defined as a 24 h period before a pain exacerbation report and a control period was a 24 h period before every 10-day interval report. A case or control period was marked as missing if questionnaire data were not entered within 48 h when there was a pain exacerbation reported. Risk factor assessment questionnaires for control periods and the case periods were the same for all online visits.

Shoes were categorized into three different styles based on heel height (<2.5 cm or 1 inch; 2.5–5 cm or 1–2 inches; >5 cm or 2 inches). Participants were asked to report which type(s) of shoes, including the heel height, they wore during the previous 24 h, and how long they were worn. To facilitate more accurate reporting of shoe heel height, participants could choose from a list of different types of shoes on the study website. The questionnaire was derived from that used in the population-based Framingham Foot Study [22,23]. We also used the International Physical Activity Questionnaire (IPAQ) short form to evaluate physical activity during the previous seven days, as physical activity could influence the choice of shoe and OA pain [24,25]. The IPAQ short form assesses the types of intensity of physical activity and sitting time that people do as part of their daily lives, which are considered to estimate total physical activity in Metabolic Equivalent of Tasks (METs)—min/week and time spent sitting. The reliability and validity tests suggest that these measures have acceptable measurement properties [26]. Three levels of physical activity (1. low, 2. moderate, 3. high) based on this questionnaire were used to classify participants.

2.4. Statistical Methods

Baseline characteristics were summarized as mean (standard deviation, SD) for continuous variables and frequency (%) of each response for categorical variables. Characteristics were summarized for all participants enrolled in the study. However, participants who did not provide data on both case periods, and control periods were excluded from the subsequent analysis. Independent sample *t*-tests, chi-square tests and nonparametric tests (Wilcoxon rank-sum test) were performed to compare those participants with both case/control periods and those without. The analyses pooled recurrent event (pain exacerbations) data under the assumption that the within-participant correlation was accounted for by conditioning on participant-specific variables (observed or unobserved), and observed time-varying factors.

Any time overlap between the case and control periods was avoided. When a pain exacerbation occurred on the same day of a normal 10-day interval report, the previous 24 h period was marked as a case period but not a control period. As each participant could contribute multiple case and control periods, an m:n matched case-control study design was used to assess the association (Figure 1). The associations of wearing shoes with higher heel heights, and the time of high-heel wearing over a 24-hour period with hip pain exacerbation, were assessed using separate conditional logistic regression models and adjusted for physical activity level. A cut-off of 6 h of wearing a heel height ≥ 2.5 cm in the past 24 h was chosen arbitrarily to differentiate those who wore them for a relatively long time. Odds ratios (ORs) were reported with corresponding 95% confidence intervals (CIs). Post-hoc analyses were also conducted to assess the interaction between the intensity of pain (at its mildest) measured at baseline, and the risk of pain exacerbation associated with heel height and the duration of wearing shoes with higher heels. Analyses were conducted using STATA version 15 (StataCorp LLC., College Station, TX, USA).

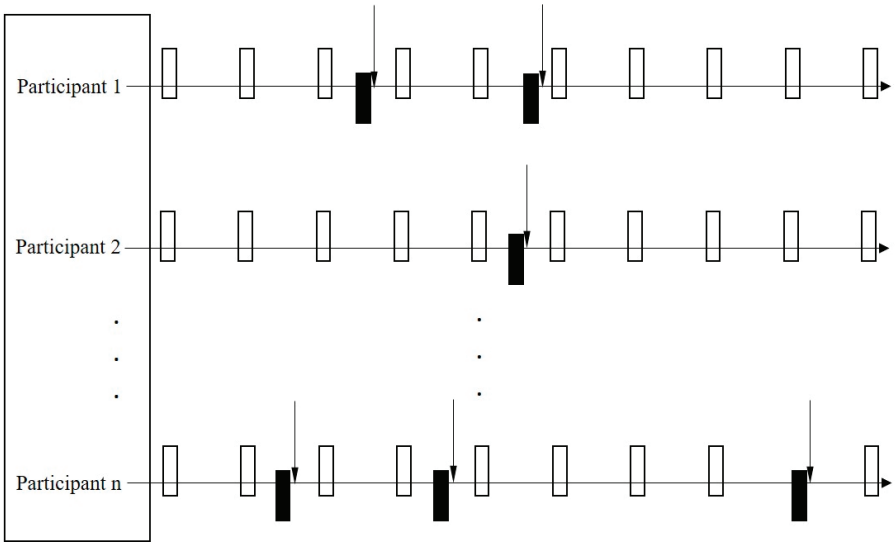

Figure 1. Case crossover study design (m:n matched). Each participant was followed up for 90 days. ▯ represents every control period; ▮ represents every case period; ↓ represents every pain exacerbation report time point.

3. Results

Among 252 participants recruited, 137 (54.4%) had at least one episode of pain exacerbation (and at least one control period), and were included in the analysis. The majority were female (114, 83.2%) and Caucasian (130, 95%). On average, these participants were 62.6 (SD: 9.8) years old with a self-reported body mass index (BMI) of 29.0 (6.3) kg/m^2. More than 60% of the participants had received a tertiary education (higher than high school education) and 78.8% performed light physical work (sedentary work or standing occupation). More than half of the participants had a Kellgren and Lawrence grade 3 hip OA. The baseline mean (SD) mildest and worst pain levels were 2.5 (2.0) out of 10 and 8.0 (1.7), respectively. Participants included in the final analysis (n = 137) showed a higher baseline pain intensity (at worst) compared with those excluded (n = 115) (t-test $p < 0.001$) (Table 1).

Table 1. Baseline characteristics of participants.

Characteristics (Mean (SD or %))	Participants Included (n = 137)	Participants Excluded (n = 115)
Age (years)	62.6 (9.8)	61.8 (8.3)
Female, n (%)	114 (83.2%)	85 (73.9%)
BMI (kg/m^2)	29.0 (6.3)	28.3 (5.8)
Index hip (right)	74 (54%)	69 (60%)
Race		
Caucasian	130 (95%)	112 (97.4%)
Others	7 (5%)	3 (2.6%)
Education level		
Less than high school	22 (16.1%)	15 (13%)
Completed high school	28 (20.4%)	30 (26.1%)
Higher than high school	87 (63.5%)	70 (60.9%)
Occupational physical workload level		
Sedentary (mostly sitting)	58 (42.3%)	53 (46.1%)
Standing occupation, physically light	50 (36.5%)	40 (34.8%)
Manual work	26 (19%)	21 (18.3%)
Heavy manual work	3 (2.2%)	1 (0.9%)
Kellgren and Lawrence grade		
2	48 (35.0%)	36 (31.3%)
3	79 (57.7%)	63 (54.8%)
4	10 (7.3%)	16 (13.9%)
Baseline pain level (0–10)		
Mildest	2.5 (2.0)	2.0 (1.7)
Worst *	8.0 (1.7)	7.1 (2.0)
IPAQ (1–3); median (IQR)	3 (2)	3 (2)

* Independent sample t-test ($p < 0.001$). IPAQ: International Physical Activity Questionnaire; IQR: Interquartile range.

There was only a small proportion of periods where participants reported wearing shoes with heels > 5 cm (10 case and 30 control periods, both less than 5%). As a result, we combined heel heights > 5 cm with 2.5–5 cm in the analysis.

In univariable conditional logistic regression analysis, wearing shoes with heel height ≥ 2.5 cm during the past 24 h was significantly associated with a decreased odds of pain exacerbations (heel height ≥ 2.5 cm vs. heel height < 2.5 cm: OR: 0.54, 95% CI: 0.30 to 0.99). There is no substantial difference after adjusting for physical activity level in the past 7 days (OR: 0.59, 95% CI: 0.32 to 1.08), and the effect estimates remained similar (Table 2).

Table 2. Association between wearing shoes with heel height ≥ 2.5 cm (yes vs. no) over the past 24 h and pain exacerbation (n = 137).

Heel Height ≥ 2.5 cm	Control Periods	Case Periods	OR (95% CI)	p Values	Adjusted OR (95% CI) *	p Values
No	798 (84%)	308 (87%)	1.0 (Reference)	0.046	1.0 (Reference)	0.090
Yes	153 (16%)	46 (13%)	0.54 (0.30, 0.99)		0.59 (0.32, 1.08)	

* Adjusted for average physical activity level in the past 7 days.

A longer duration of time wearing shoes with heel height ≥ 2.5 cm during the past 24 h was associated with a lower risk of hip pain exacerbations. The odds ratios were 1.0 (reference), 0.94 and 0.28, respectively, for wearing times of 0 h, 0–6 h and >6 h (p for linear trend = 0.003) (Table 3). This remained significant after adjusting for physical activity level in the past 7 days.

The post hoc interaction analyses did not show any effect of baseline pain level (at its mildest) and risk of pain exacerbation with heel height or the wearing time (p values > 0.05).

Table 3. Association between duration of time wearing shoes with a heel height ≥ 2.5 cm over the past 24 h and pain exacerbation (n = 137).

Duration of Wear Time	Control Periods	Case Periods	OR (95% CI)	p for Trend	Adjusted OR (95% CI) *	p for Trend
0 h	798 (84%)	308 (87%)	1.0 (Reference)		1.0 (Reference)	
0–6 h	52 (6%)	20 (6%)	0.94 (0.47, 1.88)	0.003	1.03 (0.50, 2.10)	0.006
>6 h	101 (10%)	26 (7%)	0.28 (0.12, 0.64)		0.31 (0.13, 0.71)	

* Adjusted for average physical activity level in the past 7 days.

4. Discussion

We found that wearing shoes with a heel height ≥ 2.5 cm in the past 24 h, compared with a heel height < 2.5 cm, was inversely associated with hip pain exacerbation in persons with symptomatic hip OA. This association remained even after adjusting for differences in physical activity during the previous week. There was also an independent dose—response relationship, wherein the longer the person wore shoes with a heel height ≥ 2.5 cm during the past 24 h, the lower the risk of pain exacerbation.

To our knowledge, there is no previous study investigating the relationship between shoe heel height and symptoms in people with hip OA. A recent systematic review of the effects of footwear on joint pain and function in older adults with lower limb OA found a limited number of randomized control trials evaluating wedge insoles, shock-absorbing insoles and hardness of shoe soles, but none specifically relating to heel height [27].

Our study failed to confirm our hypotheses, and surprisingly found a converse relationship, which seems somewhat counter-intuitive based on results of studies in women without OA. In this population, wearing shoes with higher heels has numerous biomechanical and neuromuscular effects on the lower limb, including the hip and pelvis, which could appear detrimental to the musculoskeletal system. In the frontal plane, a significant increase in the hip abduction moment after heel strike was found from walking in a high-heel condition (up to 9 cm) compared with barefoot walking [28,29]. In the sagittal plane, a prolonged hip flexion moment was found in high-heel conditions during the late stance phase which resulted in a 23% increase in concentric hip flexor muscle work [30]. Hip angles were larger during the stance phase in the high-heel condition [31]. Increased lumbar erector spinae muscle activity and lower pelvic range of motion were also found to be associated with wearing high heels [32]. Nonetheless, effects of higher heel shoes in women without OA may differ to those with hip OA. Furthermore, the high-heel shoes tested were much higher than those worn by the participants in our study.

Gait adaptations in people with hip OA might provide one plausible mechanism by which wearing shoes with a higher heel might be associated with a lower risk of hip pain exacerbation [31,33]. People with hip OA show gait adaptations that may be strategies to avoid pain and decrease hip joint loads, and/or result from a limitation in passive motion, particularly hip flexion contracture [33]. Possible ways to compensate for inadequate hip extension are to increase lumbar lordosis and tilt the pelvis anteriorly, strategies which can occur via wearing higher heels [34]. Thus, wearing shoes with higher heels (up to a point) might indirectly help to compensate for inadequate hip extension, decreasing pain and leading to fewer pain exacerbations. However, an important caveat to this theory is that there is significant variability between studies, with many reporting that high-heeled (>4 cm) walking showed a decrease or no change in lumbar lordosis [35,36].

We should acknowledge that another possible reason for these counter-intuitive results, is that wearing shoes with high heels means that people spend less time walking around or doing intense activities which could lead to less likelihood of having pain exacerbations by avoiding aggravating the hip joint. In our study, the inverse association between heel height and hip pain exacerbations remained comparable even after adjusting for differences in physical activity during the previous week. People with OA have higher pain levels and fatigue symptoms and decreased physical activity when compared with controls, and fatigue had a strong negative relationship to

physical activity [37]. High heel wearing could result in compensatory changes, such as increased lower limb muscle activity [28], which contributes to the higher energy cost when walking in high heels. Our previous study showed that fatigue was associated with pain exacerbations in persons with symptomatic hip OA [17], and this implies a potential indirect association between physical activity level with pain exacerbations through fatigue. What should be noted is that most people, especially women who wear high heels on a regular basis, spend the majority of time sitting rather than walking, which might contribute to a lower physical activity level [36]. In our study, we found that the effect of heel height on pain exacerbations became smaller, which can be explained by the reduced physical activity and less potential fatigue when wearing relatively higher heels. This might also explain the association between the duration of high heel wearing and pain exacerbations in our study.

Our study has limitations that should be acknowledged. Firstly, the assessment of shoe heel height was crude and relied on self-report. We did not assess the heel height relative to the ground or relative to the forefoot, and the base of the heel in terms of its width (wide-based vs. narrow-based heel), which could influence gait biomechanics. Secondly, a very small proportion of participants reported wearing shoes with heel heights > 5 cm. Thus, we cannot make conclusions about the association of wearing high-heeled shoes on pain exacerbations in hip OA, and it may be that the associations differ from those with shoe heel heights that are less extreme. Thirdly, the analyses do not take into account the shoe-wearing habits of individuals (such as whether they are a frequent wearer of shoes with higher heels), as an adaptation to footwear type might influence the relationship. There is potential confounding by indication, which might lead to distorted associations, as we did not have data before and on the day they wore higher shoes. Moreover, this relationship could be confounded by some other time-varying variables that we have not measured. Although the effect estimate remained similar after adjusting for physical activity level, we cannot confirm the role (confounder or mediator) of physical activity in this relationship based on the current analyses we have done.

Although we found these results in our study, we are not able to recommend people with hip OA to wear high heels up to 5 cm for longer periods, as this might lead to more severe constant pain and other issues related to foot, knee, low back and neck. Further research is warranted to clearly determine the advantageous heel height point before more definitive conclusions can be drawn.

5. Conclusions

Although we have limited data relating to wearing high heels > 5 cm, our results suggest that heels up to 5 cm could be worn with less probability of hip pain exacerbations in people with symptomatic hip OA. A longer time of wearing shoes with heels ≥ 2.5 cm in the past 24 h was also inversely associated with hip pain exacerbations. By applying a case-crossover design, all the confounding by constant characteristics such as age, sex, BMI, etc. can be eliminated during a 90-day follow-up by the pairing of cases to themselves. The findings of this study need to be corroborated in additional investigations before definitive recommendations about footwear for this patient group can be made.

Author Contributions: Conceptualization, D.J.H., K.F., Y.Z. and K.L.B.; clinical investigation: K.F., B.R.M., K.L.B., L.A.D. and D.J.H.; writing—original draft preparation, K.F.; writing—review and editing, B.R.M., K.L.B., Y.Z., K.D.G., K.M., L.A.D., S.R.R. and D.J.H.; funding acquisition, D.J.H. and K.L.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by a National Health and Medical Research Council program grant; grant number APP631717.

Acknowledgments: D.J.H. is supported by a National Health and Medical Research Council (NHMRC) Practitioner Fellowship. K.F. received a scholarship from the China Scholarship Council (CSC).

Conflicts of Interest: D.J.H. provides consulting advice to Pfizer, Lilly, Merck Serono and TLCbio.

References

1. Pereira, D.; Peleteiro, B.; Araujo, J.; Branco, J.; Santos, R.A.; Ramos, E. The effect of osteoarthritis definition on prevalence and incidence estimates: A systematic review. *Osteoarthr. Cartil.* **2011**, *19*, 1270–1285. [CrossRef]

2. Zhang, Y.; Jordan, J.M. Epidemiology of osteoarthritis. *Clin. Geriatr. Med.* **2010**, *26*, 355–369. [CrossRef] [PubMed]
3. Neogi, T. The epidemiology and impact of pain in osteoarthritis. *Osteoarthr. Cartil.* **2013**, *21*, 1145–1153. [CrossRef] [PubMed]
4. Trouvin, A.-P.; Marty, M.; Goupille, P.; Perrot, S. Determinants of daily pain trajectories and relationship with pain acceptability in hip and knee osteoarthritis. A national prospective cohort study on 886 patients. *Joint Bone Spine* **2019**, *86*, 245–250. [CrossRef]
5. O'Neill, T.W.; Felson, D.T. Mechanisms of osteoarthritis (oa) pain. *Curr. Osteoporos. Rep.* **2018**, *16*, 611–616. [CrossRef] [PubMed]
6. Trouvin, A.P.; Perrot, S. Pain in osteoarthritis. Implications for optimal management. *Joint Bone Spine* **2018**, *85*, 429–434. [CrossRef]
7. Hunter, D.J.; Bierma-Zeinstra, S. Osteoarthritis. *Lancet* **2019**, *393*, 1745–1759. [CrossRef]
8. Bennell, K. Physiotherapy management of hip osteoarthritis. *J. Physiother.* **2013**, *59*, 145–157. [CrossRef]
9. Sinclair, J.; Brooks, D.; Butters, B. Effects of different heel heights on lower extremity joint loading in experienced and in-experienced users: A musculoskeletal simulation analysis. *Sport Sci. Health* **2019**, *15*, 237–248. [CrossRef]
10. Birmingham, T.B.; Marriott, K.A.; Leitch, K.M.; Moyer, R.F.; Lorbergs, A.L.; Walton, D.M.; Willits, K.; Litchfield, R.B.; Getgood, A.; Fowler, P.J.; et al. Association between knee load and pain: Within-patient, between-knees, case-control study in patients with knee osteoarthritis. *Arthritis Care Res. (Hoboken)* **2019**, *71*, 647–650. [CrossRef]
11. Bierma-Zeinstra, S.M.A.; Koes, B.W. Risk factors and prognostic factors of hip and knee osteoarthritis. *Nat. Clin. Pract. Rheumatol.* **2007**, *3*, 78–85. [CrossRef]
12. Fu, K.; Makovey, J.; Metcalf, B.; Bennell, K.; Zhang, Y.; Asher, R.; Robbins, S.; Deveza, L.; Hunter, D.J. Role of hip injury and giving way in pain exacerbation in hip osteoarthritis: An internet-based case–crossover study. *Arthritis Care Res.* **2019**, *71*, 742–747. [CrossRef] [PubMed]
13. Ferreira, M.L.; Zhang, Y.; Metcalf, B.; Makovey, J.; Bennell, K.L.; March, L.; Hunter, D.J. The influence of weather on the risk of pain exacerbation in patients with knee osteoarthritis—A case-crossover study. *Osteoarthr. Cartil.* **2016**, *24*, 2042–2047. [CrossRef]
14. Zobel, I.; Erfani, T.; Bennell, K.L.; Makovey, J.; Metcalf, B.; Chen, J.S.; March, L.; Zhang, Y.; Eckstein, F.; Hunter, D.J. Relationship of buckling and knee injury to pain exacerbation in knee osteoarthritis: A web-based case-crossover study. *Interact. J. Med. Res.* **2016**, *5*, e17. [CrossRef] [PubMed]
15. Makovey, J.; Metcalf, B.; Zhang, Y.; Chen, J.S.; Bennell, K.; March, L.; Hunter, D.J. Web-based study of risk factors for pain exacerbation in osteoarthritis of the knee (spark-web): Design and rationale. *JMIR Res. Protoc.* **2015**, *4*, e80. [CrossRef]
16. Ricci, J.A.; Stewart, W.F.; Chee, E.; Leotta, C.; Foley, K.; Hochberg, M.C. Pain exacerbation as a major source of lost productive time in us workers with arthritis. *Arthritis Rheumatol.* **2005**, *53*, 673–681. [CrossRef]
17. Fu, K.; Makovey, J.; Metcalf, B.; Bennell, K.L.; Zhang, Y.; Asher, R.; Robbins, S.R.; Deveza, L.A.; Cistulli, P.A.; Hunter, D.J. Sleep quality and fatigue are associated with pain exacerbations of hip osteoarthritis: An internet-based case-crossover study. *J. Rheumatol.* **2019**, *46*, 1524–1530. [CrossRef]
18. Altman, R.; Alarcon, G.; Appelrouth, D.; Bloch, D.; Borenstein, D.; Brandt, K.; Brown, C.; Cooke, T.D.; Daniel, W.; Feldman, D.; et al. The american college of rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheumatol.* **1991**, *34*, 505–514. [CrossRef]
19. Kellgren, J.H.; Lawrence, J.S. Radiological assessment of osteo-arthrosis. *Ann. Rheum. Dis.* **1957**, *16*, 494–502. [CrossRef]
20. Downie, W.W.; Leatham, P.A.; Rhind, V.M.; Wright, V.; Branco, J.A.; Anderson, J.A. Studies with pain rating scales. *Ann. Rheum. Dis.* **1978**, *37*, 378–381. [CrossRef]
21. Alghadir, A.H.; Anwer, S.; Iqbal, A.; Iqbal, Z.A. Test-retest reliability, validity, and minimum detectable change of visual analog, numerical rating, and verbal rating scales for measurement of osteoarthritic knee pain. *J. Pain Res.* **2018**, *11*, 851–856. [CrossRef] [PubMed]
22. Hagedorn, T.J.; Dufour, A.B.; Golightly, Y.M.; Riskowski, J.L.; Hillstrom, H.J.; Casey, V.A.; Hannan, M.T. Factors affecting center of pressure in older adults: The framingham foot study. *J. Foot Ankle Res.* **2013**, *6*, 18. [CrossRef]

23. Dufour, A.B.; Broe, K.E.; Nguyen, U.S.; Gagnon, D.R.; Hillstrom, H.J.; Walker, A.H.; Kivell, E.; Hannan, M.T. Foot pain: Is current or past shoewear a factor? *Arthritis Rheumatol.* **2009**, *61*, 1352–1358. [CrossRef] [PubMed]
24. Zautra, A.J.; Fasman, R.; Parish, B.P.; Davis, M.C. Daily fatigue in women with osteoarthritis, rheumatoid arthritis, and fibromyalgia. *Pain* **2007**, *128*, 128–135. [CrossRef] [PubMed]
25. Snijders, G.F.; van den Ende, C.H.; Fransen, J.; van Riel, P.L.; Stukstette, M.J.; Defoort, K.C.; Arts-Sanders, M.A.; van den Hoogen, F.H.; den Broeder, A.A. Fatigue in knee and hip osteoarthritis: The role of pain and physical function. *Rheumatology (Oxford)* **2011**, *50*, 1894–1900. [CrossRef]
26. Craig, C.L.; Marshall, A.L.; Sjostrom, M.; Bauman, A.E.; Booth, M.L.; Ainsworth, B.E.; Pratt, M.; Ekelund, U.; Yngve, A.; Sallis, J.F.; et al. International physical activity questionnaire: 12-country reliability and validity. *Med. Sci. Sports Exerc.* **2003**, *35*, 1381–1395. [CrossRef] [PubMed]
27. Wagner, A.; Luna, S. Effect of footwear on joint pain and function in older adults with lower extremity osteoarthritis. *J. Geriatr. Phys. Ther.* **2018**, *41*, 85–101. [CrossRef]
28. Simonsen, E.B.; Svendsen, M.B.; Norreslet, A.; Baldvinsson, H.K.; Heilskov-Hansen, T.; Larsen, P.K.; Alkjaer, T.; Henriksen, M. Walking on high heels changes muscle activity and the dynamics of human walking significantly. *J. Appl. Biomech.* **2012**, *28*, 20–28. [CrossRef]
29. Barkema, D.D.; Derrick, T.R.; Martin, P.E. Heel height affects lower extremity frontal plane joint moments during walking. *Gait Posture* **2012**, *35*, 483–488. [CrossRef]
30. Esenyel, M.; Walsh, K.; Walden, J.G.; Gitter, A. Kinetics of high-heeled gait. *J. Am. Podiatr. Med. Assoc.* **2003**, *93*, 27–32. [CrossRef] [PubMed]
31. Wiedemeijer, M.M.; Otten, E. Effects of high heeled shoes on gait. A review. *Gait Posture* **2018**, *61*, 423–430. [CrossRef] [PubMed]
32. Mika, A.; Oleksy, L.; Mika, P.; Marchewka, A.; Clark, B.C. The effect of walking in high- and low-heeled shoes on erector spinae activity and pelvis kinematics during gait. *Am. J. Phys. Med. Rehabil.* **2012**, *91*, 425–434. [CrossRef]
33. Hurwitz, D.E.; Hulet, C.H.; Andriacchi, T.P.; Rosenberg, A.G.; Galante, J.O. Gait compensations in patients with osteoarthritis of the hip and their relationship to pain and passive hip motion. *J. Orthop. Res.* **1997**, *15*, 629–635. [CrossRef] [PubMed]
34. Hulet, C.; Hurwitz, D.E.; Andriacchi, T.P.; Galante, J.O.; Vielpeau, C. Functional gait adaptations in patients with painful hip. *Rev. Chir. Orthop. Reparatrice Appar. Mot.* **2000**, *86*, 581–589.
35. Baaklini, E.; Angst, M.; Schellenberg, F.; Hitz, M.; Schmid, S.; Tal, A.; Taylor, W.R.; Lorenzetti, S. High-heeled walking decreases lumbar lordosis. *Gait Posture* **2017**, *55*, 12–14. [CrossRef] [PubMed]
36. Cronin, N.J. The effects of high heeled shoes on female gait: A review. *J. Electromyogr. Kinesiol.* **2014**, *24*, 258–263. [CrossRef]
37. Murphy, S.L.; Smith, D.M.; Clauw, D.J.; Alexander, N.B. The impact of momentary pain and fatigue on physical activity in women with osteoarthritis. *Arthritis Care Res.* **2008**, *59*, 849–856. [CrossRef]

Journal of
Clinical Medicine

MDPI

Article

Self-Assessment of Competence and Referral Behavior for Musculoskeletal Injections among Dutch General Practitioners

Emely Spruit [1], Marianne F. Mol [1], P. Koen Bos [2], Sita M.A. Bierma-Zeinstra [1,2], Patrick Krastman [1] and Jos Runhaar [1,*]

[1] Department of General Practice, Erasmus MC, University Medical Center, 3000 CA Rotterdam, The Netherlands; e.spruit@erasmusmc.nl (E.S.); marianne.f.mol@gmail.com (M.F.M.); s.bierma-zeinstra@erasmusmc.nl (S.M.A.B.-Z.); p.krastman@erasmusmc.nl (P.K.)

[2] Department of Orthopaedic Surgery, Erasmus MC, University Medical Center, 3000 CA Rotterdam, The Netherlands; p.k.bos@erasmusmc.nl

* Correspondence: j.runhaar@erasmusmc.nl

Received: 19 May 2020; Accepted: 15 June 2020; Published: 16 June 2020

Abstract: General practitioners (GPs) are qualified and trained to administer therapeutic musculoskeletal injections when indicated. However, it is unknown to what extend Dutch GPs feel competent to administer these injections in clinical practice. Reluctance among GPs to inject might lead to unnecessary and costly referral to secondary care. An online and offline questionnaire was spread among Dutch GPs, querying demographics, GPs' self-assessment of injection competence, the number of administered/referred injections and management strategy for musculoskeletal injections. A total of 355 GPs responded. In total, 81% of the GPs considered themselves competent in administering musculoskeletal injections. Self-assessed incompetent GPs performed less injections the last month than self-assessed competent GPs (1.2 ± 1.4 vs 4.8 ± 4.6 injections, $P < 0.001$). Additionally, they referred four times more often to a colleague GP (0.4 ± 1.0 vs 0.1 ± 0.6 injections per month, $P < 0.001$) and twice as often to secondary care (1.0 ± 1.3 vs 0.5 ± 0.9 injections per month, $P = 0.001$). Self-assessed incompetence was associated with female sex (OR [95% CI] = 4.94 [2.39, 10.21]) and part-time work (OR [95% CI] = 2.58 [1.43, 4.66]). The most frequently addressed barriers were a lack of confidence in injection skills, lack of practical training, and uncertainty about the effectiveness and diagnosis of musculoskeletal injections. Although most GPs considered themselves competent to administer musculoskeletal injections, the referral rate to secondary care for several injections was strikingly high. To decrease secondary care referrals, addressing some of the most frequently indicated barriers is highly recommended.

Keywords: musculoskeletal disorders; therapeutic injections; competence

1. Introduction

Musculoskeletal problems are common in general practice. In the Netherlands, yearly 700 consultations per thousand registered patients concern problems of the musculoskeletal system [1]. The costs of productivity loss and the burden of disease due to musculoskeletal disorders are high [2–4]. Moreover, the prevalence of musculoskeletal diseases is increasing [5]. Therefore, effective diagnostic assessment and treatment of these disorders is of paramount importance. Conservative therapies such as physiotherapy and analgesics are the first treatments of choice [6]. Unfortunately, painkillers are not always effective and non-steroidal anti-inflammatory drugs and opioids have numerous side effects and contraindications [7].

Therapeutic injections are an option when non-drug therapies and painkillers fail or are not recommended. In primary care, musculoskeletal injections normally consist of corticosteroids with or

without the addition of local analgesics. The indications for musculoskeletal injections in primary care are diverse. Intra-articular injections are, for example, applied for patients diagnosed with osteoarthritis or as an addition in frozen shoulder treatment [6]. Furthermore, soft tissue injections are used, among other conditions, for carpal tunnel syndrome, de Quervain's tendinitis, subacromial bursitis and trigger finger [6]. The long-term effectiveness of musculoskeletal injections is often questioned. [8] Nevertheless, there is evidence that musculoskeletal injections do have short-term effects and are therefore advocated in multiple primary care guidelines [6,9–11]. These injections have few adverse effects and, if administered adequately, rarely lead to complications. [9] However, the correct use of musculoskeletal injections for the treatment of common disorders in primary care requires competence and self-confidence by the doctor that administrates these injections.

In Northern Ireland and England, research has been done on the administration of musculoskeletal injections by general practitioners (GPs) [12,13]. These studies concluded that the majority of GPs performed most musculoskeletal injections themselves, rather than referring to a colleague or to secondary care [12,13]. Among Northern Irish and English GPs, female GPs, urban GPs and part-time working GPs were less likely to perform musculoskeletal injections [12,13]. Reported barriers for administering injections were little confidence in injection skills and difficulty in maintaining skills, leading to over-referral to secondary care [12–14].

There are few data on the number of musculoskeletal injections by Dutch GPs. A Dutch study on septic arthritis following intra-articular injections demonstrated higher administration of injections by medical specialists (524 injections by 12 specialists) compared to GPs (170 injections by 23 GPs) during a follow-up period [15]. This study, together with the indicated Northern Irish and English studies, give the impression that there is a variation in self-confidence among GPs to perform musculoskeletal injections. We hypothesized that self-assessed incompetent GPs perform less injections and refer more often to secondary care. Moreover, we hypothesize that the female sex, working in an urban practice, working part-time, not being a GP trainer or specialized in the musculoskeletal system would be associated to self-assessed incompetence. Through questionnaires among Dutch GPs, we aimed to answer the following research questions:

1. To what extent do Dutch GPs feel competent in administering musculoskeletal injections?
2. How does self-assessed (in)competence affect their clinical treatment and referral behavior?
3. Which factors are associated with self-assessed incompetence?
4. Which barriers and facilitators to administer musculoskeletal injections do Dutch GPs experience?

2. Methods

2.1. Development of the Questionnaire

A cross-sectional survey through a self-administered online or paper questionnaire, inspired by the survey of Liddell et al. [13], was developed for Dutch GPs (see Appendix A). The questionnaire included questions on demographic data (sex, organization, work setting, full-time equivalent (FTE), patient population, number of years since the completion of GP training, being a GP trainer and specialization in musculoskeletal disease), the number of musculoskeletal injections performed during the past month and the management strategy (injection by GP self, referral to other GP, or referral to secondary care) for a set of 18 selected musculoskeletal diseases for which an injection is indicated or optional in primary care (inspired by Liddell et al. [13]).

To assess competence, the GPs were asked to what extent they agreed with the statement 'I consider myself competent in performing musculoskeletal injections'. The outcome was measured using a five-point Likert scale with answer options 'completely disagree', 'disagree', 'neutral', 'agree' and 'completely agree'. Additionally, the questionnaire contained questions about the experienced barriers for musculoskeletal injections and possible ways to overcome these barriers (facilitators). A pilot among three GPs and two GPs in training was performed to make sure the questions were unequivocal.

2.2. Study Population and Recruitment

To draw valid conclusions, we estimated a total of 200 completed questionnaires would be necessary, based on previous research [12,13]. Data were collected during the period from 12 November 2018 to 14 January 2019. Through HAweb (an online platform for Dutch GPs), a link to the questionnaire was spread among all 12378 members. Reminders were posted 1, 2.5 and 4 weeks after the first post. Additionally, a paper version of the questionnaire was sent by mail to all 636 GPs in the Haaglanden and Amsterdam–Amstelland regions (randomly selected regions in the Netherlands). Furthermore, the link to the questionnaire was spread using Twitter and Facebook. Finally, forty-two GP trainers who are affiliated with our department were asked to fill out the questionnaire. The total size of the targeted population was 12,378, as all targeted GPs are registered on HAweb.

All GPs participated voluntarily in this study. Ethical approval to conduct the study was not necessary.

2.3. Statistical Analysis

To examine whether the sample of responding GPs was representative of the total Dutch GP population, the demographic data were compared to the data of the total Dutch GP population (by eyeballing), obtained through the Netherlands Institute for Health Services Research (NIVEL) [16].

The normality of the questionnaire data was checked using the Kolmogorov–Smirnov test. Using descriptive statistics, the frequency of competent/incompetent GPs was calculated. To compare the referral behavior of competent and incompetent GPs, Mann–Whitney U and chi-squared tests were used to test for the significance of differences between groups on the average number of (referred) injections and the percentage of referrals per indication.

Demographic factors were compared between competent and incompetent GPs, first using chi-squared tests and subsequently with a multivariable logistic regression analysis containing all significant factors. To examine the magnitude of the associations between the competence groups, the number of injections administered and those referred, a linear regression was used, adjusted for significant factors from the multivariable model. Lastly, descriptive statistics were used to demonstrate the frequencies of barriers and facilitators to inject. All data were analyzed using SPSS version 24. The significance level in the analysis was set at $P < 0.05$.

3. Results

3.1. Responses

A total of 355 returned questionnaires were analyzed. Fifty-four questionnaires were incompletely filled in. All completed questions from the incomplete questionnaires were used in the analyses.

3.2. Representativeness

Sex and work setting were representative of Dutch GPs (Table 1). The other demographic characteristics differed slightly; less full-timers and more recently graduated GPs were included in our study.

Table 1. Frequencies of demographics of the study and Netherlands Institute for Health Services Research (NIVEL).

Demographic		Current Study % [a]	NIVEL %
Sex	Man	42.0	49 [b]
	Woman	52.1	51 [b]
	Neutral	0.3	-
Organization	Solo practice	19.2	17.8 [b]
	Duo practice	23.7	40.3 [b]
	Group practice	20.8	41.9 [b]
	Health center	15.5	-
	Other [c]	11.5	-
Work setting	Rural	26.8	30.2 [d]
	Urban	67.6	69.8 [d]
FTE	0–0.20	1.7	0.2
	0.21–0.40	5.9	2.8
	0.41–0.60	20.6	15.9
	0.61–0.80	30.7	28.5
	0.81–1	35.2	52.6
Patient population [e]	< standard practice	13.2	-
	= standard practice	25.9	-
	> standard practice	55.2	-
Completion GP training	<5 years	23.1	13.6 [f]
	5–15 years	33.2	29.4 [f]
	16–25 years	18.9	18.5 [fg]
	>25 years	19.4	38.6 [fg]
Trainer	Yes	24.5	-
	No	70.1	-
GP musculoskeletal system	Yes	1.7	-
	No	93.0	-

[a] Percentages do not add up to 100% due to missing data. [b] Data calculated on regular established general practitioners. [c] Other organizations include observers, nursing home, non-practicing etc. [d] Rural is defined as little urban and not urban. Urban is defined as moderately urban, strong urban and very strong urban. [e] Standard practice = 2095 patients. [f] Data calculated on the basis of Table 2 of the NIVEL brochure: the number of general practitioners who have graduated in the Netherlands for each graduation year or graduation period for their status as at 1 January 2016. [g] Data concern the period 16–26 years instead of 16–25 years and the period 27–42 years instead of > 25 years in connection with specified graduation periods instead of graduation year.

Table 2. Percentages of general practitioners (GPs) who referred an injection to secondary care per indication.

Indication	All GPs % (N)	Competent % (N)	Incompetent % (N)	*P*-Value
Indications for which an injection is recommended by the Dutch College of General Practitioners				
Carpal tunnel syndrome	36.1 (330)	32.2 (267)	52.4 (63)	0.003
Knee osteoarthritis	26.7 (333)	23.0 (270)	42.9 (63)	0.001
Plantar fasciitis	22.7 (326)	20.9 (263)	30.2 (63)	0.116
De Quervain's tenosynovitis	22.5 (329)	16.1 (267)	50.0 (62)	0.000
Supraspinatus tendinitis	20.0 (320)	16.9 (260)	33.3 (60)	0.004
Trigger finger/thumb	17.1 (333)	12.6 (270)	36.5 (63)	0.000
Osteoarthritis shoulder	16.9 (325)	14.5 (262)	27.0 (63)	0.018
Trochanteric bursitis	5.8 (330)	4.1 (268)	12.9 (62)	0.007
Bursitis shoulder	2.1 (333)	0.7 (270)	7.9 (63)	0.000
Indications for which an injection is not recommended by the Dutch College of General Practitioners				
Ankle osteoarthritis	50.0 (324)	50.2 (261)	49.2 (63)	0.888
Osteoarthritis MCP/PIP/DIP	48.9 (321)	44.2 (258)	68.3 (63)	0.001
Osteoarthritis CMC	48.9 (319)	45.1 (257)	64.5 (62)	0.006
Sacroiliitis	44.4 (420)	43.4 (258)	48.4 (62)	0.479
Achilles tendinitis	33.1 (323)	31.2 (260)	41.3 (63)	0.126
Prepatellar bursitis	11.9 (327)	9.1 (264)	23.8 (63)	0.001
Medial epicondylitis	8.5 (329)	7.5 (266)	12.7 (63)	0.185
Lateral epicondylitis	7.0 (328)	6.4 (265)	9.5 (63)	0.385
Olecranon bursitis	5.8 (327)	4.9 (265)	9.7 (62)	0.148

3.3. Competence

The distribution of the answers to the competence question was 2.0% completely disagree, 5.4% disagree, 11.8% neutral, 57.7% agree and 23.1% completely agree. After completion of GP training (which includes injection training), all GPs are considered capable of injecting. Therefore, 'neutral' was considered as incompetent and categories were defined by combining 'completely disagree', 'disagree' and 'neutral' as 'incompetent' versus 'agree' and 'completely agree' as 'competent', resulting in 80.8% of the GPs considering themselves competent.

3.4. Number of Injections

3.4.1. Injections Aministered

GPs (N = 339) performed an average of 4.1 (SD = 4.4) musculoskeletal injections during the past month. Forty-four GPs did not administer any injections in this period. The number of administered injections differed significantly between competent and incompetent GPs (4.8 ± 4.6 vs. 1.2 ± 1.4, $P < 0.001$).

3.4.2. Injections Referred to GP Colleague

On average, 0.2 (SD = 0.7) injections were referred to a GP colleague in the past month. A significant difference was observed between the competence groups; 0.1 (SD = 0.6) injections were referred to a GP colleague in the past month by the competent GPs and 0.4 (SD = 1.0) by the incompetent GPs ($P < 0.001$).

3.4.3. Injections Referred to Scondary Care

The mean number of injections that were referred to the secondary care in the past month was 0.6 (SD = 1.0). GPs who considered themselves competent referred 0.5 (SD = 0.9) injections to secondary care in the past month, compared to 1.0 (SD = 1.3) referred injections by incompetent GPs ($P = 0.001$).

3.5. Injection Indications

Table 2 shows the percentages of GPs who would refer to secondary care for an injection for different indications. The most referred injection indications were ankle osteoarthritis and metacarpophalangeal (MCP)/proximal interphalangeal (PIP)/distal interphalangeal (DIP)/carpometacarpal (CMC) osteoarthritis. GPs referred the least often for an injection of shoulder bursitis. For eleven of the eighteen injection indications, a significantly higher percentage referred to secondary care among the incompetent GPs compared to the competent GPs.

3.6. Factors Associated with Incompetence and Referral Behavior

Table 3 shows the differences in demographic between the competence groups. In the multivariable analysis, factors associated with incompetence were female sex (OR [95% CI] = 4.94 [2.39, 10.21]) and part-time work (low FTE) (OR [95% CI] = 2.58 [1.43, 4.66]). Given the association with sex, stratified regression analyses were done. In the sex-stratified analysis, FTE differed significantly between male competent and incompetent GPs and years since completion of GP training differed significantly among female GPs. Male and female GPs who considered themselves competent administered a significantly higher number of injections than their incompetent counterparts (B [95% CI] = 4.76 [1.21, 8.30] for men and B [95% CI] = 2.01 [1.31, 2.70] for women). The differences in referrals to a colleague GP or to secondary care between male competent and incompetent GPs were not significant (respectively, B [95% CI] = −0.37 [−0.89, 0.14], B [95% CI] = −0.08 [−0.68, 0.52]), while female competent GPs referred significantly less than female incompetent GPs (respectively, B [95% CI] = −0.31 [−0.54, −0.09], B [95% CI] = −0.55 [−0.89, −0.21]).

Table 3. Percentages of self-assessed (in)competent GPs in performing injections per demographic group.

Demographic		N	Competent %	Incompetent %	*P*-Value
Sex	Man	149	93.3	6.7	-
	Woman	185	70.8	29.2	0.000
Organization	Solo practice	70	81.4	18.6	-
	Duo practice	85	78.8	21.2	-
	Group practice	74	79.7	20.3	-
	Health center	55	83.6	16.4	-
	Other	41	80.5	19.5	0.967
Work setting	Rural	95	78.9	21.1	-
	Urban	240	81.2	18.8	0.631
Full-time equivalent (FTE)	0–0.20	6	66.7	33.3	-
	0.21–0.40	21	61.9	38.1	-
	0.41–0.60	73	68.5	31.5	-
	0.61–0.80	109	84.4	15.6	-
	0.81–1	125	80.2	19.2	0.001
Patient population	< standard practice	47	76.6	23.4	-
	= standard practice	92	83.7	16.3	-
	> Standard practice	196	80.1	19.9	0.584
Completion GP training	< 5 years	82	81.7	18.3	-
	5–15 years	118	78.8	21.2	-
	16–25 years	67	73.1	26.9	-
	> 25 years	69	89.9	10.1	0.090
Trainer	Yes	87	81.6	18.4	-
	No	249	80.3	19.7	0.793

3.7. Barriers

The most common barriers to perform musculoskeletal injections experienced by GPs were a lack of practical training, a lack of confidence in skills, a lack of confidence in diagnosis and uncertainty about

the effectiveness of the injection (Figure 1). Few GPs had concerns about medicolegal issues or had a bad experience with injections due to complications. Overall, 28% of the general practitioners indicated that they did not experience any barriers to perform musculoskeletal injections. When analyzed for men and women separately, 45% of the male and 16% of the female GPs did not experience any barriers. Moreover, strikingly more female than male GPs reported a lack of confidence in their skills (50% vs. 14%).

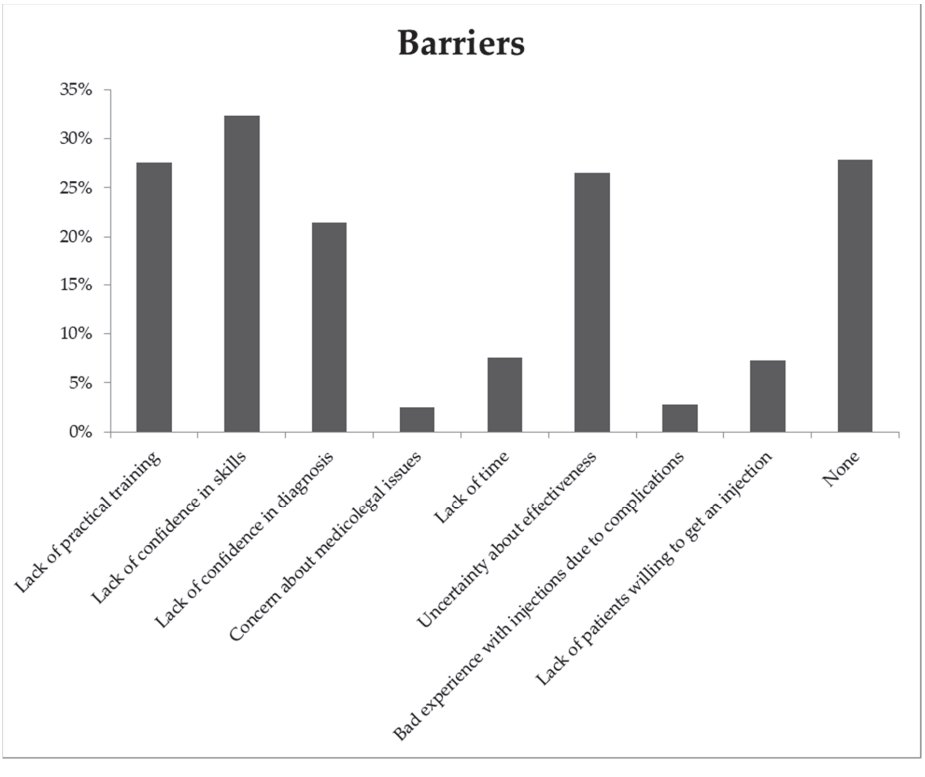

Figure 1. Percentage of GPs reporting selected barriers to performing musculoskeletal injections (N = 355).

3.8. Facilitators

In total, 41% of the GPs considered training in musculoskeletal injection by a rheumatologist or an orthopedic surgeon a possible facilitator (Figure 2). In addition, training by an experienced GP colleague and the possibility to perform the injection intramuscularly instead of intra-articular were often indicated. Overall, facilitators slightly differed between the sexes. Only the option to administer an injection intramuscularly instead of intra-articular showed a clear difference between male and female GPs (respectively, 25% and 49%).

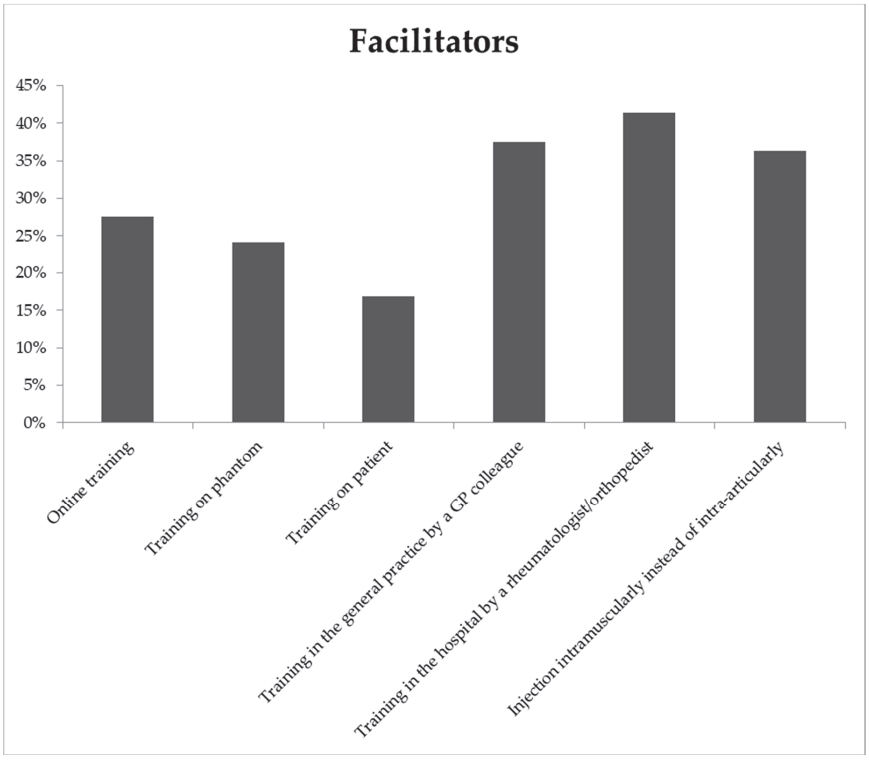

Figure 2. Percentage of GPs reporting selected facilitators to performing musculoskeletal injections (N = 355).

4. Discussion

This study demonstrated that one in five GPs considered themselves incompetent in performing musculoskeletal injections. Self-assessed incompetent GPs referred twice as many injections to secondary care than self-assessed competent GPs. The main barriers for GPs to perform joint and soft tissue injections were a lack of practical training and a lack of confidence in their own skills.

A greater percentage of female GPs considered themselves incompetent. In general, women tend to underestimate their skills more often, while men tend to overestimate themselves [17,18]. Looking at the actual performance between the sexes, multiple studies confirmed that women and men are equally skilled, or women even outperform men [19–23]. In the Netherlands, the proportion of female GPs has increased from 36% in 2007 to 51% in 2016 [16]. Therefore, it is important that women consider themselves competent in administering musculoskeletal injections to decrease referrals to secondary care. According to Sharp et al., GPs' perception of practical skills competence can be increased by performing more procedures [24]. For this reason, injection training for insecure female GPs could be recommended.

Part-time work was also associated with self-assessed incompetence, independent of female sex. Since the hours worked per Dutch GP is decreasing, more GP are working part-time [25]. To keep up with all clinical skills can be hard for a GP, especially when the skills are not used regularly [26]. Part-time workers will encounter musculoskeletal injections less frequently. To prevent unnecessary and costly referrals to secondary care, keeping up with injection skills should be a priority for GPs. Otherwise, referral to colleague GPs (e.g., those specialized in musculoskeletal disorders) should be facilitated.

4.1. Comparison with Other Studies

4.1.1. Number of Injections

Liddell et al. concluded that GPs carried out a median number of 17.0 musculoskeletal injections in the last year [13]. The number of injections administered by Dutch GPs was higher, with a mean number of 4.1 injections per month. It is possible that English guidelines recommend a musculoskeletal injection less often compared to Dutch GP guidelines or that English GPs have more/stronger personal barriers to injecting. Unfortunately, Liddell et al. and Gormley et al. did not examine referrals to secondary care [12,13].

In accordance with previous studies, we found that female GPs performed significantly fewer injections than male GPs [12,13]. Since these studies did not question self-assessed competence, it is not possible to compare the competence in musculoskeletal injections between English and Dutch GPs [12,13].

4.1.2. Barriers

In agreement with previous studies, we found that GPs' most frequently reported barriers to carry out musculoskeletal injections were a lack of confidence in skills and a lack of practical training [12,13]. In contrast, uncertainty about the effectiveness of injections was a common barrier among Dutch GPs [12,13]. This is noteworthy, as the Dutch College of General Practitioners (NHG) provides guidelines with evidence-based recommendations regarding musculoskeletal injections. Despite this, GPs indicated to question these NHG guidelines, though they are usually well adhered to these guidelines [6,27].

4.1.3. Training

English GPs preferred training on patients, but this facilitator was the least popular among Dutch GPs [13]. Previously, training on patients was deemed superior to training on a phantom for improving confidence in performing musculoskeletal injections [12]. However, training on phantoms is more feasible, because the training could be given in large groups and GPs can easily practice the injection multiple times consecutively.

4.2. Limitations

First of all, there is a possibility of non-response bias in our study. In particular, more GPs who have a special interest in injections or in the musculoskeletal system could have completed the questionnaire. In the Netherlands, there is a dedicated group of general practitioners who specialize in musculoskeletal disease. Colleague GPs can refer patients to them or consult them with questions about musculoskeletal pathology. As expected, the prevalence of these specialized GPs in our study was higher than among the entire Dutch GP population. This might have led to an overestimation of self-assessed competence in the current study. This overestimation could be even larger, as self-assessed incompetent GPs might feel reluctant to fill out the questionnaire (truthfully), despite the fact that the survey was anonymous.

Our study included slightly less full-timers when compared to Dutch primary care. As mentioned earlier, part-time work was associated with self-assessed incompetence. Therefore, the percentage of competent GPs could be underestimated.

Moreover, it is unclear whether GPs are capable of adequate self-assessment of their injection skills. In a study among GPs, Janssen et al. found that there was only a moderate correlation between self-assessed competence of technical skills and actual competence (as measured with a performance-based test) [28].

Furthermore, the lack of information collected on working conditions is a limitation as well. The level of colleague support or the psychosocial work environment may also influence the decision to refer injections to a colleague.

Finally, we decided to dichotomize GPs' competence by combining the answer categories 'completely disagree', 'disagree' and 'neutral' versus 'agree' and 'completely agree'. Obviously, when 'neutral' was added to the definition of competence, more GPs consider themselves competent (92.7%), but this does not change the significance of any of the outcomes.

4.3. Implications for Practice and Further Research

Although the self-assessed competence among GPs was high, many patients were referred to secondary care for musculoskeletal injections by incompetent GPs. To decrease referrals to secondary care, more training on injection skills is recommended, both in GP training and as refresher courses. NHG-led courses already exist for injections for shoulder, de Quervain's tenosynovitis, carpal tunnel syndrome, trigger finger/thumb, knee and trochanter pain syndrome [6]. Referral to specialized GPs instead of to secondary care seems a good option for incompetent GPs. Another possibility is that orthopedists visit primary care to assess and treat patients together with the general practitioner.

In recent years, the effectiveness of intramuscular injections for musculoskeletal disease, compared to traditional intra-articular injections, has been studied. If an intramuscular injection is not inferior to an intra-articular injection, it may be easier for GPs to carry out musculoskeletal injections in the future, as injecting intramuscularly was considered a facilitator by GPs to perform musculoskeletal injections. Furthermore, the clinical effectiveness of intramuscular injections has already been shown for hip osteoarthritis [29] and the effectiveness of an intra-articular injection versus intramuscular injection in patients with rotator cuff disease showed no significant differences between the methods of administration [30]. Recently, a randomized controlled trial on the non-inferiority of an intramuscular injection in comparison with an intra-articular injection in patients with knee osteoarthritis finished data collection [31].

Finally, uncertainty about the effectiveness of musculoskeletal injections was a common barrier among GPs in the current study. A more in-depth analysis on the reasons why GPs question the effectiveness of musculoskeletal injections would be of interest.

Author Contributions: Conceptualization, M.F.M., S.M.A.B.-Z. and J.R.; methodology, E.S., M.F.M., S.M.A.B.-Z. and J.R.; formal analysis, E.S., M.F.M., and J.R.; investigation, E.S. and M.F.M.; data curation, E.S., M.F.M., and J.R.; writing—original draft preparation, E.S.; writing—review and editing, M.F.M., P.K.B., S.M.A.B.-Z., P.K., J.R.; supervision, M.F.M. and J.R.; project administration, J.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

Appendix A

Questionnaire

1. To what extent do you agree with the following statement? (1 answer possible) I consider myself competent in administering musculoskeletal injections.

    - ☐ completely disagree
    - ☐ disagree
    - ☐ neutral
    - ☐ agree
    - ☐ completely agree

2. Which barriers do you experience in administering musculoskeletal injections? (multiple answers possible)

    - ☐ Lack of training
    - ☐ Lack of confidence in skills

- ☐ Lack of confidence in diagnosis
- ☐ Medical-legal issues
- ☐ Lack of time
- ☐ Uncertainty about effectiveness of an injection
- ☐ Bad experience with injections due to complications
- ☐ Lack of patients willing to get an injection
- ☐ None

3. What would facilitate administering musculoskeletal injections for you? (multiple answers possible)

- ☐ Online training
- ☐ Training on a phantom
- ☐ Training on a patient
- ☐ Training in primary care by a GP colleague
- ☐ Training in secondary care by a rheumatologists/orthopaedist
- ☐ Possibility to perform an intramuscular injection instead of an intra articular injection

4. How many musculoskeletal injections did you administer in the past month? (1 answer possible)

- ☐ None
- ☐ Indicate how many (enter only numbers)

5. How many musculoskeletal injections did you refer to a GP colleague in the past month? (1 answer possible)

- ☐ None
- ☐ Indicate how many (enter only numbers)

6. How many musculoskeletal injections did you refer to secondary care (rheumatologist/orthopaedist) in the past month?

- ☐ None
- ☐ Indicate how many (enter only numbers)

7. What is your sex? (1 answer possible)

- ☐ Man
- ☐ Woman
- ☐ Neutral

8. What is the organisation of your practice? (1 answer possible)

- ☐ Healthcare centre
- ☐ Duo practice
- ☐ Solo practice
- ☐ Group practice
- ☐ Other, namely:

9. What is the setting of your practice? (1 answer possible)

- ☐ Rural
- ☐ Urban

10. How many FTE (fulltime-equivalent) do you work as a GP? (1 answer possible) (fulltime-equivalent)

- □ 0–0.20
- □ 0.21–0.40
- □ 0.41–0.60
- □ 0.61–0.80
- □ 0.81–1

11. What is the size of your patient population (standard practice = 2095)? (1 answer possible)

- □ Smaller than the standard practice
- □ Equal to the standard practice
- □ Bigger than the standard practice

12. How many years ago did you graduate from the GP training? (1 answer possible)

- □ < 5 years
- □ 5–15 years
- □ 16–25 years
- □ >25 years

13. Are you a GP trainer? (1 answer possible)

- □ Yes
- □ No

14. Are you a GP specialised in the musculoskeletal system? (1 answer possible)

- □ Yes
- □ No

15. Finally, for the following indications, please indicate whether you would administer the injection yourself, would refer for the injection to a colleague GP or to secondary care (rheumatologist/orthopaedist), or would not advise an injection/indication did not occur in your practice (multiple answers possible per indication).

Indication	Self	Colleague GP	Secondary Care	No Injection/Did Not occur
Bursitis shoulder	□	□	□	□
Osteoarthritis shoulder	□	□	□	□
Supraspinatus tendinitis	□	□	□	□
Lateral epicondylitis	□	□	□	□
Medial epicondylitis	□	□	□	□
Olecranon bursitis	□	□	□	□
Carpal tunnel syndrome	□	□	□	□
Trigger finger/thumb	□	□	□	□
De Quervain's tenosynovitis	□	□	□	□
Osteoarthritis MCP/PIP/DIP	□	□	□	□
Osteoarthritis CMC	□	□	□	□
Trochanteric bursitis	□	□	□	□
Sacro-iliitis	□	□	□	□
Knee osteoarthritis	□	□	□	□
Prepatellar bursitis	□	□	□	□
Achilles tendinitis	□	□	□	□
Ankle osteoarthritis	□	□	□	□
Plantar fasciitis	□	□	□	□

References

1. Meeste Consulten bij Huisarts voor Klachten aan het Bewegingsapparaat. *NIVEL Zorgregistraties*, infographic. Available online: https://www.nivel.nl/nl/nieuws/nederlanders-vooral-met-beweegklachten-naar-de-huisarts (accessed on 1 September 2019).
2. Hermans, J.; Koopmanschap, M.A.; Bierma-Zeinstra, S.M.; van Linge, J.H.; Verhaar, J.A.; Reijman, M.; Burdorf, A. Productivity costs and medical costs among working patients with knee osteoarthritis. *Arthritis Care Res. (Hoboken)* **2012**, *64*, 853–861. [CrossRef] [PubMed]
3. March, L.; Smith, E.U.; Hoy, D.G.; Cross, M.J.; Sanchez-Riera, L.; Blyth, F.; Buchbinder, R.; Vos, T.; Woolf, A.D. Burden of disability due to musculoskeletal (MSK) disorders. *Best Pract. Res. Clin. Rheumatol.* **2014**, *28*, 353–366. [CrossRef] [PubMed]
4. Lubeck, D.P. The costs of musculoskeletal disease: Health needs assessment and health economics. *Best Pract. Res. Clin. Rheumatol.* **2003**, *17*, 529–539. [CrossRef]
5. GBD 2016. Analysis of the Global Burden of Disease Study 2016. *Lancet* **2017**, *390*, 1211–1259.
6. Nederlandse Huisartsen Genootschap (NHG) Richtlijnen. Available online: https://www.nhg.org (accessed on 1 September 2019).
7. Curatolo, M.; Bogduk, N. Pharmacologic pain treatment of musculoskeletal disorders: Current perspectives and future prospects. *Clin. J. Pain* **2001**, *17*, 25–32. [CrossRef] [PubMed]
8. Foster, Z.J.; Voss, T.T.; Hatch, J.; Frimodig, A. Corticosteroid Injections for Common Musculoskeletal Conditions. *Am. Fam. Physician* **2015**, *92*, 694–699.
9. Babatunde, O.O.; Jordan, J.L.; Van der Windt, D.A.; Hill, J.C.; Foster, N.E.; Protheroe, J. Effective treatment options for musculoskeletal pain in primary care: A systematic overview of current evidence. *PLoS ONE* **2017**, *12*, e0178621. [CrossRef]
10. Bannuru, R.R.; Osani, M.C.; Vaysbrot, E.E.; Arden, N.K.; Bennell, K.; Bierma-Zeinstra, S.M.A.; Kraus, V.B.; Lohmander, L.S.; Abbott, J.H.; Bhandari, M.; et al. OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. *Osteoarthr. Cartil.* **2019**, *27*, 1578–1589. [CrossRef]
11. National Clinical Guideline Centre (UK). *Osteoarthritis: Care and Management in Adults*; National Institute for Health and Care Excellence (UK): London, UK, 2014.
12. Gormley, G.J.; Corrigan, M.; Steele, W.K.; Stevenson, M.; Taggart, A.J. Joint and soft tissue injections in the community: Questionnaire survey of general practitioners' experiences and attitudes. *Ann. Rheum. Dis.* **2003**, *62*, 61–64. [CrossRef]
13. Liddell, W.G.; Carmichael, C.R.; McHugh, N.J. Joint and soft tissue injections: A survey of general practitioners. *Rheumatology (Oxford)* **2005**, *44*, 1043–1046. [CrossRef]
14. Brown, J.S.; Smith, R.R.; Cantor, T.; Chesover, D.; Yearsley, R. General practitioners as providers of minor surgery–a success story? *Br. J. Gen. Pract.* **1997**, *47*, 205–210. [PubMed]
15. Bruens, M.L.; van der Zaag-Loonen, H.J.; Steenstra, F.; Stemerding, A.M.; Wijngaarden, S. Septic arthritis after intra-articular injection is rare: Does the Taskforce Infection Prevention use a sledgehammer to crack a nut? *Ned. Tijdschr. Geneeskd.* **2016**, *160*, D789. [PubMed]
16. Van der Velden, L.; Kasteleijn, A.; Kenens, R. Cijfers uit de registratie van huisartsen, peiling 2016. *Proc. Natl. Acad. Sci. USA* **2017**.
17. Bridges, J.S. Sex Differences in Occupational Performance Expectations. *Psychol. Women Q.* **1988**, *12*, 75–90. [CrossRef]
18. Beyer, S. Gender differences in the accuracy of self-evaluations of performance. *J. Personal. Soc. Psychol.* **1990**, *59*, 960–970. [CrossRef]
19. Lind, D.S.; Rekkas, S.; Bui, V.; Lam, T.; Beierle, E.; Copeland, E.M. Competency-based student self-assessment on a surgery rotation. *J. Surg. Res.* **2002**, *105*, 31–34. [CrossRef]
20. Blanch, D.C.; Hall, J.A.; Roter, D.L.; Frankel, R.M. Medical student gender and issues of confidence. *Patient Educ. Couns.* **2008**, *72*, 374–381. [CrossRef]
21. Minter, R.M.; Gruppen, L.D.; Napolitano, K.S.; Gauger, P.G. Gender differences in the self-assessment of surgical residents. *Am. J. Surg.* **2005**, *189*, 647–650. [CrossRef]
22. Kilminster, S.; Downes, J.; Gough, B.; Murdoch-Eaton, D.; Roberts, T. Women in medicine—Is there a problem? A literature review of the changing gender composition, structures and occupational cultures in medicine. *Med. Educ.* **2007**, *41*, 39–49. [CrossRef]

23. Whittle, S.R.; Eaton, D.G. Attitudes towards transferable skills in medical undergraduates. *Med. Educ.* **2001**, *35*, 148–153. [CrossRef]
24. Sharp, L.K.; Wang, R.; Lipsky, M.S. Perception of competency to perform procedures and future practice intent: A national survey of family practice residents. *Acad. Med.* **2003**, *78*, 926–932. [CrossRef] [PubMed]
25. Van der Velden, L.; Batenburg, R.S. Aantal huisartsen en aantal FTE van huisartsen vanaf 2007 tot en met 2016: Werken er nu meer of minder huisartsen dan 10 jaar geleden en werken zij nu meer of minder FTE? *Proc. Natl. Acad. Sci. USA* **2017**.
26. Arthur, W., Jr.; Bennett, W., Jr.; Stanush, P.L.; McNelly, T.L. Factors That Influence Skill Decay and Retention: A Quantitative Review and Analysis. *Hum. Perform.* **1998**, *11*, 57–101. [CrossRef]
27. Lugtenberg, M.; Burgers, J.S.; Besters, C.F.; Han, D.; Westert, G.P. Perceived barriers to guideline adherence: A survey among general practitioners. *BMC Fam. Pract.* **2011**, *12*, 98–2296. [CrossRef]
28. Jansen, J.J.; Tan, L.H.; van der Vleuten, C.P.; van Luijk, S.J.; Rethans, J.J.; Grol, R.P. Assessment of competence in technical clinical skills of general practitioners. *Med. Educ.* **1995**, *29*, 247–253. [CrossRef]
29. Dorleijn, D.M.J.; Luijsterburg, P.A.J.; Reijman, M.; Kloppenburg, M.; Verhaar, J.A.N.; Bindels, P.J.E.; Bos, P.K.; Bierma-Zeinstra, S.M.A. Intramuscular glucocorticoid injection versus placebo injection in hip osteoarthritis: A 12-week blinded randomised controlled trial. *Ann. Rheum. Dis.* **2018**, *77*, 875–882. [CrossRef]
30. Ekeberg, O.M.; Bautz-Holter, E.; Tveita, E.K.; Juel, N.G.; Kvalheim, S.; Brox, J.I. Subacromial ultrasound guided or systemic steroid injection for rotator cuff disease: Randomised double blind study. *BMJ* **2009**, *338*, a3112. [CrossRef]
31. Mol, M.F.; Runhaar, J.; Bos, P.K.; Dorleijn, D.M.J.; Vis, M.; Gussekloo, J.; Bindels, P.J.E.; Bierma-Zeinstra, S.M.A. Effectiveness of intramuscular gluteal glucocorticoid injection versus intra-articular glucocorticoid injection in knee osteoarthritis: Design of a multicenter randomized, 24 weeks comparative parallel-group trial. *BMC Musculoskelet. Disord.* **2020**, *21*, 225. [CrossRef]

Journal of
Clinical Medicine

MDPI

Article

Long-Term Surgical Results of Skip Pedicle Screw Fixation for Patients with Adolescent Idiopathic Scoliosis: A Minimum-Ten-Year Follow-Up Study

Masashi Uehara [1,*], Shugo Kuraishi [1], Shota Ikegami [1], Hiroki Oba [1], Takashi Takizawa [1], Ryo Munakata [1], Terue Hatakenaka [1], Michihiko Koseki [2] and Jun Takahashi [1]

[1] Department of Orthopaedic Surgery, Shinshu University School of Medicine, 3-1-1 Asahi, Matsumoto, Nagano 390-8621, Japan; kura@apricot.ocn.ne.jp (S.K.); sh.ikegami@gmail.com (S.I.); oba1@hotmail.co.jp (H.O.); takizawahal449@gmail.com (T.T.); battlemiddle3@yahoo.co.jp (R.M.); hatakenaka@shinshu-u.ac.jp (T.H.); jtaka@shinshu-u.ac.jp (J.T.)

[2] Department of Mechanical Engineering and Robotics, Faculty of Textile Science and Technology, Shinshu University, 3-15-1 Tokida, Ueda, Nagano 390-8621, Japan; koseki@shinshu-u.ac.jp

* Correspondence: masashi_u560613@yahoo.co.jp; Tel.: +81-263-37-2659

Received: 22 October 2020; Accepted: 10 December 2020; Published: 10 December 2020

Abstract: Skip pedicle screw fixation for adolescent idiopathic scoliosis (AIS) requires fewer screws and can reduce the risk of neurovascular injury as compared with segmental pedicle screw fixation. However, the long-term impact of screw number reduction on correction and clinical results is unclear. This study examined the 10-year post-operative outcomes of skip pedicle screw fixation for patients with AIS. We reviewed the outcomes of 30 patients who underwent skip pedicle screw fixation for AIS. Radiological and clinical findings were assessed before and immediately, 2 years, and 10 years after surgery in the remaining 25 patients. The mean Cobb angle of the main curve preoperatively and immediately, 2 years, and 10 years post-operatively was 59.4°, 23.4°, 25.8°, and 25.60°, respectively, and was significantly improved at all post-surgical time points (all $p < 0.001$). The mean correction rate immediately after surgery was 60.8%, and the correction loss rate at the observation end point was 4.8%. The Cobb angle of the lumbar curve was significantly improved immediately after surgery, and the correction persisted until 10 years post-operatively. Remarkable gains were observed for most Scoliosis Research Society-22 patient questionnaire sub-scores at the final follow-up versus preoperative assessments. In conclusion, good correction of the AIS deformity by skip pedicle screw fixation was well maintained over a long follow-up period of 10 years, with clinically meaningful gains in Society-22 patient questionnaire sub-scores.

Keywords: long-term results; adolescent idiopathic scoliosis; skip pedicle fixation; 10 years; posterior fusion; surgery

1. Introduction

Segmental pedicle screw fixation using numerous thoracic pedicle screws as described by Suk et al. [1] is a useful option for posterior spinal fusion in adolescent idiopathic scoliosis (AIS). The posterior approach offers stronger fixation strength to enable shorter fusion length and better correction [2]. On the other hand, pedicle screw fixation has also been associated with a risk of injury to neurovascular and visceral structures, such as the spinal cord, nerve root, lung, and aorta [1,3–5]. To avoid those serious complications, Takahashi et al., employ skip pedicle screw fixation for the AIS deformity since it requires fewer screws (Figure 1) and can reduce surgical costs [6–9]. Uehara et al., earlier described the two-year results of computer-assisted skip pedicle screw fixation in 62 consecutive cases of AIS containing all Lenke classification types apart from 5C [9]. However, evidence on the

long-term outcomes of skip pedicle screw fixation, especially scoliosis correction loss, implant failure including screw breakage/pull-out, and decreased Scoliosis Research Society-22 patient questionnaire (SRS-22r) scores, is scarce.

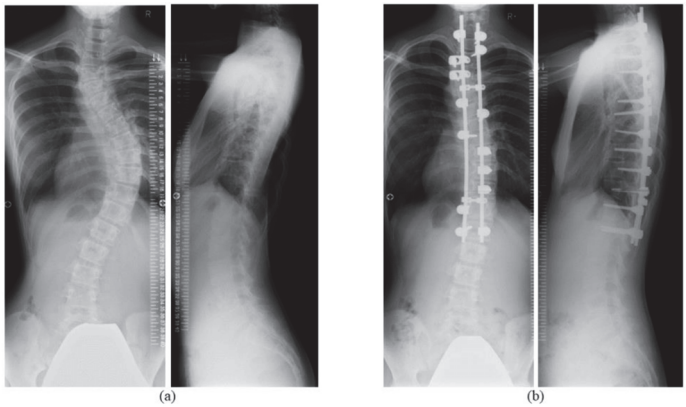

Figure 1. Representative skip pedicle screw fixation for adolescent idiopathic scoliosis. (**a**) Preoperative Cobb angle of the main thoracic curve was 53°. (**b**) Skip pedicle screw fixation from T1 to L1 improved the scoliotic curve to 17°.

The purpose of this study is to examine if the corrective position retention and good clinical outcomes of skip pedicle screw fixation are maintained even at 10 years after surgery. To further validate the usefulness of skip pedicle screw fixation for AIS, we examined the radiological and clinical findings of patients after a long observation period of at least 10 years.

2. Methods

2.1. Patients

This study was conducted at our institution. We earlier described the 2-year results of skip pedicle screw fixation in 62 consecutive cases of AIS [9]. Among 30 consecutive patients receiving preoperative computed tomography (CT)-based navigation-assisted skip pedicle screw fixation for AIS at our hospital between August 2005 and March 2010, 25 patients (2 male and 23 female; mean ± standard deviation (SD) age: 14.4 ± 2.1 years, range: 12–19 years) who were followed for a minimum of 10 years (follow-up rate: 83.3%) were retrospectively reviewed. All surgeries were performed by one senior spine surgeon specializing in scoliosis (J.T.). The inclusion criteria were patients who underwent AIS surgery and follow-up at our hospital. The exclusion criteria were follow-up of less than 10 years and Lenke type 5 curves. Five patients without follow-up of at least 10 years due to relocation were excluded. Classification according to Lenke grading was type 1A in seven patients, 1B, 2A, and 6C in three patients each, 1C in four patients, 4B in one patient, 3C and 4C in two patients each. Patients with Lenke type 5 curves were excluded from this investigation since they often received segmental pedicle screw fixation. Prior approval of the study was obtained from the investigational review board of our hospital (No. 3500). Written informed consent for publication was obtained from all patients prior to this study. No financial support or equivalent was received for this investigation. The complete database of the cohort can be accessed at the Zenodo repository (doi.org/10.5281/zenodo.4295859).

2.2. Surgical Technique

Pedicle screw insertion into the upper and lower ends of the instrumentation area was performed bilaterally. Screw placement in the vertebral levels other than the upper instrumented vertebra, lowest instrumented vertebra, and apex was judged based on the size and rigidity of the curve and

was skipped when possible. Screw insertion was skipped near the apical vertebrae of each curve and in the junctional zone between structural and compensatory curves that were not close to fused end vertebrae. We excluded pedicles with an outer diameter of less than the thinnest screw diameter during navigation planning. Pedicle screws were routinely inserted near the concave apical vertebrae. For cases in which a 4.5 mm-diameter pedicle screw appeared difficult to insert in navigation planning, tape was used instead. However, in cases where only one vertebra was skipped, no tape was used. Since pedicle diameter is wider near the convex vertebrae, we basically used pedicle screws and skipped one vertebra per inserted vertebra. Screw holes were made using a Kotani probe [10] with CT-based navigation, while pedicle screws were inserted without navigation. Bone grafting employed local bone. Ponte osteotomy was not performed in this cohort. The lower instrumented vertebra (LIV) was decided as the vertebra which last touched the central sacral vertical line when the lumbar modifier was Lenke 1A or 2A. For Lenke 1B and 1C curves, the LIV was determined as the stable vertebra or one vertebra distal to the stable vertebra. For Lenke 3C, 4B, 4C, and 6C curves, we routinely set the LIV as L3.

2.3. Evaluation and Statistical Analysis

The first author (M.U.), who was not associated with any surgery in the cohort, performed the evaluation. Cobb angle [11] and thoracic kyphosis from T5 to T12, absolute value of clavicular angle (CA), distance from the C7 plumb line to the central sacral vertical line (C7PL), LIV tilt, and other radiological parameters were evaluated before and immediately, 2 years, and 10 years after surgery. The main curve flexibility was calculated as: (Cobb angle in standing position—Cobb angle in side bending)/Cobb angle in standing position × 100%. CT was performed at 6 months postoperatively to confirm bony fusion of the facet and to determine whether the patients could be permitted to exercise. Clinical outcomes were examined using SRS-22r scores [12,13] before and at 2 and 10 years post-operatively. Results for the Japanese Orthopaedic Association Back Pain Evaluation Questionnaire (JOABPEQ), which included a visual analog scale (VAS) for back pain and complications requiring surgical revision, were assessed at the study end point. The JOABPEQ contains five domains (pain-related disorder, lumbar spine dysfunction, gait disturbance, social life disturbance, and psychological disorder) whose scores range from 0 (lowest) to 100 (highest) [14]. We also evaluated whether JOABPEQ scores for back pain were influenced by poor overall correction, such as coronal decompensation, lumbar decompensation, the adding-on phenomenon, and early degeneration. Coronal decompensation was considered as C7PL deviation exceeding 20 mm [15]. Lumbar decompensation was defined as progression of thoracolumbar/lumbar Cobb angle by 10° or more versus immediately postoperative Cobb angle [16]. The adding-on phenomenon was judged as follows: (1) a ≥5° aggravation of scoliotic Cobb angle and a shift in curve-end vertebrae towards the caudal side, or (2) a ≥5° wedge-shaped change in the vertebral disc adjacent to the lower side of the fused vertebrae during postoperative follow-up [17]. Pre- and post-surgical data were compared using the paired t-test with Bonferroni correction, and data from the two groups were compared by means of Welch's t-test with EZR software (Saitama Medical Center, Jichi Medical University, Saitama, Japan), a graphical user interface for R (The Foundation for Statistical Computing, Vienna, Austria). As a modified version of R commander, EZR adds statistical functions frequently used in biostatistics. A p-value of <0.05 was considered statistically significant.

3. Results

The cohort's preoperative radiological and clinical features are summarized in Table 1. The preoperative Cobb angle of the main thoracic curve was 59.4°, and the main thoracic curve flexibility was 38.2%. The mean ± SD number of fused vertebrae was 10.8 ± 1.9 (range: 8–13), mean surgical time was 294 ± 80 min (range: 168–420 min), mean blood loss volume was 1304 ± 672 mL (range: 500–3050 mL), and mean screw density (i.e., number of screws per vertebra fused) was 1.33 ± 0.22 (range: 1.00–1.77). The vertebral level of the LIV was T12 in five cases, L1 in eight cases,

L2 in three cases, and L3 in nine cases. Among the 250 inserted pedicle screws, the grade 2 and 3 perforation (i.e., >2 mm) rate was 6.8%, and the grade 3 perforation (i.e., >4 mm) rate was 4.4% based on post-operative CT. No severe neurovascular injuries from screw misplacement were observed.

Table 1. Preoperative radiological and clinical features.

	Mean ± Standard Deviation (Range)
Age, years	14.4 ± 2.1 (12–19)
Sex, male/female	2:23
Cobb angle of main thoracic curve, °	59.5 ± 15.1 (44–94)
Main curve flexibility, %	38.2 ± 19.4 (6.3–96.1)
Cobb angle of lumbar curve, °	40.5 ± 19.9 (12–82)
Thoracic kyphotic angle (T5–T12), °	9.1 ± 8.3 (−8–29)
Clavicular angle, °	2.2 ± 1.9 (0–7.9)
C7PL, cm	1.4 ± 1.1 (0–5.0)

C7PL: distance from the C7 plumb line to the central sacral vertical line.

The Cobb angle of the main thoracic curve was significantly improved immediately, at 2 years, and at 10 years after correction vs. preoperative measurements (all $p < 0.001$) (Figure 2, Table 2). The mean correction rate of the main curve at the above time points was 60.8 ± 2.2%, 56.3 ± 2.5%, and 56.8 ± 2.2%, respectively (Table 2). The correction loss rate between immediately and 10 years after surgery was 4.8 ± 3.6%.

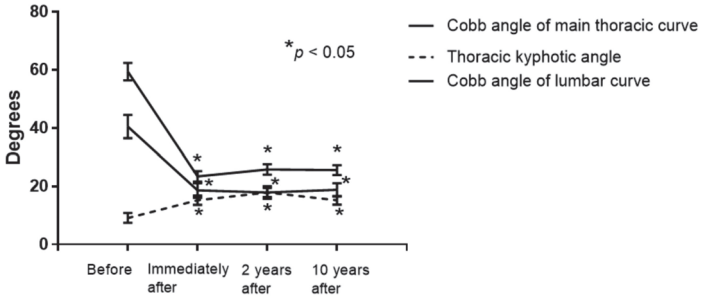

Figure 2. The mean Cobb angle of the main thoracic curve before and at 2 and 10 years after surgery was 59.4°, 25.8°, and 25.6°, respectively. The main thoracic curve Cobb angle was significantly improved at both post-operative time points ($p < 0.001$). The respective mean T5–12 thoracic kyphotic angle before and at 2 and 10 years after surgery was 9.1°, 17.9°, and 15.2°. The Cobb angle of the lumbar curve was significantly improved immediately after surgery, which persisted until 10 years post-operatively.

Table 2. Surgical results of skip pedicle screw fixation.

Radiological and Physical Evaluation	Before Surgery	Immediately after Surgery	Two Years after Surgery	Ten Years after Surgery
Cobb angle of main thoracic curve, °	59.4 ± 3.0	23.4 ± 1.8 **	25.8 ± 1.7 **	25.6 ± 1.7 **
Correction rate of Cobb angle (main thoracic curve), %	N/A	60.8 ± 2.2	56.3 ± 2.5	56.8 ± 2.2
Thoracic kyphotic angle (T5–T12), °	9.1 ± 1.7	15.2 ± 1.6 **	17.9 ± 1.5 **	15.2 ± 1.5 **
Clavicular angle, °	2.2 ± 0.4	3.2 ± 0.5	2.4 ± 0.4	2.2 ± 0.4
C7PL, cm	1.4 ± 0.2	1.3 ± 0.2	0.7 ± 0.2 *	1.4 ± 0.4
Cobb angle of lumbar curve, °	40.5 ± 4.0	18.6 ± 2.5 **	17.9 ± 2.2 **	18.8 ± 2.3 **
LIV tilt, °	20.1 ± 1.5	7.8 ± 1.4 **	7.3 ± 1.4 **	7.7 ± 1.2 **

All values are expressed as the mean ± standard error; LIV: lower instrumented vertebra, N/A: not applicable; * $p < 0.05$ vs. before surgery, ** $p < 0.001$ vs. before surgery.

The T5–12 thoracic kyphotic angle was significantly improved at all post-operative time points over baseline (all $p < 0.05$) (Figure 2, Table 2). Absolute CA values were similar, corresponding to 2.2 ± 0.4° preoperatively, 3.2 ± 0.5° immediately after surgery, 2.4 ± 0.4° at 2 years after surgery, and 2.2 ± 0.4° at the final follow-up. Similarly, absolute C7PL values remained comparable, with values of 1.4 ± 0.2 cm, 1.3 ± 0.2 cm, 0.7 ± 0.2 cm, and 1.4 ± 0.4 cm, respectively (Table 2).

The Cobb angle of the lumbar curve and LIV tilt were significantly improved immediately after surgery, which persisted until 10 years post-operatively (Figure 2, Table 2).

SRS-22r scores increased significantly in all domains from preoperatively to 2 and 10 years after surgery (all $p < 0.05$) (Table 3). At the study end point, the mean satisfaction score was 4.02 ± 0.17 (range: 2–5.0), and the mean total score was 4.29 ± 0.09 (range: 3.1–4.95) (Table 3).

Table 3. Clinical results of SRS-22r scores.

SRS-22r Domain	Before Surgery	Two Years after Surgery	Ten Years after Surgery
Function	4.31 ± 0.12	4.74 ± 0.05 **	4.65 ± 0.09 *
Pain	4.01 ± 0.12	4.68 ± 0.07 **	4.40 ± 0.13 *
Self-image	2.75 ± 0.11	3.88 ± 0.13 ***	3.89 ± 0.15 ***
Mental health	3.73 ± 0.18	4.62 ± 0.07 ***	4.31 ± 0.12 *
Sub-total	3.69 ± 0.10	4.48 ± 0.05 ***	4.31 ± 0.09 ***
Satisfaction	N/A	4.08 ± 0.15	4.02 ± 0.17
Total	N/A	4.44 ± 0.05	4.29 ± 0.09

All values are expressed as the mean ± standard error; SRS-22r: Scoliosis Research Society-22 patient questionnaire, N/A: not applicable; * $p < 0.05$ vs. before surgery, ** $p < 0.01$ vs. before surgery, *** $p < 0.001$ vs. before surgery.

JOABPEQ scores at 10 years after surgery were 86.8 ± 4.2 for pain-related disorder, 94.0 ± 2.4 for lumbar spine dysfunction, 95.8 ± 2.7 for gait disturbance, 84.4 ± 3.6 for social life disturbance, and 72.9 ± 3.4 for psychological disorder (Table 4). Three (12.0%) patients experienced coronal decompensation, five (20.0%) patients had lumbar decompensation, two (8.0%) patients exhibited adding-on, and three (12.0%) patients had early degeneration. The JOABPEQ score for low back pain in patients with coronal decompensation was 100 ± 0, which was significantly higher than in patients without it (85.0 ± 21.8) ($p = 0.004$). The JOABPEQ scores for low back pain in patients with and without lumbar decompensation were comparable, with values of 91.4 ± 19.2 and 85.7 ± 21.8, respectively ($p = 0.58$). The JOABPEQ score for low back pain in patients with adding-on was significantly higher than in patients without it (100 ± 0.0 vs. 85.7 ± 21.6, $p = 0.004$). The respective JOABPEQ scores for low back pain in patients with and without early degeneration were similar, corresponding to 62.0 ± 32.9 and 90.2 ± 17.4 ($p = 0.27$). The mean VAS score for low back pain was 19.6 ± 5.4 (range: 0−80) and was zero in 13 (52%) patients. Bone union was confirmed by CT in all cases. Screw breakage without symptoms requiring surgical revision was seen in two patients. No severe neurovascular complications from surgical invasion or indication for surgical revision were recorded.

Table 4. Comparison of JOABPEQ scores with reference values for low back pain patients.

JOABPEQ Domain	Present Series	Reference Value [18]
Pain-related disorder	86.8 ± 4.2	42.9
Lumbar spine dysfunction	94.0 ± 2.4	58.3
Gait disturbance	95.8 ± 2.7	50.0
Social life disturbance	84.4 ± 3.6	51.4
Psychological disorder	72.9 ± 3.4	47.6

All values are expressed as the mean ± standard error; JOABPEQ: Japanese Orthopaedic Association Back Pain Evaluation Questionnaire.

4. Discussion

Skip pedicle screw fixation for AIS correction significantly improved both radiological and clinical parameters at a minimum of 10 years after surgery using fewer inserted screws than in segmental fixation and without complications requiring surgical revision.

The primary goals of surgical AIS treatment are preventing curve progression and correcting the deformity. Segmental pedicle screw fixation is frequently used for rod anchoring in posterior fixation of the scoliotic deformity in AIS. The rate of neurovascular complications from misplaced screws ranges from 0% to 1.3%, being attributed mainly to narrow pedicle diameter and spinal rotation [2,18–20]. We have been employing skip pedicle screw fixation for AIS to reduce the number of screw violations and the associated risk of major adverse events [6]. Behrbalk et al., reported that a low screw density technique was as safe and effective as a high screw density technique for posterior-only correction of Scheuermann kyphosis and could provide significant cost savings [21]. Cheung et al., reported that fulcrum bending correction could estimate curve correction in AIS surgery using alternate-level thoracic pedicle screws [22]. In our study, the preoperative Cobb angle of the main curve was 59.4°, and the main curve flexibility was 38.2%. We defined the planned correction angle as the difference between the Cobb angle in side bending and the target Cobb angle, and the number of screws was determined with reference to a previously reported formula, i.e., planned correction angle/1.7° [23]. In the present study, mean screw density was 1.33, grade 3 screw perforation rate was 4.4%, and grade 2 and 3 combined perforation rate was 6.8%. No neurovascular injuries were encountered during 10 years of observation. Although preoperative CT-based navigation is useful for safely inserting pedicle screws, performing a CT scan of all vertebrae for the length of the fusion to check pedicle anatomy and size carries the potential risk of long-term damage from radiation.

Although skip pedicle screw fixation is considered safer than segmental screw fixation, a trade-off with diminished long-term correction and clinical effects may exist. A two-year observational radiological study of surgical AIS correction showed that an all-pedicle screw system provided better maintenance of corrective parameters than did hybrid instrumentation surgery using pedicle screws, hooks, and sublaminar wire or tape as an anchor, with final correction rates of 70.41% and 60.00%, respectively [24]. Preoperative flexibility was approximately 50% in both groups [24]. Hwang et al., also reported satisfactory results for skip pedicle screw insertion, obtaining 69% for correction rate and 2% for correction loss rate at 5 years [25]. In this study, the correction rate was 59.8% immediately post-operatively, and the 10-year correction loss rate was 4.8%, which were less favorable than Hwang's findings but comparable to those of the hybrid group. Moreover, the correction loss at 10 years was low at 2.2°. The absence of Ponte osteotomy procedures in this series may have reduced the correction rate. Screw breakage was seen in two patients, who experienced a mean correction loss of 10.5° at 10 years but maintained good SRS-22r scores. Regarding the changes occurring from 2 years to 10 years post-operatively, the end vertebrae were not fused in some cases, so it was conceivable that alterations were caused by angular deformations in the cranial or caudal adjacent intervertebrae of the unfused end vertebrae.

Ishikawa et al., reported that the final Cobb angle of the thoracolumbar/lumbar curve was significantly correlated with the immediately post-operative LIV tilt [26]. On the other hand, Skaggs et al., described that LIV tilt was not associated with post-surgical lumbar Cobb angle [27]. In our study, the Cobb angle of the lumbar curve and the LIV tilt were significantly improved until 10 years after surgery. However, 12.0% of the patients had coronal decompensation, 20.0% of the patients had lumbar decompensation, 8.0% of the patients had adding-on, and 12.0% of the patients had early degeneration. Furthermore, comparisons of the LIV tilt immediately after surgery in patients with and without lumbar decompensation revealed similar findings (6.4 ± 5.3° and 8.2 ± 7.3°, respectively; $p = 0.63$).

The JOABPEQ domain scores in our cohort were all considerably higher than the reference values of low-back-pain patients [28] (Table 4). Furthermore, all scores were greater than the minimum clinically important differences (MCIDs) for JOABPEQ (20.4 for low back pain, 15.6 for lumbar

function, 16.8 for walking ability, 13.4 for social life function, and 9.4 for mental health) reported by Ogura et al. [29]. The score for psychological disorder was the lowest, possibly due to the stresses of child care or working life in our young adult cohort. Hence, the long-term correction ability of skip pedicle screw fixation was considered sufficient, especially since the preoperative flexibility in our cohort (approximately 40%) was comparably lower than that of earlier reports. Furthermore, evaluations on whether the JOABPEQ scores for back pain were influenced by poor overall correction, such as coronal decompensation, lumbar decompensation, the adding-on phenomenon, and early degeneration, revealed that coronal decompensation and adding-on were not relevant adverse effects at 10 years after surgery.

Regarding long-term radiological results, the main thoracic curve Cobb angle improved from 51° to 16°, with a correction loss rate of 5% in a five-year study on segmental pedicle screw fixation for thoracic scoliosis [30]. However, data on the long-term benefits of skip pedicle screw fixation are scarce. We observed that this method significantly improved and maintained the coronal and sagittal radiological parameters of AIS cases at 10 years after surgery without severe neurovascular events or the need for surgical revision. Although a formal cost-effectiveness analysis was not performed, fewer implants were presumed as more economical for the patients.

With regard to the long-term clinical results of surgical AIS treatment, there are several reports on the Harrington method, posterior fixation with the rod and hook system, and anterior fusion, and few on pedicle screw fixation. In a comparative control study using the SRS-24, AIS patients undergoing correction surgery by the Harrington method had findings equivalent to those of healthy subjects for all parameters, and the magnitude of the scoliotic angle was not associated with quality of life (QOL) scores [31]. In a similar study with healthy subjects using the SRS-22, however, Akazawa et al., witnessed that patients had comparable scores for pain and mental health, but inferior results for function and self-image [32]. Using the Short Form 36 Health Survey, Götze et al., observed that adolescents who received correction surgery by the Harrington method had similar physical function, but lower mental health, in relation to German national standard values. Curve type and magnitude of the scoliotic angle were unrelated to QOL [33]. The present study evaluated 10-year post-operative clinical results using SRS-22r scores to reveal that self-image and sub-total were significantly ameliorated at 2 years and 10 years after surgery. Although some patients were subjected to long fusion procedures, the function and pain scores were also improved after correction. Thus, skip pedicle screw fixation for AIS produces good long-term correction without function loss or pain worsening.

We lastly examined whether the SRS-22r score changes in our cohort were clinically meaningful. In earlier studies on MCID for clinical domain outcomes, MCID was 0.20 for pain (area under the receiver operating characteristic curve (AUC) = 0.723), 0.08 for function (AUC = 0.648), and 0.98 for self-image (AUC = 0.629) [34]. The changes in the scores for pain, function, and self-image were all higher than their respective MCID, confirming that skip pedicle fixation produced clinically meaningful improvements.

Newton et al., noted that posterior segmental instrumentation for AIS could significantly decrease thoracic kyphosis [35]. Such a finding was absent in our study, with the mean thoracic kyphotic angle before and at 2 and 10 years after surgery being 9.1°, 15.2°, and 17.9°, respectively. Skip pedicle screw fixation may also have the advantage of improving thoracic kyphosis. If more rigid cobalt chromium rods had been used, larger thoracic kyphosis might have been possible.

The main limitations of this study were no control group, a limited sample size, and a retrospective design. No blinding of the assessor was also a limitation of this study. Although five patients were lost to follow-up, the resulting 83.3% follow-up rate was considered satisfactory based on Solberg et al., reporting that a follow-up loss of 22% did not bias conclusions about the effects of treatment [36]. Thus, our analysis on the outcomes of 25 AIS patients indicated good long-term radiological and clinical results for skip pedicle screw fixation, with almost all parameters being well conserved up to 10 years post-operatively.

5. Conclusions

Skip pedicle screw fixation for AIS provided significant and sustained radiological and clinical improvements at 10 years after surgery without serious complications.

Author Contributions: Conceptualization, M.U. and J.T.; methodology, M.U., S.K., S.I., H.O., T.T., R.M., T.H., M.K., and J.T.; validation, M.U., and S.K., S.I., H.O.; formal analysis, M.U. and R.M.; investigation, M.U., T.T., R.M., T.H., M.K., and J.T.; data curation, M.U., S.K., S.I., H.O., T.T., R.M., T.H., M.K., and J.T.; writing-original draft preparation, M.U.; writing-review and editing, S.K., S.I., H.O., T.T., R.M., T.H., M.K., and J.T.; and supervision, J.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

1. Suk, S.I.; Lee, C.K.; Kim, W.J.; Chung, Y.J.; Park, Y.B. Segmental pedicle screw fixation in the treatment of thoracic idiopathic scoliosis. *Spine* **1995**, *20*, 1399–1405. [CrossRef] [PubMed]
2. Halm, H.; Niemeyer, T.; Link, T.; Liljenqvist, U. Segmental pedicle screw instrumentation in idiopathic thoracolumbar and lumbar scoliosis. *Eur. Spine J.* **2000**, *9*, 191–197. [CrossRef] [PubMed]
3. Liljenqvist, U.R.; Halm, H.F.; Link, T.M. Pedicle screw instrumentation of the thoracic spine in idiopathic scoliosis. *Spine* **1997**, *22*, 2239–2245. [CrossRef] [PubMed]
4. Heini, P.; Scholl, E.; Wyler, D.; Eggli, S. Fatal cardiac tamponade associated with posterior spinal instrumentation. *Spine* **1998**, *23*, 2226–2230. [CrossRef] [PubMed]
5. Papin, P.; Aebi, M. Unusual presentation of spinal cord compression related to misplaced pedicle screws in thoracic scoliosis. *Eur. Spine J.* **1999**, *8*, 156–159. [CrossRef] [PubMed]
6. Takahashi, J.; Hirabayashi, H.; Hashidate, H.; Ogihara, N.; Kato, H. Accuracy of multilevel registration in image-guided pedicle screw insertion for adolescent idiopathic scoliosis. *Spine* **2010**, *35*, 347–352. [CrossRef] [PubMed]
7. Takahashi, J.; Ikegami, S.; Kuraishi, S.; Shimizu, M.; Futatsugi, T.; Kato, H. Skip pedicle screw fixation combined with Ponte osteotomy for adolescent idiopathic scoliosis. *Eur. Spine J.* **2014**, *23*, 2689–2695. [CrossRef]
8. Mukaiyama, K.; Takahashi, J.; Hirabayashi, H.; Ogihara, N.; Kuraishi, S.; Shimizu, M.; Kato, H. Factors influencing the residual rib hump after posterior spinal fusion for adolescent idiopathic scoliosis with Lenke 1 and 2 curves. *J. Orthop. Sci.* **2013**, *18*, 687–692. [CrossRef]
9. Uehara, M.; Takahashi, J.; Kuraishi, S.; Shimizu, M.; Ikegami, S.; Futatsugi, T.; Oba, H.; Kato, H. Computer-assisted skip pedicle screw fixation for adolescent idiopathic scoliosis. *J. Orthop. Sci.* **2017**, *22*, 218–223. [CrossRef]
10. Kotani, Y.; Abumi, K.; Ito, M.; Minami, A. Improved accuracy of computer-assisted cervical pedicle screw insertion. *J. Neurosurg.* **2003**, *99* (Suppl. 3), 257–263. [CrossRef]
11. Cobb, J.R. Outline for the study of scoliosis. *Am. Acad. Orthop. Surg. Inst. Course Lect.* **1948**, *5*, 261–275.
12. Haher, T.R.; Gorup, J.M.; Shin, T.M.; Homel, P.; Merola, A.A.; Grogan, D.P.; Pugh, L.; Lowe, T.G.; Murray, M. Results of the Scoliosis Research Society instrument for evaluation of surgical outcome in adolescent idiopathic scoliosis, A multicenter study of 244 patients. *Spine* **1999**, *24*, 1435–1440. [CrossRef] [PubMed]
13. Hashimoto, H.; Sase, T.; Arai, Y.; Maruyama, T.; Isobe, K.; Shouno, Y. Validation of a Japanese version of the Scoliosis Research Society-22 Patient Questionnaire among idiopathic scoliosis patients in Japan. *Spine* **2007**, *32*, E141–E146. [CrossRef]
14. Fukui, M.; Chiba, K.; Kawakami, M.; Kikuchi, S.; Konno, S.; Miyamoto, M.; Seichi, A.; Shimamura, T.; Shirado, O.; Taguchi, T.; et al. Subcommittee of the Clinical Outcome Committee of the Japanese Orthopaedic Association on Low Back Pain and Cervical Myelopathy Evaluation. JOA back pain evaluation questionnaire (JOABPEQ)/JOA Cervical Myelopathy evaluation questionnaire (JOACMEQ), The report on the development of revised versions, April 16, 2007, The Subcommittee of the clinical outcome committee of the Japanese orthopaedic association on low back pain and Cervical Myelopathy evaluation. *J. Orthop. Sci.* **2009**, *14*, 348–365. [PubMed]

15. Li, S.; Chen, Z.H.; Qiu, Y.; Xu, L.; Chen, X.; Du, C.Z.; Zhu, Z.Z.; Sun, X. Coronal decompensation after posterior-only thoracolumbar hemivertebra resection and short fusion in young children with congenital scoliosis. *Spine* **2018**, *43*, 654–660. [CrossRef] [PubMed]
16. Kwan, M.K.; Chiu, C.K.; Tan, P.H.; Chian, X.H.; Ler, X.Y.; Ng, Y.H.; Ng, S.J.; Goh, S.H.; Chan, C.Y.W. Radiological and clinical outcome of selective thoracic fusion for patients with Lenke 1C and 2C adolescent idiopathic scoliosis with a minimum follow-up of 2 years. *Spine J.* **2018**, *18*, 2239–2246. [CrossRef] [PubMed]
17. Cho, R.H.; Yaszay, B.; Bartley, C.E.; Bastrom, T.P.; Newton, P.O. Which Lenke 1A curves are at the greatest risk for adding-on … and why? *Spine* **2012**, *37*, 1384–1390. [CrossRef] [PubMed]
18. Shimizu, M.; Takahashi, J.; Ikegami, S.; Kuraishi, S.; Futatsugi, T.; Kato, H. Are pedicle screw perforation rates influenced by registered or unregistered vertebrae in multilevel registration using a CT-based navigation system in the setting of scoliosis? *Eur. Spine J.* **2014**, *23*, 2211–2217. [CrossRef]
19. Belmont, P.J., Jr.; Klemme, W.R.; Dhawan, A.; Polly, D.W., Jr. In vivo accuracy of thoracic pedicle screws. *Spine* **2001**, *26*, 2340–2346. [CrossRef]
20. Suk, S.I.; Kim, W.J.; Lee, S.M.; Kim, J.H.; Chung, E.R. Thoracic pedicle screw fixation in spinal deformities: Are they really safe? *Spine* **2001**, *26*, 2049–2057. [CrossRef]
21. Behrbalk, E.; Uri, O.; Parks, R.M.; Grevitt, M.P.; Rickert, M.; Boszczyk, B.M. Posterior-only correction of Scheuermann kyphosis using pedicle screws: Economical optimization through screw density reduction. *Eur. Spine J.* **2014**, *23*, 2203–2210. [CrossRef] [PubMed]
22. Cheung, K.M.; Natarajan, D.; Samartzis, D.; Wong, Y.W.; Cheung, W.Y.; Luk, K.D. Predictability of the fulcrum bending radiograph in scoliosis correction with alternate-level pedicle screw fixation. *J. Bone Joint Surg. Am.* **2010**, *92*, 169–176. [CrossRef] [PubMed]
23. Uehara, M.; Takahashi, J.; Ikegami, S.; Oba, H.; Kuraishi, S.; Futatsugi, T.; Takizawa, T.; Munakata, R.; Koseki, M.; Kato, H. Determination of optimal screw number based on correction angle for main thoracic curve in adolescent idiopathic scoliosis. *J. Orthop. Sci.* **2019**, *24*, 415–419. [CrossRef] [PubMed]
24. Crawford, A.H.; Lykissas, M.G.; Gao, X.; Eismann, E.; Anadio, J. All-pedicle screw versus hybrid instrumentation in adolescent idiopathic scoliosis surgery: A comparative radiographical study with a minimum 2-year follow-up. *Spine* **2013**, *38*, 1199–1208. [CrossRef]
25. Hwang, C.J.; Lee, C.K.; Chang, B.S.; Kim, M.S.; Yeom, J.S.; Choi, J.M. Minimum 5-year follow-up results of skipped pedicle screw fixation for flexible idiopathic scoliosis. *J. Neurosurg. Spine* **2011**, *15*, 146–150. [CrossRef]
26. Ishikawa, M.; Cao, K.; Pang, L.; Watanabe, K.; Yagi, M.; Hosogane, N.; Machida, M.; Shiono, Y.; Nishiyama, M.; Fukui, Y.; et al. Postoperative behavior of thoracolumbar/lumbar curve and coronal balance after posterior thoracic fusion for Lenke 1C and 2C adolescent idiopathic scoliosis. *J. Orthop. Sci.* **2015**, *20*, 31–37. [CrossRef]
27. Skaggs, D.L.; Seehausen, D.A.; Yamaguchi, K.T., Jr.; Hah, R.J.; Wright, M.L.; Bumpass, D.B.; Kim, H.J.; Andras, L.M.; Vitale, M.G.; Lenke, L.G. Assessment of Lowest Instrumented Vertebra Tilt on Radiographic Measurements in Lenke "C" Modifier Curves Undergoing Selective Thoracic Fusion in Adolescent Idiopathic Scoliosis. *Spine Deform.* **2016**, *4*, 125–130. [CrossRef]
28. Tominaga, R.; Sekiguchi, M.; Yonemoto, K.; Kakuma, T.; Konno, S.I. Establishment of reference scores and interquartile ranges for the Japanese Orthopaedic Association Back Pain Evaluation Questionnaire (JOABPEQ) in patients with low back pain. *J. Orthop. Sci.* **2018**, *23*, 643–648. [CrossRef]
29. Ogura, Y.; Ogura, K.; Kobayashi, Y.; Kitagawa, T.; Yonezawa, Y.; Takahashi, Y.; Yoshida, K.; Yasuda, A.; Shinozaki, Y.; Ogawa, J. Minimally clinically important differences for the Japanese Orthopaedic Association Back Pain Evaluation Questionnaire (JOABPEQ) following decompression surgery for lumbar spinal stenosis. *J. Clin. Neurosci.* **2019**, *69*, 93–96. [CrossRef]
30. Bharucha, N.J.; Lonner, B.S.; Auerbach, J.D.; Kean, K.E.; Trobisch, P.D. Low-density versus high-density thoracic pedicle screw constructs in adolescent idiopathic scoliosis: Do more screws lead to a better outcome? *Spine J.* **2013**, *13*, 375–381. [CrossRef]
31. Mariconda, M.; Galasso, O.; Barca, P.; Milano, C. Minimum 20-year follow-up results of Harrington rod fusion for idiopathic scoliosis. *Eur. Spine J.* **2005**, *14*, 854–861. [CrossRef]
32. Akazawa, T.; Minami, S.; Kotani, T.; Nemoto, T.; Koshi, T.; Takahashi, K. Long-term clinical outcome of surgery for adolescent idiopathic scoliosis 21 to 41 years later. *Spine* **2012**, *37*, 402–405. [CrossRef] [PubMed]
33. Götze, C.; Liljenqvist, U.R.; Slomka, A.; Götze, H.G.; Steinbeck, J. Quality of life and back pain: Outcome 16.7 years after Harrington instrumentation. *Spine* **2002**, *27*, 1456–1463. [CrossRef] [PubMed]

34. Carreon, L.Y.; Sanders, J.O.; Diab, M.; Sucato, D.J.; Sturm, P.F.; Glassman, S.D.; Spinal Deformity Study Group. The minimum clinically important difference in Scoliosis Research Society-22 Appearance, Activity, and Pain domains after surgical correction of adolescent idiopathic scoliosis. *Spine* **2010**, *35*, 2079–2083. [CrossRef] [PubMed]
35. Newton, P.O.; Yaszay, B.; Upasani, V.V.; Pawelek, J.B.; Bastrom, T.P.; Lenke, L.G.; Lowe, T.; Crawford, A.; Betz, R.; Harms Study Group; et al. Preservation of thoracic kyphosis is critical to maintain lumbar lordosis in the surgical treatment of adolescent idiopathic scoliosis. *Spine* **2010**, *35*, 1365–1370. [CrossRef] [PubMed]
36. Solberg, T.K.; Sørlie, A.; Sjaavik, K.; Nygaard, Ø.P.; Ingebrigtsen, T. Would loss to follow-up bias the outcome evaluation of patients operated for degenerative disorders of the lumbar spine? *Acta Orthop.* **2011**, *82*, 56–63. [CrossRef] [PubMed]

Journal of
Clinical Medicine

MDPI

Article

Knee Injury and Osteoarthritis Outcome Score (KOOS) Responder Criteria and Minimal Detectable Change 3–12 Years Following a Youth Sport-Related Knee Injury

Clodagh M. Toomey [1,2,3,4,*], Jackie L. Whittaker [1,5,6], Luz Palacios-Derflingher [1,7] and Carolyn A. Emery [1,3,4,7,8]

[1] Sport Injury Prevention Research Centre, Faculty of Kinesiology, University of Calgary, Calgary, AB T2N 1N4, Canada; jackie.whittaker@ubc.ca (J.L.W.); lmpalaci@ucalgary.ca (L.P.-D.); caemery@ucalgary.ca (C.A.E.)
[2] School of Allied Health, Faculty of Education and Health Sciences, University of Limerick, V94 T9PX Limerick, Ireland
[3] McCaig Institute for Bone and Joint Health, University of Calgary, Calgary, AB T2N 4Z6, Canada
[4] Alberta Children's Hospital Research Institute, University of Calgary, Calgary, AB T2N 4N1, Canada
[5] Department of Physical Therapy, Faculty of Medicine, University of British Columbia, Vancouver, BC V6T 1Z3, Canada
[6] Arthritis Research Canada, Richmond, BC V6X 2C7, Canada
[7] Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB T2N 4N1, Canada
[8] Department of Pediatrics, Cumming School of Medicine, University of Calgary, Calgary, AB T2N 4N1, Canada
* Correspondence: clodagh.toomey@ucalgary.ca

Citation: Toomey, C.M.; Whittaker, J.L.; Palacios-Derflingher, L.; Emery, C.A. Knee Injury and Osteoarthritis Outcome Score (KOOS) Responder Criteria and Minimal Detectable Change 3–12 Years Following a Youth Sport-Related Knee Injury. *J. Clin. Med.* **2021**, *10*, 522. https://doi.org/10.3390/jcm10030522

Academic Editor: Jos Runhaar
Received: 20 December 2020
Accepted: 29 January 2021
Published: 1 February 2021

Abstract: The applicability of thresholds that constitute an acceptable score or meaningful change on the Knee injury and Osteoarthritis Outcome Score (KOOS) in cohorts ≥ 5 years following knee injury is not well understood. The primary objective of this study was to evaluate the association between intra-articular knee injury type and two different KOOS pain thresholds (patient acceptable symptom state (PASS) and Englund symptomatic knee criteria) in the Alberta Youth Prevention of Osteoarthritis (PrE-OA) cohort, which includes participants 3–12 years following a youth sport-related knee injury and uninjured controls with similar age, sex and sport characteristics. Analyses accounted for sex, time since injury and the interaction between time since injury and injury type. Secondary objectives were to report proportions meeting thresholds for KOOS outcomes and minimal detectable change (MDC) from published test–retest reliability data, over a 1–4-year follow-up. Two hundred and fifty-three (253) participants (124 injured, 129 controls) were included in analyses, of which 153 (77 injured, 76 controls) had follow-up data. Similar odds were observed for presence of pain (below PASS threshold) in participants with anterior cruciate ligament (ACL)/meniscus injury (odds ratio (OR) 4.2 (97.5% confidence interval (CI): 1.8, 9.9)) and other knee injuries (OR 4.9 (97.5% CI: 1.2, 21.0)), while there were higher odds for presence of Englund "symptomatic knee" criteria in participants with ACL/meniscus injury (OR 13.6 (97.5% CI: 2.9, 63.4)) than other knee injuries (OR 7.3 (97.5% CI: 0.8, 63.7)) compared to controls. After a median 23.4 (8 to 42) month follow-up, 35% of previously injured participants had at least one KOOS sub-scale score that worsened by more than the MDC published threshold. Despite limited research, this study shows that individuals with youth sport knee injuries other than ACL or meniscus injury may also experience significant pain and symptoms 3–12 years following injury. Replication and further follow-up are needed to identify a possible clinical trajectory towards osteoarthritis.

Keywords: knee; ACL; injury; osteoarthritis; KOOS; musculoskeletal disorders; symptoms; pain

1. Introduction

Much sport medicine and musculoskeletal health research has focused on rehabilitation and recovery following an anterior cruciate ligament (ACL) injury due to the immedi-

ate and long-term effects on activity participation and health. The negative consequences on health outcomes that are present in this population in the short-term (1–2 years) [1,2] and medium-term (3–12 years) [3–5] interval following injury, implicate knee injuries as one of the largest public health injury-related burdens. Of foremost concern is the higher likelihood of development of knee joint osteoarthritis (OA) after sustaining an ACL or other traumatic intra-articular knee injury [6], sentencing a proportion of these individuals to a lifetime of knee pain, symptoms, reduced function and quality of life (QOL). Of these patient-reported outcomes (PRO), knee pain is the most common symptom of OA [7], but defining what constitutes clinically significant pain is not well-understood in individuals following injury and prior to onset of potential joint disease.

The Knee injury and Osteoarthritis Outcome Score (KOOS) is a common PRO used to track the most prevalent clinical features across the timeline from knee injury to OA [8,9]. An alternative and more clinically interpretable method of observing KOOS has been to use defined thresholds of "patient acceptable symptom state" (PASS) in cohorts up to 5 years post-injury to understand what is deemed "acceptable" to a patient on each 0–100 sub-scale. Ingelsrud et al. [10] used anchor-based questions of a current satisfactory state regarding knee function in the Norwegian Knee Ligament Registry (NKLR) and took mean scores for KOOS subscales, where 'yes' was answered. Muller et al. [11] used a similar method in an ACL cohort 1–5 years from injury using the receiver operating characteristic method. Reports of PASS in these ACL cohorts construe that symptoms tend to stabilize and do not change much from 1 to 2 years postoperatively. Other methods to identify KOOS thresholds have used a consensus expert panel to define patients with a symptomatic knee, significant enough to seek medical attention, as applied by Englund et al. [12] to a cohort 16 years post-meniscectomy. Examining these thresholds in the medium-term interval following injury may provide an indication of the trajectory of PRO scores in a population at risk of OA. In addition, knowledge of outcomes in knee injuries other than ACL or meniscus injury is currently lacking.

The interpretation of longitudinal changes in PROs in the years following a knee injury is critical to identify individuals who may be on a trajectory to post-traumatic OA. This can be achieved using the minimal detectable change (MDC), to judge change that surpasses the instrument test–retest reliability [13]. These values may help to identify patients whose scores worsen over time and require an early post-traumatic OA intervention. It may also be helpful for clinicians to understand the baseline variables of other physiological (body composition) and performance (knee strength and function) outcomes in those who worsen over time, as this may aid with identification of at-risk individuals and mechanisms to target during intervention. This type of analysis has not previously been applied to longitudinal data of cohorts with different types of knee injuries.

The primary objective of this cohort study was to evaluate the association of intra-articular knee injury history type (uninjured, ACL and/or meniscal injury or other knee injury) with two types of KOOS criteria to define a painful or symptomatic knee, adjusting for time since injury and sex in a cohort 3–12 years after injury. Secondary objectives were to report the prevalence of acceptable or non-symptomatic PRO's using defined KOOS thresholds and comparing all injury types in this cohort. An additional secondary objective was to evaluate the MDC in KOOS outcomes over a 1–4-year longitudinal follow-up, describing the clinical features and functional performance of injured participants who worsened on at least one sub-scale over time.

2. Materials and Methods

2.1. Participants and Recruitment

Participants include the entire cohort of the Alberta Youth Prevention of Early Osteoarthritis (PrE-OA) study, a historical longitudinal cohort study carried out at the Sport Injury Prevention Research Centre, Faculty of Kinesiology, University of Calgary. The PrE-OA cohort consists of youth and young adults who sustained an intra-articular knee injury 3–12 years prior to study recruitment, while participating in sport under the age

of 18 years. Uninjured controls of similar age (≤1 year), sex and sport (at the time of injury) were recruited to match the injured participants. Detailed recruitment procedures have been described previously [4,14]. Knee injury was defined as a clinical diagnosis of knee ligament, meniscal, or other intra-articular tibiofemoral or patellofemoral injury that occurred playing sport and required both medical consultation and disrupted regular sport participation. Uninjured participants reported no previous time-loss knee injury. Participant exclusion criteria included pregnancy, non-steroidal anti-inflammatory use or cortisone injection within three months prior to testing, a musculoskeletal injury within the previous three months prior to testing that resulted in time loss (work, school or sport), diagnosis of other arthritides, or any current medical problem that prevented participation in the functional testing aspect of the study (e.g., neurological conditions). Participants were recruited between 2013 and 2017 for their initial visit and were invited to return for repeat testing annually for two subsequent visits. If a participant withdrew from the study before completion, every effort was made to replace that participant with an individual with the same characteristics of age (≤1 year), sex and sport at the time of injury. Participants were no longer considered 'uninjured' if they sustained a knee injury in the follow-up period. All tests were performed on the same day with random order of stations and rests between tests to minimize fatigue. Ethics approval was granted from the Conjoint Health Research Ethics Board at the University of Calgary, Canada (REB-14-2212), and all participants provided signed informed consent or assent where applicable.

2.2. Outcome Measures

Participants completed a study questionnaire that gathered demographic details, sport participation and medical, injury and surgery history detail. The Knee Injury and Osteoarthritis Outcome Score (KOOS) survey was completed, which provided information on knee function related to five sub-scales: pain, other symptoms, activities of daily living (ADL), sport and recreation and quality of life (QOL). The KOOS has been validated in knee injury populations and has high test–retest reliability [9], with higher scores indicating better outcomes. Two criteria were applied to the KOOS scores to determine the proportion of participants meeting an acceptable threshold (PASS) [10,11] for each subscale and KOOS4 [10,15] (average KOOS score of all subscales except KOOS ADL). Additional criteria were applied to identify a "symptomatic knee" based on Englund et al.'s [12] thresholds—defined by having QOL and at least two other subscales below the cut-off. For longitudinal assessment, minimal detectable change (MDC) criteria were used to observe change that surpassed the test–retest reliability of KOOS. As a conservative approach, the upper limits reported by Collins et al. [13] were used. Thresholds for each criterion are presented in Table 1.

Table 1. Thresholds for The Knee injury and Osteoarthritis Outcome Score (KOOS) corresponding to the patient acceptable symptom state (PASS) and symptomatic knee for cross-sectional data, and minimal detectable change (MDC) for longitudinal change.

	Thresholds			
	PASS [10] (Ingelsrud et al.)	**PASS [11] (Muller et al.)**	**Symptomatic Knee [12] ¥ (Englund et al.)**	**MDC [13] (Collins et al.)**
Pain	89	88.9	86.1	6.1
Symptoms	83	57.1	85.7	8.5
ADL	95	100	86.8	8.0
Sport/Rec	72	75	85.0	12.0
QOL	73	62.5	87.5	7.2
KOOS 4	79	-	-	-

ADL, activities of daily living; Sport/Rec, sport and recreation; QOL, quality of life. ¥ Definition of "symptomatic knee" based on QOL threshold plus at least two other subscale thresholds.

As reported in previous publications [4,14], participants' knee extensor strength, hop performance and body composition were measured. Normalized knee extensor isometric strength was assessed using handheld isometric dynamometry (model 01163; Lafayette Instrument, Lafayette, IN, USA) as described previously [3]. The peak isometric strength (N) scores were converted to torque values (N/m; force distance between joint line and dynamometer position) and normalized to body weight ($N/m/kg^{-1}$). The maximum distance across two trials of the Triple Single Leg Hop (TLSH) test [3] was recorded for each leg and expressed as a percentage of leg length. Total body adiposity was assessed using dual-energy X-ray absorptiometry (DXA; Hologic Discovery (Hologic, Inc., Marlborough, MA, USA)). Whole-body fat mass (kg) was divided by height (m) squared to derive a fat mass index (FMI; kg/m^2), to allow an accurate representation of adiposity that is not confounded by the presence of lean mass.

2.3. Statistical Analyses

All data were managed in REDCap (Research Electronic Data Capture) and statistical analyses were completed using Stata Version 15.1 (StataCorp LP, College Station, TX, USA). All outcomes were described using median with interquartile range (IQR) by injury type. Control participants were excluded from the analysis if there was no injured case with matching age, sex and sport criteria. For the primary objective and cross-sectional analysis, outcomes collected at the final measurement (furthest time-point from injury) were included due to a greater likelihood of observing changes related to post-traumatic OA. Multivariable logistic regression clustered on matched characteristics of sex and sport, and adjusted for time since exposure, sex and the interaction between injury type and time since exposure, was used to evaluate the odds (odds ratio (OR) and 97.5% confidence interval (CI)) of significant pain (Model 1) or a symptomatic knee (Model 2) by knee injury type (no knee injury history, grade III ACL tear and/or meniscal injury history, other knee injury history). Bonferroni adjustment was applied to CIs since there were two primary outcomes ($\alpha = 0.025$). The outcome for Model 1 was the KOOS pain PASS criteria consistent from both Ingelsrud et al. [10] and Muller et al. [11] due to a high prevalence of pain in symptomatic OA. The outcome for Model 2 was the Englund et al. [12] criteria for a symptomatic knee. For the control participants, time since exposure was coded the same as that of the matched injured participants on recruitment corresponding to an equivalent injury-free time. Age was not included in the models due to collinearity with time since exposure. A backwards-stepwise elimination approach was employed, and the most parsimonious models reported using likelihood ratio tests and checking for a 10% difference in beta coefficients. For longitudinal analyses, the change in KOOS scores from the first and final follow-up for all participants was calculated and the frequency and proportions of individuals whose KOOS scores changed greater than MDC in either direction were reported. Baseline clinical features of injured participants who had $\geq$1 KOOS outcome that worsened over time using MDC were described alongside those who improved or stayed the same using median and IQR, or frequencies and proportions.

3. Results

The final dataset was comprised of 253 participants (124 previously injured participants ($n = 89$ ACL and/or meniscal injury and $n = 35$ other knee injury) and 129 uninjured controls). A flowchart of participant recruitment and follow-up for these analyses is shown in Figure 1. Females comprised 55% of the sample and the median participant age was 24.2 years (range 15–30), median body mass index (BMI) 24.4 kg/m^2 (range 18.1–38.9) at the furthest time-point from exposure. The median age of previously injured participants was 15.4 years (range 9–18 years) at the time of injury and 7.8 years (range 3–12.6 years) post-injury at the time of data collection. Injury type comprised mainly of third-degree ACL sprain ($n = 69$; 56%), all of which underwent surgical reconstruction. Of these, 51 (74%) also had an associated meniscal lesion. A further 16% ($n = 20$) of the total injured participants had a meniscal injury which did not include a third-degree ACL tear ($n = 17$ isolated

meniscal injury and $n = 3$ associated with another injury). Eight of these underwent arthroscopic surgery. The 'Other Injury' category included other ligamentous injuries (first- to third-degree medial and lateral collateral ligament sprain, or first- to second-degree ACL or posterior collateral ligament sprain without meniscal involvement) ($n = 15$; 12%), patellar dislocations and subluxations ($n = 18$; 15%) and intra-articular fracture ($n = 2$). Injuries were sustained while playing 14 different sports: the most common sports were soccer (50%), basketball (18%) and ski/snowboarding (9%) for females, and ice hockey (42%), soccer (18%) and football (11%) for males. Participants were grouped into 21 clusters based on sex and sport, ranging between 2 and 69 participants per cluster.

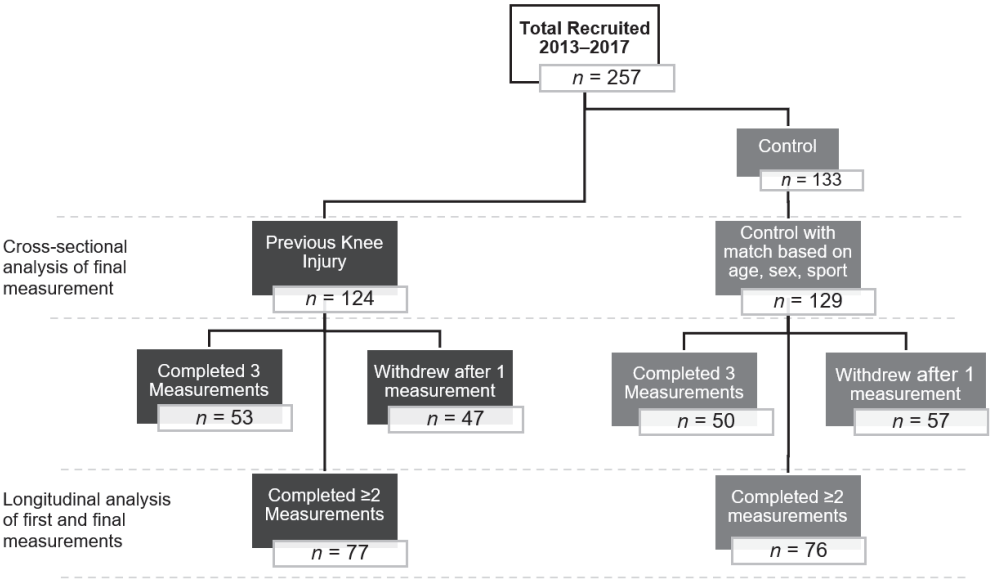

Figure 1. Flowchart of participant recruitment and follow-up for the Alberta Youth Prevention of Early Osteoarthritis (PrE-OA) cohort.

Table 2 displays descriptive statistics for KOOS outcomes by injury type at the furthest time-point from injury, and the proportion of participants meeting thresholds by each criterion. For participants with an ACL and/or meniscal injury, the odds of not achieving an "acceptable" KOOS pain PASS score (Model 1) were 4.2 times (97.5% CI: 1.8, 9.9) the odds for uninjured controls, while those with other knee injuries had 4.9 times (97.5% CI: 1.2, 21.0) the odds of uninjured controls. Sex, time since exposure or interaction terms did not significantly influence these findings. In Model 2, the odds of meeting criteria for a symptomatic knee in those with an ACL and/or meniscus injury were 13.6 times (97.5% CI: 2.9, 63.4) the odds of uninjured controls, while those with other knee injuries had 7.3 times (97.5% CI: 0.8, 63.7) the odds of uninjured controls. Including sex as a covariate improved the likelihood ratio and r^2 of the model but was not a significant predictor (OR 0.5 (97.5% CI: 0.2, 1.2)). Time since exposure or interaction terms did not influence Model 2.

A graphical comparison of median KOOS subscales with thresholds is shown in Figure 2. Comparison is also made with published data from the Multicenter Orthopedic Outcomes Network (MOON) cohort [16] to observe any similarities with cohorts at a further time-point from injury (10 years post-ACL).

Table 2. Descriptive statistics and proportion of participants in the Alberta Youth Prevention of Early Osteoarthritis (PrE-OA) cohort meeting patient acceptable symptom state (PASS) for KOOS outcomes by Ingelsrud [10] and Muller [11] thresholds and proportion passing Englund [12] thresholds.

KOOS	Uninjured (n = 129)				ACL and/or Meniscus (n = 89)				Other Knee Injury (n = 35)			
	Median (IQR)	PASS% Ingelsrud	PASS% Muller	% Englund	Median (IQR)	PASS % Ingelsrud	PASS% Muller	% Englund	Median (IQR)	PASS % Ingelsrud	PASS% Muller	% Englund
Pain	100 (5.6)	91.5	91.5	95.4	94.4 (13.9)	71.9	71.9	80.9	97.2 (11.1)	68.6	68.6	88.6
Sym.	96.4 (10.7)	85.3	100	85.3	85.7 (17.9)	59.6	94.4	59.6	89.3 (10.7)	71.4	100	71.4
ADL	100 (0)	96.1	77.5	99.2	98.5 (4.4)	79.8	48.3	93.3	100 (2.9)	88.6	65.7	94.3
Sport	100 (2.8)	100	99.2	96.9	91.7 (13.9)	95.5	92.1	79.8	97.2 (11.1)	100	100	85.7
QOL	100 (2.8)	100	100	97.7	88.9 (11.1)	96.6	100	61.8	94.4 (5.5)	100	100	85.7
KOOS 4	98.2 (5.1)	98.5	N/A	N/A	90.7 (11.3)	80.9	N/A	N/A	93.7 (6.5)	88.6	N/A	N/A
* Englund	N/A	N/A	N/A	98.5	N/A	N/A	N/A	82.0	N/A	N/A	N/A	91.4

ADL, activities of daily living; IQR, interquartile range; KOOS, knee injury and osteoarthritis outcome score; N/A, not applicable; QOL, quality of life; Sport, sport and recreation; Sym, symptoms. * Proportion that did not have a "symptomatic knee" based on Englund criteria of meeting QOL subscale threshold and at least 2 other subscale thresholds.

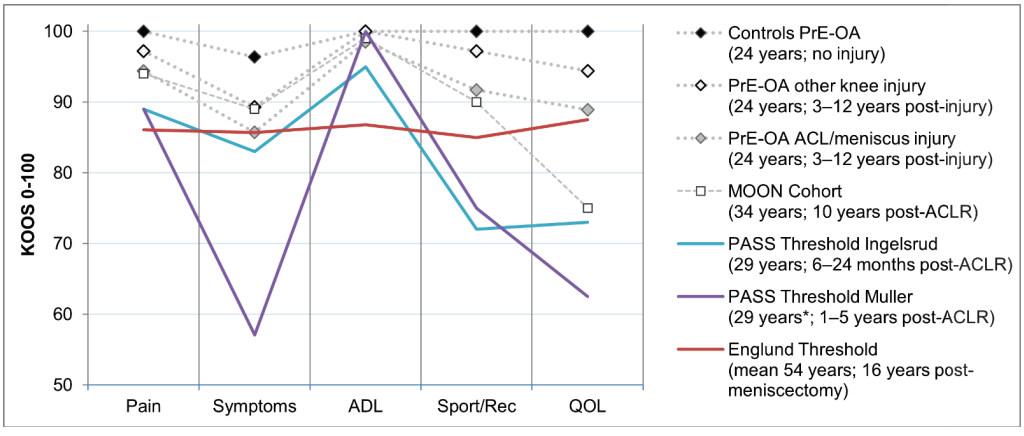

Figure 2. Median of KOOS subscales by injury type in the PrE-OA cohort compared to the Multicenter Orthopedic Outcomes Network (MOON) cohort and KOOS thresholds with median age and time since injury. * Estimate of age based on mean baseline age and time since injury.

A total of 153 participants completed at least two testing sessions (n = 76 uninjured, n = 77 previously injured). The median time between follow-up was 23.4 months (range 8 to 42 months). Reasons for study withdrawal included relocation (16%; 5 cases, 11, controls), not interested (8%; 4 cases, 4 controls), too busy (17%; 9 cases, 9 controls), could not contact (47%; 19 cases, 29 controls), new injury or re-injury (6%; 3 cases, 3 controls), unknown reason (3%; 1 case, 2 controls), pregnancy/recent birth (3%; 1 case, 2 controls) and personal reasons (1%; 1 case). Baseline characteristics for participants who were lost to follow-up are compared with those who returned for follow-up in Appendix A.

The proportion of participants whose KOOS scores were deemed to have "changed" based on MDC criteria are presented in Table 3. To observe clinical features of participants with worse KOOS scores on follow-up, descriptive statistics are presented for injured youth and young adults with at least one KOOS subscale that worsened (n = 27) in Table 4. There were nine injured participants who worsened on at least two subscales, four who worsened on at least three and one who worsened on four subscales. Participants worsened on the symptoms' subscale most commonly (n = 18), followed by pain and QOL (n = 16 for each), ADL (n = 4) and sport and recreation (n = 4).

Table 3. Frequency and proportion (n (%)) of uninjured and injured participants with improved or worse KOOS scores by minimal detectable change (MDC) in a 23.4 months (range 8 to 42 months) follow-up.

	Uninjured (n = 76) MDC n (%)		Previous Injury (n = 77) MDC n (%)	
	Improved	Worse	Improved	Worse
Pain	4 (5.3)	5 (6.6)	14 (18.2)	11 (14.3)
Symptoms	9 (11.8)	5 (6.6)	14 (18.2)	13 (16.9)
ADL	1(1.3)	1(1.3)	5 (6.5)	3 (3.9)
Sport/Rec	1(1.3)	1(1.3)	5 (6.5)	3 (3.9)
QOL	6 (7.9)	5 (6.6)	7 (9.1)	11 (14.3)
KOOS 4	-	-	-	-

Table 4. Baseline clinical features for injured participants who had worse KOOS scores at follow-up (at least one KOOS subscale worse than the minimal detectable change) and other injured and uninjured participants using mean (95% confidence interval (CI)) or median (25th, 75th centile).

Outcome at Baseline	Injured Participants with Worse KOOS Score(s) at Follow-Up (n = 27)	Injured Participants with No Change or Improved KOOS Scores at Follow-Up (n = 50)	Uninjured (n = 76)
Sex (%)	56% Female	54% Female	53% Female
Age (years) *	22.6 (20.4, 24.5)	23.6 (21.5, 24.9)	23.0 (20.5, 24.5)
Time since injury (years) *	6.0 (4.8, 8.7)	7.0 (5.4, 8.0)	-
Time between tests (m) *	23.6 (18.4, 24.8)	23.4 (17.5, 25.1)	23.3 (14.5, 25.8)
Injury type	82% ACL and/or meniscus 18% Other knee injury	64% ACL and/or meniscus 36% Other knee injury	-
Body mass index (kg/m^2)	25.4 (24.0, 26.9)	24.8 (23.8, 25.9)	23.4 (22.8, 24.0)
Fat mass index (kg/m^2)	6.10 (5.09, 7.10)	5.56 (4.84, 6.27)	4.67 (4.24, 5.09)
Knee extension strength (N/m/kg^{-1})	1.77 (1.51, 2.03) $^{\Delta}$	1.89 (1.74, 2.04)	2.00 (1.88, 2.13)
Triple single leg hop (cm)	411.3 (380, 442)	427.6 (405, 451)	443.0 (424.2, 461.7) $^{¥}$
KOOS			
Pain *	97.2 (91.7, 100)	93.1 (86.1, 97.2)	100 (97.2, 100)
Symptoms *	85.7 (78.6, 96.4)	89.3 (75.0, 96.4)	96.4 (89.3, 100)
ADL *	98.5 (97.1, 100)	98.5 (95.6, 100)	100 (100, 100)
Sport/Rec *	97.2 (88.9, 100)	94.4 (86.1, 97.2)	100 (95.8, 100)
QOL *	91.7 (83.3, 97.2)	91.7 (96.1, 94.4)	100 (97.2, 100)

* Median (IQR), $^{\Delta}$ n = 26, $^{¥}$ n = 75.

4. Discussion

This study applied a number of different criteria to KOOS outcomes to determine a painful or symptomatic knee in the 3–12-year interval following intra-articular knee injury. It is the first study to apply these methods during this interval and across a variety of knee injuries and matched controls, in an effort to characterize the period between injury and potential knee OA development in youth and young adults.

Multivariable logistic regression showed higher odds of having a painful or symptomatic knee for each criteria in both those with an ACL and/or meniscus injury and other

sub-groups of injury compared to uninjured controls. The focus of knee injury research and major cohort studies has traditionally been on ACL or meniscus injury due to a high injury burden and significant risk for joint disease. The results of this study suggest that other sub-groups of injury (e.g., collateral ligament and patellar dislocations and subluxations) may carry a high risk for a painful or symptomatic knee requiring medical attention. A lower proportion in this sub-group may have lasting problems across sub-scales: there were 9% (n = 3) reporting a symptomatic knee, compared to 18% (n = 16) in the ACL/meniscus group. However, of note, 31% (n = 11) had significant pain on the PASS criteria compared to 28% (n = 25) in the ACL/meniscus group. Since the experience of pain drives healthcare usage and is the top concern of people living with knee OA [17], individuals with other types of knee injury are also worthy of further investigation and research. While it is not known which patients will progress to development of joint disease, it appears as though individuals with traumatic knee injuries that are typically considered "less severe" compared to ACL and meniscus injuries have a similar level of pain and symptoms at 3–12 years following injury.

Sex was not observed to be a significant covariate in the regression models, even though it has been reported that females often report worse knee KOOS outcomes in the first two years following ACL injury [18]. Despite the wide range of post-exposure years included and an expectation that more adverse outcomes consistent with osteoarthritis would be present further from injury, time since injury did not influence the findings. It is worth considering that time since injury may have an impact on the actual thresholds applied since the population assessed is different to the reference populations. Nonetheless, this finding appears to be consistent with other large prospective cohort studies of ACL injury. The MOON cohort reported similar median (IQR) scores for KOOS pain at two years (92 (83–97)), six years (94 (86–100)) and ten years (94 (86–100)) of follow-up [18]. The proportion of individuals in the MOON cohort who satisfied Englund criteria for a symptomatic knee were 43% at two years and 39% at six years [19]. This is higher than the reported 18% of ACL/meniscus injury participants in the current study and is likely because this threshold is primarily based on a low-QOL KOOS sub-scale. As observed in Figure 2, median scores for 10-year data in the MOON cohort show similarities to the PrE-OA cohort for pain, symptoms, ADL and sports/recreation subscales, but are much lower for the QOL sub-scale. It is unknown why this difference between cohorts exists, but could be related to regional differences between populations, healthcare experiences or a younger age ($\leq$18 years) at the time of injury in the PrE-OA cohort in comparison to a baseline median age of 24 years in the MOON cohort. Younger patients have been shown to score higher outcomes post-ACL injury [20].

The development and application of thresholds for PROs is an important step in the clinical identification of patients deemed at risk of adverse symptoms or requiring intervention. However, it is unknown whether thresholds developed on a specific cohort at one point in time can be applied to other cohorts at different time-points from injury. For this reason, three different criteria were applied to the current data, at a median of 7.8 years following knee injury. Across KOOS sub-scales, 59.6% to 96.6% (ACL and/or meniscus) and 68.6% to 100% (other knee injury) had satisfactory scores above the Ingelsrud et al. PASS thresholds established on the NKLR cohort two years after ACL reconstruction. PASS proportions were quite similar using Muller et al. thresholds, with the exception of symptoms, which had a much lower threshold (94.4% to 100%), and ADL, which had a high threshold (48.3% to 65.7%). Although this reference data had a larger follow-up time of 3.4 years (range 1–5 years), these two thresholds did not appear to fit well with the current data or compared to the NKLR cohort. This may be due, in part, to the different analysis approach taken using a receiver operating characteristic (ROC) curve, with lower sensitivity and specificity values reported for symptoms and ADL. One other study [21] applied Muller et al. PASS thresholds to an ACL cohort 10 years following injury in the Swedish National Knee Ligament Register (SNKLR). Proportions of participants reporting satisfactory outcome by sub-scale matched well with the current study, but again were

lower on QOL (100% for PrE-OA ACL/meniscus vs. 67.9% for SNKLR), and also sport and recreation (92.1% for PrE-OA ACL/meniscus vs. 57.2% for SNKLR). Similarly, this is likely explained by an older age at time of injury (median 24.8 years).

In order to better understand the time period following knee injury, a longitudinal assessment of PRO's was explored as a secondary objective. Change in scores was assessed using MDC criteria over a 1–4-year follow-up to identify the extent of outcome decline over a short period. Results were quite variable on different sub-scales and revealed a similar proportion of injured participants that improved (6.5–18.2%) as those that worsened (3.9–16.9%). This may suggest that fluctuations or flare-ups in symptoms are a common feature in this timeframe. Surprisingly, the number of uninjured participants who had an important or detectable improvement (1.3–11.8%) or worsening (1.3–6.6%) on follow-up was higher than expected and underscores the importance of having a control group in these studies. However, the low number of individuals involved warrant some caution regarding interpretation.

Of the injured participants, 35% worsened on at least one KOOS sub-scale in the 1–4-year follow-up. While we did not power the study to detect significant differences on this outcome, observation of baseline characteristics showed that the group who worsened may have been younger, with a higher proportion of ACL and/or meniscus injuries. Mean values for body mass index and fat mass index were higher in those who worsened, consistent with previous evidence of body mass index as a risk factor for worse two-, six- [19] and ten-year outcomes [18]. However, this was unlikely to be a statistically significant difference in these sub-groups. Values for knee extension strength and triple single leg hop appear to be lower at baseline for those with worse KOOS outcomes at follow-up, but this is unlikely to be clinically or statistically significant. Further research is required to confirm this data and may help clinicians to distinguish patients that are likely to require more attention and intervention.

A limitation of this study is that responder criteria and thresholds from an ACL cohort two years post-reconstruction, an ACL cohort one to five years post-reconstruction and a meniscectomy cohort 16 years post-surgery were applied to youth and young adults 3–12 years following knee injury. It is important to consider that thresholds used were not population-specific and this is acknowledged in the interpretation of these findings. Responder criteria thresholds have not previously been applied to injury types other than ACL or meniscus injury. We investigated criteria developed on other cohorts in order to provide comparison across populations and for lack of a more specific option to define post-injury knee problems or early symptomatic OA. An advantage of this approach was the ability to compare published thresholds alongside uninjured matched controls and sub-groups of injury. This evaluation would benefit from replication with a larger number of participants for sub-group analysis. As demonstrated by the wide confidence intervals for multivariable logistic regression, there were likely not enough events per variable which introduced sparse data bias to the analysis. Despite this, a significant association was observed. In addition, numbers were not large enough to statistically account for covariates other than sex and time since injury, such as body or fat mass index, surgery, re-injury or physical activity. Age was not included in models due to collinearity with time since exposure. However, a sensitivity analysis with age included instead of time since exposure did not change the results. Observation of a cohort who were inured under the age of 18 years was a strength of this study since incidence of ACL injury is increasing most rapidly in this age group [22,23], and they may suffer from more long-term health outcomes [24]. Due to the time period under investigation, it was not possible to control for the specific mechanism of injury or type of rehabilitation in the acute injury stage, which may influence the impact on cartilaginous structures, recovery and thus long-term outcomes. Loss to one year of follow-up in this study was 38% for injured participants and 44% for controls, which poses a threat to validity of the longitudinal assessment. Considerable effort was made by the study coordinators to retain participants and withdrawal was mainly due to relocation, university attendance and contact phone numbers that were no longer in use.

While the number lost to follow-up is higher compared to other knee injury cohorts, the nature and length of the testing protocol (full procedure included blood draw, biomechanics assessment, clinical exam, shuttle run and range of questionnaires in addition to the above) created a higher participant burden than questionnaires alone. However, age, sex, type of injury and time since injury were very similar for those who were retained vs. lost to follow-up. In an attempt to capture more adverse changes over a longer time period, the time between longitudinal follow-ups was wide-ranging (8 to 42 months). In order to retain numbers for analyses, a more targeted approach (e.g., one-year or two-year change) was not used, but future analysis may benefit from this approach.

5. Conclusions

In this younger cohort of individuals with a 3–12-year history of different types of intra-articular knee injury, a high proportion (28–31%) were experiencing significant pain, while 9–18% met the criteria for a symptomatic knee, significant enough to seek medical attention. Odds for significant pain were equally high in those with ACL and/or meniscus injury and other (e.g., collateral ligament) injury compared to controls, demonstrating that a "less severe" injury may also require follow-up and intervention. Over a 1–4-year follow-up, 36% of injured participants worsened on at least one KOOS sub-scale. However, more research is required to validate thresholds specific to populations in the interval between youth knee injury and risk of joint osteoarthritis to accurately observe clinically significant fluctuations in symptoms in this time period. This will aid the development of targeted interventions to decrease the long-term public health burden of youth knee injury.

Author Contributions: Conceptualization, C.M.T., J.L.W., L.P.-D., and C.A.E.; methodology and data collection, C.M.T., J.L.W., and C.A.E.; writing of original draft, C.M.T.; writing—reviewing and editing, C.M.T., J.L.W., L.P.-D., and C.A.E.; funding acquisition, C.A.E. and J.L.W. All authors have read and agreed to the published version of the manuscript.

Funding: The Alberta Youth PrE-OA study was funded by the Canadian Institutes of Health Research (MOP 133597), the Alberta Osteoarthritis Team supported by Alberta Innovates and the Alberta Children's Hospital Research Institute Chair in Pediatric Rehabilitation (Alberta Children's Hospital Foundation). The Sport Injury Prevention Research Centre (SIPRC) is one of ten International Olympic Committee Research Centers focused on Injury and Illness Prevention in Sport. C.A.E. holds a chair in Pediatric Rehabilitation funded by the Alberta Children's Hospital Foundation. J.L.W is supported by a Michael Smith Foundation for Health Research Scholar Award.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Conjoint Health Research Ethics Board at the University of Calgary, Canada (REB-14-2212).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Acknowledgments: The authors acknowledge the assistance of research coordinators Gabriella Nasuti, Jamie Rishaug, and Lisa Loos, and the numerous research assistants/students at the SIPRC as well as the participants.

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. Baseline characteristics of those who completed at least two measurements vs. participants who were lost to follow-up and eligible for study objectives.

Outcome	Completed Follow-Up (n = 153)	Lost to Follow-Up (n = 100)
Injured/Uninjured (% injured)	77/76 (50% injured)	47/53 (47% injured)
Female/Male (% Female)	82/71 (54% female)	56/44 (56% female)
Age at study entry (years) (Median; 25th, 75th percentile)	23.1 (21.3, 24.5)	22.6 (20.9, 24.5)
Time since injury (years) (Median; 25th, 75th percentile)	6.9 (5.4, 8.3)	6.5 (5.2, 8.2)
Type of injury	70% ACL and/or meniscus 30% Other injury	74% ACL and/or meniscus 26% Other injury

References

1. Granan, L.-P.; Forssblad, M.; Lind, M.; Engebretsen, L. The Scandinavian ACL registries 2004–2007: Baseline epidemiology. *Acta Ortopaedica* **2009**, *80*, 563–567. [CrossRef]
2. Failla, M.J.; Logerstedt, D.S.; Grindem, H.; Axe, M.J.; Risberg, M.A.; Engebretsen, L.; Huston, L.J.; Spindler, K.P.; Snyder-Mackler, L. Does Extended Preoperative Rehabilitation Influence Outcomes 2 Years after ACL Reconstruction? A Comparative Effectiveness Study between the MOON and Delaware-Oslo ACL Cohorts. *Am. J. Sports Med.* **2016**, *44*, 2608–2614. [CrossRef] [PubMed]
3. Whittaker, J.L.; Toomey, C.M.; Nettel-Aguirre, A.; Jaremko, J.L.; Doyle-Baker, P.K.; Woodhouse, L.J.; Emery, C.A. Health-related Outcomes after a Youth Sport-related Knee Injury. *Med Sci Sports Exerc.* **2019**, *51*, 255–263. [CrossRef]
4. Toomey, C.M.; Whittaker, J.L.; Nettel-aguirre, A.; Reimer, R.A.; Woodhouse, L.J.; Ghali, B.; Doyle-Baker, P.K.; Emery, C.A. Higher Fat Mass Is Associated With a History of Knee Injury in Youth Sport. *J. Orthop. Sport Phys. Ther.* **2017**, *47*, 80–87. [CrossRef] [PubMed]
5. Lohmander, L.S.; Ostenberg, A.; Englund, M.; Roos, H. High Prevalence of Knee Osteoarthritis, Pain, and Functional Limitations in Female Soccer Players Twelve Years After Anterior Cruciate Ligament Injury. *Arthritis Rheum.* **2004**, *50*, 3145–3152. [CrossRef] [PubMed]
6. Poulsen, E.; Goncalves, G.H.; Bricca, A.; Roos, E.M.; Thorlund, J.B.; Juhl, C.B. Knee osteoarthritis risk is increased 4-6 fold after knee injury—A systematic review and meta-analysis. *Br. J. Sports Med.* **2019**. [CrossRef]
7. Hawker, G.A.; Stewart, L.; French, M.R.; Cibere, J.; Jordan, J.M.; March, L.; Suarez-Almazor, M.; Gooberman-Hill, R. Understanding the pain experience in hip and knee osteoarthritis—An OARSI / OMERACT initiative. *Osteoarthr. Cartil* **2008**, *16*, 415–422. [CrossRef]
8. Roos, E.M.; Lohmander, L.S. The Knee injury and Osteoarthritis Outcome Score (KOOS): From joint injury to osteoarthritis Pilot study. *Health Qual. Life Outcomes* **2003**, *64*. [CrossRef]
9. Collins, N.J.; Prinsen, C.A.C.; Christensen, R.; Bartels, E.M.; Terwee, C.B.; Roos, E.M. Knee Injury and Osteoarthritis Outcome Score (KOOS): Systematic review and meta-analysis of measurement properties. *Osteoarthr. Cartil* **2016**, *24*, 1317–1329. [CrossRef]
10. Ingelsrud, L.H.; Granan, L.; Terwee, C.B.; Engebretsen, L.; Roos, E.M. Proportion of Patients Reporting Acceptable Symptoms or Treatment Failure and Their Associated KOOS Values at 6 to 24 Months After Anterior Cruciate Ligament Reconstruction. *Am. J. Sports Med.* **2015**, *43*, 1902–1907. [CrossRef]
11. Muller, B.; Yabroudi, M.A.; Lynch, A.; Lai, C.-L.; van Dijk, C.N.; Fu, F.H.; Irrgang, J.J. Defining Thresholds for the Patient Acceptable Symptom State for the IKDC Subjective Knee Form and KOOS for Patients Who Underwent ACL Reconstruction. *Am. J. Sports Med.* **2016**, *44*, 2820–2826. [CrossRef] [PubMed]
12. Englund, M.; Roos, E.M.; Lohmander, L.S. Impact of type of meniscal tear on radiographic and symptomatic knee osteoarthritis: A sixteen-year followup of meniscectomy with matched controls. *Arthritis Rheum.* **2003**, *48*, 2178–2187. [CrossRef] [PubMed]
13. Collins, N.J.; Misra, D.; Felson, D.T.; Crossley, K.M.; Roos, E.M. Measures of Knee Function: International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Knee Ou. *Arthritis Care Res.* **2011**, *63*, S208–S228. [CrossRef] [PubMed]
14. Whittaker, J.L.; Woodhouse, L.J.; Nettel-aguirre, A.; Emery, C.A. Outcomes associated with early post-traumatic osteoarthritis and other negative health consequences 3-10 years following knee joint injury in youth sport. *Osteoarthr. Cartil* **2015**, *23*, 1122–1129. [CrossRef] [PubMed]
15. Roos, E.M.; Boyle, E.; Frobell, R.B.; Lohmander, L.S.; Ingelsrud, L.H. It is good to feel better, but better to feel good: Whether a patient finds treatment 'successful' or not depends on the questions researchers ask. *Br. J. Sports Med.* **2019**, *53*, 1474–1478. [CrossRef]
16. Spindler, K.P.; Huston, L.J.; Chagin, K.M.; Kattan, M.W.; Reinke, E.K.; Amendola, A.; Andrish, J.T.; Brophy, R.H.; Cox, C.L.; Dunn, W.R. Ten-Year Outcomes and Risk Factors After Anterior Cruciate Ligament Reconstruction: A MOON Longitudinal Prospective Cohort Study. *Am. J. Sports Med.* **2018**, *46*, 815–825.

17. Hawker, G.A.; Badley, E.M.; Croxford, R.; Coyte, P.C.; Glazier, R.H.; Guan, J.; Harvey, B.J.; Williams, J.I.; Wright, J.G. A population-based nested case-control study of the costs of hip and knee replacement surgery. *Med. Care* **2009**, *47*, 732–741. [CrossRef]
18. Ageberg, E.; Forssblad, M.; Herbertsson, P.; Roos, E. Sex Differences in Patient-Reported Outcomes After Anterior Cruciate Ligament Reconstruction Data From the Swedish Knee Ligament Register. *Am. J. Sports Med.* **2010**, *38*, 1334–1342. [CrossRef]
19. Wasserstein, D.; Huston, L.J.; Nwosu, S.; Group, M.; Kaeding, C.C.; Parker, R.D.; Wright, R.W.; Andrish, J.T.; Marx, R.G.; Amendola, A.; et al. KOOS pain as a marker for significant knee pain two and six years after primary ACL reconstruction: A Multicenter Orthopaedic Outcomes Network (MOON) prospective longitudinal cohort study. *Osteoarthr. Cartil* **2015**, *23*, 1674–1684. [CrossRef]
20. Webster, K.E.; Hewett, T.E. Meta-Analysis of Meta-Analyses of Anterior Cruciate Ligament Injury Reduction Training Programs. *J. Orthop. Res.* **2018**, *36*, 2696–2708. [CrossRef]
21. Hamrin Senorski, E.; Svantesson, E.; Spindler, K.P.; Alentorn-Geli, E.; Sundemo, D.; Westin, O.; Karlsson, J.; Samuelsson, K. Ten-Year Risk Factors for Inferior Knee Injury and Osteoarthritis Outcome Score After Anterior Cruciate Ligament Reconstruction: A Study of 874 Patients From the Swedish National Knee Ligament Register. *Am. J. Sports Med.* **2018**, *46*, 2851–2858. [CrossRef] [PubMed]
22. Zbrojkiewicz, D.; Vertullo, C.; Grayson, J.E. Increasing rates of anterior cruciate ligament reconstruction in young Australians, 2000–2015. *Med. J. Aust.* **2018**, *208*, 354–358. [CrossRef] [PubMed]
23. Shaw, L.; Finch, C.F.; Bekker, S. Infographic: Trends in paediatric and adolescent ACL injuries. *Br. J. Sports Med.* **2019**, *53*, 228. [CrossRef] [PubMed]
24. Maffulli, N.; Longo, U.G.; Gougoulias, N.; Loppini, M.; Denaro, V. Long-term health outcomes of youth sports injuries. *Br. J. Sports Med.* **2010**, *44*, 21–25. [CrossRef]

Journal of
Clinical Medicine

MDPI

Article

Medical Interventions for Patellofemoral Pain and Patellofemoral Osteoarthritis: A Systematic Review

Erin M. Macri [1,2,*], Harvi F. Hart [3,4], David Thwaites [3], Christian J. Barton [3,5], Kay M. Crossley [3], Sita M.A. Bierma-Zeinstra [1] and Marienke van Middelkoop [1]

1 Department of General Practice, Erasmus University Medical Center, Doctor Molewaterplein 40, 3015 GD Rotterdam, The Netherlands; s.bierma-zeinstra@erasmusmc.nl (S.M.A.B.-Z.); m.vanmiddelkoop@erasmusmc.nl (M.v.M.)
2 Department of Family Practice, The University of British Columbia, 5950 University Boulevard, Vancouver, BC V6T 1Z3, Canada
3 Sport and Exercise Medicine Research Centre, La Trobe University, Health Sciences Building 3, Bundoora VIC 3086, Australia; h.hart@latrobe.edu.au (H.F.H.); david@completesportscare.com.au (D.T.); christian@completesportscare.com.au (C.J.B.); K.Crossley@latrobe.edu.au (K.M.C.)
4 Department of Physical Therapy, Western University, London CA N6G 2V4, UK
5 Department of Surgery, St Vincent's Hospital, University of Melbourne, Building 181, Grattan St, Melbourne VIC 3010, Australia
* Correspondence: e.macri@erasmusmc.nl; Tel.: +31-10-70-33459

Received: 16 September 2020; Accepted: 19 October 2020; Published: 23 October 2020

Abstract: Patellofemoral pain (PFP) and patellofemoral osteoarthritis (PFOA) are common, persistent conditions that may lie along a pathological spectrum. While evidence supports exercise-therapy as a core treatment for PFP and PFOA, primary care physicians commonly prescribe medication, or refer for surgical consults in persistent cases. We conducted a systematic review of medical interventions (pharmaceutical, nutraceutical, and surgical) for PFP and PFOA to inform primary care decision making. Methods: Following protocol registration, we searched seven databases for randomized clinical trials of our target interventions for PFP and PFOA. Our primary outcome was pain. We assessed risk of bias, calculated standardized mean differences (SMDs) and determined the level of evidence for each intervention. Results: We included 14 publications investigating pharmaceutical or nutraceutical interventions, and eight publications investigating surgical interventions. Two randomized control trials (RCTs) provided moderate evidence of patellofemoral arthroplasty having similar pain outcomes compared to total knee arthroplasty in isolated PFOA, with SMDs ranging from −0.3 (95% CI −0.8, 0.2, Western Ontario McMaster Pain Subscale, 1 year post-surgery) to 0.3 (−0.1, 0.7, SF-36 Bodily Pain, 2 years post-surgery). Remaining studies provided, at most, limited evidence. No efficacy was demonstrated for oral nonsteroidal anti-inflammatories or arthroscopic surgery. Conclusions: Pharmaceutical and nutraceutical prescriptions, and surgical referrals are currently being made with little supporting evidence, with some interventions showing limited efficacy. This should be considered within the broader context of evidence supporting exercise-therapy as a core treatment for PFP and PFOA.

Keywords: patellofemoral pain; patellofemoral osteoarthritis; pharmaceuticals; nutraceuticals; surgery

1. Introduction

Patellofemoral pain (PFP) is a condition characterized by diffuse peri- or retro-patellar pain that is made worse by activities that increase patellofemoral joint load such as squatting, negotiating stairs, or running [1]. PFP is associated with reduced physical activity [2,3], impaired quality of life [4,5], and psychosocial features such as anxiety and depression [3,6]. PFP may also present with structural

findings on imaging; particularly in older adults with patellofemoral osteoarthritis (PFOA) [7–9]. Annual prevalence of PFP in the general population is estimated at almost 23% [10,11], and it is the most common knee complaint seen by general practitioners [12]. At least 1.5% of over 30 million insured individuals who seek outpatient orthopedic care in the United States do so for PFP [13].

PFP is frequently persistent. In prospective studies of individuals diagnosed with PFP, 40–77% report persistent pain 6 to 20 years later [14,15]. There is an emerging hypothesis that PFOA may represent a long-term sequela of persistent PFP [7,8,16–18]. Like PFP, PFOA is associated with pain [9,19], reduced physical function [19], and lower quality of life [20], and PFOA in turn may lead to whole knee osteoarthritis (OA) [21]. These long-term consequences implore clinicians and researchers to ensure that the best evidence informs patient management, and highlights an urgent need for research aimed at improving patient outcomes.

Primary care clinicians prescribe pharmaceuticals to 25% or more of patients presenting with knee complaints such as PFP and PFOA, with underlying inflammation or structural features presumed to cause symptoms [12,22]. These prescriptions may be made in combination with exercise advice or referrals to physical therapy [12], which are recommended for evidence-based management of PFP and PFOA [23–27]. In addition to pharmaceutical agents, nutraceuticals are reported to be of therapeutic benefit in certain musculoskeletal conditions [28]. They are derived from dietary sources, can be administered in different formats (e.g., taken orally or by injection), and are regulated differently in different countries. Therefore, nutraceuticals can be prescribed and administered in clinical settings, but in many countries patients can also obtain nutraceuticals off-the-shelf and taken without specific recommendations by a medical practitioner. Finally, in cases where patients have not responded favorably to non-surgical care, clinicians may also refer them for surgical consults (10 to 12%) [12,22,29,30].

While the best evidence for exercise-therapy and adjunct physical and rehabilitation medicine modalities have been recently synthesized [31,32], a systematic review of randomized control trials (RCTs) has not been completed for pharmaceutical or nutraceutical efficacy in over 10 years [33], and since then, new treatments have emerged (e.g., platelet rich plasma). Moreover, to our knowledge, a systematic synthesis of surgical treatment efficacy has not yet been undertaken [33]. Thus, we conducted a systematic review to answer the question of whether medical interventions (pharmaceutical, nutraceutical, or surgical) are efficacious in the management of PFP and PFOA, to inform primary care evidence-based decision making and identify knowledge gaps.

2. Methods

2.1. Study Design

We designed our review using the Preferred Reporting Items for Systematic Reviews and meta-analyses (PRISMA) guidelines [34]. We prospectively registered our study protocol as two separate protocols with the International Prospective Register of Systematic Reviews (PROSPERO) [35] (#CRD42017082527, #CRD42017078575). However, on account of the similarity of the research, study design and methods, and likelihood that a single paper would be most useful to clinicians, we subsequently combined them into a single study. To combine the two protocols, we adopted the methods from each protocol that provided the most rigorous and inclusive study design. We used the former protocol to select databases to search, resulting in five additional databases. We limited types of study designs to RCT only, in order to align with current best practices [26]. We included all types of control interventions, including wait-list or placebo. We defined the primary outcome as pain, though all outcomes from both protocols were still included. Finally, we removed the exclusion criterion of traumatic knee injury from the former protocol since this was no longer appropriate in regard to surgical studies.

2.2. Study Eligibility and Search Strategy

Our search strategy was developed by two co-authors (EMM, MvM) in collaboration with a reference librarian (Table 1). We included peer-reviewed RCTs investigating pharmaceutical, nutraceutical, or surgical interventions for PFP or PFOA. We included studies investigating PFP using intentionally broad criteria based on current recommended terminology [1]. Specifically, for PFP, we included individuals diagnosed either symptomatically or structurally, and included synonyms for PFP such as patellofemoral pain syndrome and chondromalacia patella, recognizing that terminology differs among different disciplines and has changed over time [1,36]. We also included individuals with PFOA because emerging evidence suggests that many conditions affecting the patellofemoral joint may lie along a pathological spectrum, with PFOA potentially representing a late stage of chronic PFP (though treatment approaches may differ) [7,8,16–18]. We included men and women of all ages. Diagnosis by any common clinical or structural criteria was eligible. We excluded studies targeting patellar instability. We included all types of control interventions (placebo, wait-list, other interventions). Our primary outcome was pain, and secondary outcomes included function, quality of life, structural outcomes on imaging, and biomechanics. We included studies reported in English, Dutch, French, or German.

Table 1. Search string for Medline (Ovid).

(Patellofemoral Pain Syndrome/OR Chondromalacia Patellae/OR ((Patellofemoral Joint/OR patella/) AND (osteoarthritis/OR osteoarthritis, knee/OR pain/OR Joint Diseases/OR Cartilage Diseases/)) OR Chondromalacia Patellae/OR (((femoropatell* OR patell* OR retropatell*) ADJ6 (pain* OR osteoarthrit* OR osteo-arthrit* OR arthralgi* OR syndrome* OR Chondromalac* OR dysfunction* OR chondropath*)) OR ((femoropatell* OR patell* OR retropatell*) ADJ6 (degenerat*) ADJ6 (arthrit* OR cartilag*)) OR (anter* ADJ3 knee ADJ3 pain*) OR ((lateral* OR odd) ADJ3 (compress* OR facet OR pressure*) ADJ3 syndrome*)).ab,ti,kw) AND (Exp Controlled clinical trial/OR "Double-Blind Method"/OR "Single-Blind Method"/OR "Random Allocation"/OR (random* OR factorial* OR crossover* OR cross over* OR placebo* OR ((doubl* OR singl*) ADJ blind*) OR assign* OR allocat* OR volunteer* OR trial OR groups).ab,ti.) NOT (Animals/NOT Humans/)

NB Syntax was developed in MEDLINE using keywords and medical subject headings, then adapted for PEDro, EMBASE, CINAHL, CENTRAL, SPORTDiscus, and Scopus using their respective indexing vocabularies.

Two investigators (EMM, HFH) independently screened titles and abstracts from seven databases: MEDLINE, PEDro, EMBASE, CINAHL, CENTRAL, SPORTDiscus, and Scopus. The first search included all publication dates from database inception to 12 December 2018, and a search update included publications to 4 May 2020. Relevant reference lists, including those of related review papers, were also screened. The same investigators independently screened all potentially eligible full-text publications. At both stages of screening, in cases of disagreement, a third investigator was consulted (CJB).

2.3. Risk of Bias

Two investigators (EMM, DT) independently evaluated the risk of bias (RoB) using Cochrane's RoB v2.0 [37], and discrepancies were resolved with a third investigator (CJB). This tool reports on five domains of potential bias: randomization; deviations from intended interventions; missing outcome data; measurement of the outcome; and selection of the reported results. Each domain is summarized as having low or high bias, or some concerns. We contacted study authors in instances of missing information. Per Cochrane guidelines, a study was determined to be at overall high RoB if at least one domain was rated as having high RoB, or some concerns were identified across multiple domains. A study was considered at overall low RoB if all five domains had low RoB.

2.4. Data Extraction and Statistical Analyses

One author (EMM) extracted data including study population, participant characteristics, interventions, follow-up time points, and relevant outcomes. Data extraction was verified by a second author (CJB or DT). Where results were provided as legible figures, we digitized them and extracted

outcomes (WebPlotDigitizer v.4.2, San Francisco, CA, USA). We contacted study authors where information was missing. We grouped follow-up times as: short-term (0 to 6 weeks), medium-term (7 weeks to 6 months), and long-term (greater than 6 months) [38]. Where adequate data were reported (or obtained from authors), we reported effect sizes as Hedge's standardized mean differences (SMD) with 95% confidence intervals (CI) using Stata SE 15.1 (StataCorp LLC, College Station, TX, USA). SMDs enhance comparison across studies, particularly when different outcome measures are used to measure similar domains, by reporting effect sizes in units of pooled standard deviation. We conducted meta-analyses for any intervention with at least three trials of sufficient homogeneity and undertook best-evidence synthesis when this was not possible. We summarized the overall level of certainty of the evidence for each intervention using the Grading of Recommendation, Assessment, Development and Evaluation (GRADE) approach [39]. GRADE overall ratings range from a "very low" to "high" level of certainty, and are derived from consideration of five domains: RoB, consistency of effects, indirectness, imprecision, and publication bias. Using the GRADE approach represents a deviation from our original protocol that we undertook in order to better align with current best practices.

3. Results

We screened 6375 titles (Figure 1) and identified 22 eligible papers (20 studies). Meta-analysis was not possible due to study heterogeneity.

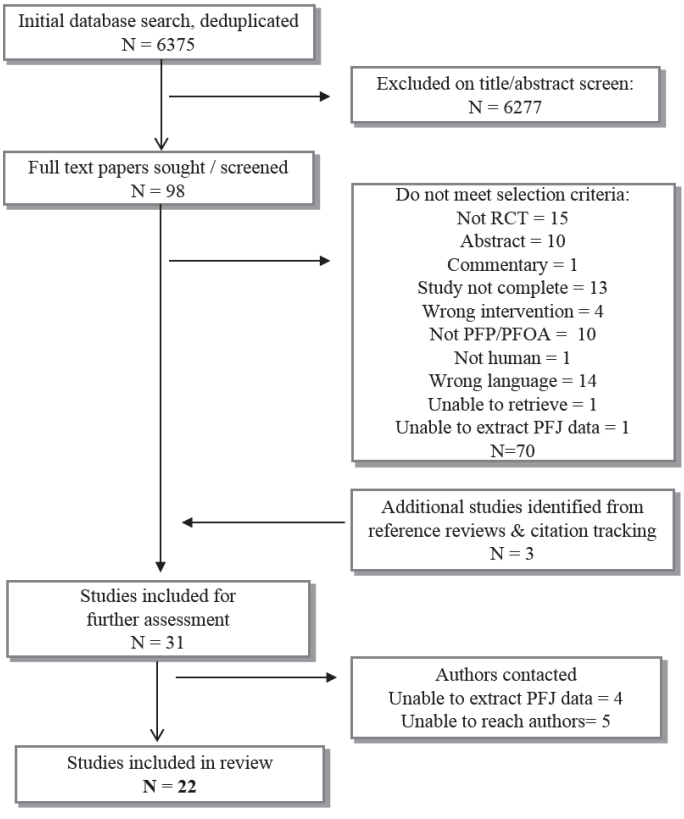

Figure 1. Flow chart of study selection.

We rated fourteen studies as high RoB, four moderate, and two low (Figure 2). Only four studies registered trial protocols [40–44]. Eight studies did not report funding source [45–52], and eight were at least partly industry-funded [40,44,52–57].

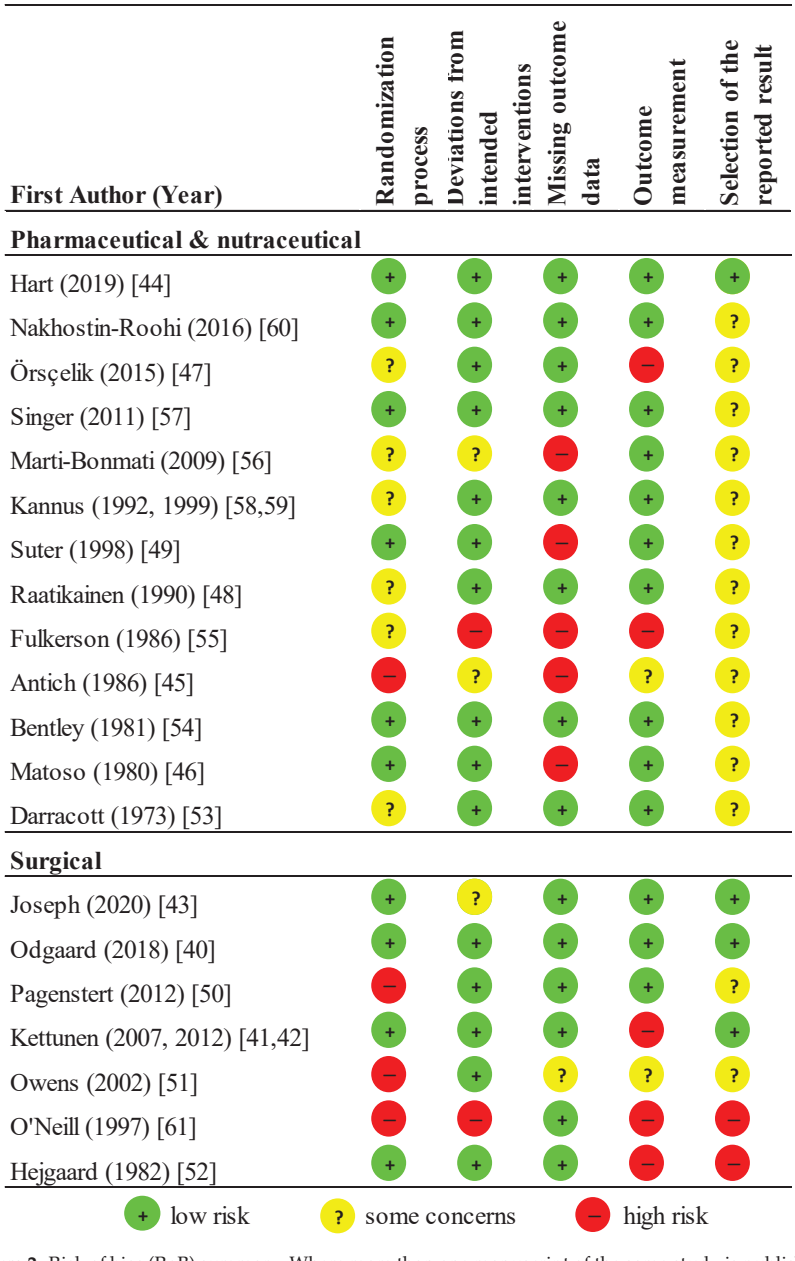

Figure 2. Risk of bias (RoB) summary. Where more than one manuscript of the same study is published, they are reported together.

Table 2. Participant characteristics.

First Author Year	Sample Randomized *n*	Sample Completed *n*		Women % (*n*)		Age Mean (SD) y		BMI Mean (SD) kg/m^2		Height Mean (SD) cm		Weight Mean (SD) kg	
Pharmaceutical and Nutraceutical		**Int**	**Cont**	**Int**	**Cont**	**Int**	**Cont**	**Int**	**Cont**	**Int**	**Cont**	**Int**	**Cont**
Hart 2019 [44]	86	42	38	76% (34)	76% (31)	26.0 (7.0)	28.1 (8.4)	26.4 (5.3)	25.8 (5.1)	168.9 (9.2)	168.3 (8.6)	75.7 (16.6)	73.5 (17.6)
Nakhostin-Roohi 2016 [60]	93	31	31 * 31 *	100% (93)		26.0 (8.9)	26.2 (10.0) 27.2 (9.5)	NR	NR NR	160.7 (11.1)	164.2 (7.2) 1.6 (8.9)	59.3 (9.5)	58.2 (11.1) 61.0 (12.8)
Örsçelik 2015 [47]	30	10	20	40% (4)	35% (7)	28.7 (6.0)	27.2 (5.7)	24.2 (3.4)	24.0 (3.0)	170.8 (6.6)	174.5 (7.9)	71.1 (12.5)	73.4 (12.8)
Singer 2011 [57]	24	14	10	57% (8)	90% (9)	31.5 (r15,48)	27.4 (r18,44)	NR	NR	NR	NR	NR	NR
Marti-Bonmati 2009 [56]	20	10	6	63% (10)		39 (18)		NR	NR	NR	NR	NR	NR
Kannus 1992 [58]	5	16	16 † 17 †	53% (28)		27 (9)		NR	NR NR	NR	NR NR	NR	NR NR
and 1999 [59]	As above	45		aa		aa		aa		aa		aa	
Suter 1998 [49]	42	19	17	31% (13)		35.6 (8.4)		NR	NR	NR	NR	NR	NR
Raitikainen 1990 [48]	31	14	15	29% (4)	20% (3)	29.1 (7.7)	30.2 (6.6)	24.6 (3.3)	25.1 (3.2)	173.6 (9.0)	177.0 (8.8)	74.8 (14.0)	78.9 (13.9)
Fulkerson 1986 [55]	56	20	16	78% (28)		32.3 (??)		NR	NR	NR	NR	NR	NR
Antich 1986 [45]	64 (86 k)	9 k	21 k ‡ 13 k ‡ 16 k ‡	NR		NR		NR	NR	NR	NR	NR	NR
Bentley 1981 [54]	30	16	13	72% (21)		25 (??)		NR	NR	NR	NR	NR	NR
Matoso 1980 [46]	33	12	10	53% (10)	64% (9)	28 (r16,50)	30 (r16,46)	NR	NR	NR	NR	NR	NR
Darracott 1973 [53]	43	23	20	0% (0)	15% (3)	22.9 (4.4)	20.8 (3.3)	NR	NR	NR	NR	NR	NR
Surgical													
Joseph 2020 [43]	64	31	29	71% (22)	90% (26)	64.7 (10.5)	64.4 (12.8)	28.9 (6.7)	29.2 (4.2)	NR	NR	NR	NR
Odgaard 2018 [40]	100	46	47	77% (77)		64 (8.9)		NR		NR		NR	

Table 2. *Cont.*

First Author Year	Sample Randomized *n*	Sample Completed *n*		Women % (*n*)		Age Mean (SD) y		BMI Mean (SD) kg/m^2		Height Mean (SD) cm		Weight Mean (SD) kg	
Pagenstert 2012 [50]	(28)	14	14	79% (11)	71% (10)	47.6 (9.9)	48.0 (11.6)	NR	NR	NR	NR	NR	NR
Kettunen 2007 [41]	56	27	25	61% (17)	64% (18)	28.4 (7.5)	28.4 (5.6)	24.1 (3.3)	23.8 (3.6)	171.7 (10.2)	172.4 (9.6)	69.0 (19.3)	71.4 (15.1)
2012 [42]	As above	24	20	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa
Owens 2002 [51]	48	20	19	100% (20)	100% (19)	36.9 (r 30,45)	37.5 (r 30,45)	NR	NR	NR	NR	NR	NR
O'Neill 1997 [61]	91	43	43	60% (26)	58% (25)	27.2 M 18 (r13,56) W 33 (r15,47)	28.7 M 22 (r13,33) W 34 (r14,59)	NR	NR	NR	NR	NR	NR
Hejgaard 1982 [52]	42	20	22	50% (10)	59% (13)	28 (r18,38)	32 (r19,50)	NR	NR	NR	NR	NR	NR

BMI = body mass index; k = knees; NR = not reported; Int = intervention; Cont = control; aa = as above; iPFOA = isolated patellofemoral osteoarthritis. * cont top = piroxicam, cont bottom = base gel. † cont top = exercise alone, cont bottom = placebo injection. ‡ cont top = iontophoresis; cont mid = US/ice; cont bottom = ice.

Table 3. Standardized mean differences (SMD) and relative risks (RR) of all included studies (SMD > 0 and RR > 1 indicate improvement of intervention group relative to control).

First Author Year	Intervention	Control	Outcome	Short Term * <6 Weeks	Medium Term 7 Weeks–6 Months	Long Term, >6 Months	Adverse Events
Pharmaceutical and Nutraceutical				**SMD (95% CI)**	**SMD (95% CI)**	**SMD (95% CI)**	
Hart 2019 [44]	Hyaluronic acid IA inj	Placebo	VAS single leg squat	−0.3 (−0.7, 0.2)	−0.3 (−0.7, 0.2)		NR
			KOOS	**−0.9 (−1.3, −0.4)**	**−0.6 (−1.0,−0.2)**		
			AKPS	**−0.6 (−1.1, −0.2)**	**−0.5 (−1.0,−0.1)**		
			Tegner	−0.3 (−0.7, 0.1)	−0.2 (−0.6, 0.3)		
			Knee extension	0.0 (−0.4, 0.4)	−0.1 (−0.6, 0.3)		
			Knee ratio	0.0 (−0.4, 0.4)	−0.3 (−0.7, 0.2)		
Nakhostin-Roohi 2016 [60]	Olive oil (phonophoresis)	Piroxicam (phonophoresis)	WOMAC Pain	NR			NR
			WOMAC Stiffness	NR			
			WOMAC Physical Function	NR			
Örsçelik 2015 [47]	PRP, 3 IA inj	PRP, 1 inj	AKPS		**1.1 (0.3, 2.0)**		Some localized pain for several weeks after injections (numbers/arm not reported)
			Quadriceps peak torque		0.1 (−0.7, 0.8)		
Singer 2011 [57]	Botox IM inj	Placebo	AKPS		**0.9 (0.1, 1.8)**		6 (43%) slight distal thigh asymmetry; few (numbers and arm NR): temporary bruising or pain last up to several days.
			Pain stairs		0.3 (−0.5, 1.1)		
			Pain squat		**1.1 (0.2, 1.9)**		
			Pain kneel		**1.4 (0.5, 2.3)**		
			Pain walk		−0.9 (−1.7, 0.0)		
Marti-Bonmati 2009 [56]	Glucosamine	Acetaminophen	AKSS		**2.5 (1.1, 3.9)**		NR
			VAS (task not specified)		**4.1 (2.2, 6.0)**		
Kannus 1992 and 1999 [58,59]	GAGPS IA inj	Placebo	No pain, 1 leg jump	RR 0.9 (0.4, 1.8)	RR 0.8 (0.5, 1.3)	NR	0 intervention vs. 1 control reactive synovitis to injection (discontinued treatment)
			No pain, 25 squats	RR 1.1 (0.7, 1.6)	RR 1.1 (0.7, 1.6)		
			Excellent rating	RR 0.9 (0.5, 1.8)	RR 1.4 (0.8, 2.4)		
			Return full activity	RR 1.4 (0.8, 2.4)	RR 1.4 (0.9, 2.0)		
			Lysholm	0.2 (−0.5,0.8)	0.2 (−0.5, 0.9)		
			Tegner	0.4 (−0.3, 1.1)	−0.1 (−0.6, 0.8)		
Suter 1998 [49]	Naproxen	Placebo	Pain with activity	0.2 (−0.9, 0.5)	0.0 (−0.7, 0.7)		NR
			Pain resisted extension	0.5 (−0.2, 1.1)			
Raatikainen 1990 [48]	GAGPS IM inj	Placebo	Pain down stairs	−0.1 (−0.6, 0.9)	0.3 (−0.4, 1.1)	**1.6 (0.8, 2.5)**	1 intervention, sweating and dizziness after 4th injection, resolved and continued
			Pain squat	0.3 (−0.4, 1.1)	0.3 (−0.4, 1.1)	**1.3 (0.5, 2.2)**	
			Physician "Improved"	RR 1.4 (0.6, 3.4)	RR 1.7 (0.5, 6.1)	**RR 3.6 (1.2, 10.4)**	
			Hindrance—sport	0.4 (−0.3, 1.2)	0.5 (−0.3, 1.2)	**0.8 (0.1, 1.6)**	
			Hindrance—life	−0.2 (−0.9, 0.6)	0.3 (−0.4, 1.1)	0.7 (−0.0, 1.5)	

Table 3. *Cont.*

First Author Year	Intervention	Control	Outcome	Short Term * <6 Weeks	Medium Term 7 Weeks–6 Months	Long Term, >6 Months	Adverse Events
Fulkerson 1986 [55]	Diflunisal	Naproxen	Significant pain relief	RR 0.9 (0.5, 1.5)			11 (55%) diflunisal (headache, gastric distress, etc.) vs. 7 (44%) naproxen (drowsiness, headache, etc.)
Antich 1986 [45]	Hexadrol (phonophoresis)	Ultrasound plus ice	Subjective improvement	[RR 0.7 (0.2, 2.2)] †			NR
Bentley 1981 [54]	Palaprin	Placebo	Global improvement	RR 1.2 (0.4, 3.4)			NR
Matoso 1980 [46]	Chloroquine	Placebo	Spontaneous pain		NR		15 (79%) intervention (6 discontinued treatment) vs. 4 (29%) controls: visual trouble, digestive trouble, headache, vertigo
Darracott 1973 [53]	Nandrolone phenylpropionate inj	Placebo	Improved	**RR 8.7 (2.3, 32.7)**			**NR**
			Surgical				
Joseph 2020 [43]	Patellofemoral arthroplasty	Total knee arthroplasty	WOMAC (1y): Pain Stiffness Function AKSS (1y): Knee Function UCLA (1y) EQ5D3L (5y) OKS (5y) Pain-free years (5y) Satisfied (5y)			 −0.3 (−0.8, 0.2) −0.2 (−0.7, 0.3) −0.1 (−0.6, 0.4) 0.1 (−0.5, 0.6) −0.2 (−0.4, 0.7) 0.3 (−0.2, 0.8) 0.2 (−0.4, 0.8) −0.1 (−0.8, 0.5) −0.6 (−1.2, 0.1) RR 0.9 (0.6, 1.2)	Superficial infections: 4 PFA, 5 TKA—all treated with antibiotics. 4 TKA required further interventions: 1 arthroscopic facetectomy, 1 manipulation under anesthesia, 2 aspiration/steroid injection
Odgaard 2018 [40]	Patellofemoral arthroplasty	Total knee arthroplasty	KOOS: Pain Symptoms ADL SPR QOL SF-36 Bodily Pain	0.4 (−0.0, 0.8) ‡ **0.6 (0.2, 1.0)** 0.2 (−0.2, 0.6) 0.3 (−0.1, 0.7) 0.0 (−0.4, 0.4) **0.5 (0.1, 0.9)**	0.4 (−0.0, 0.8) **0.8 (0.4, 1.2)** 0.1 (−0.3, 0.5) 0.3 (−0.1, 0.7) 0.3 (−0.1, 0.7) **0.4 (0.0, 0.8)**	0.1 (−0.3, 0.5) **0.5 (0.1, 0.9)** −0.1 (−0.6, 0.3) 0.3 (−0.2, 0.7) 0.1 (−0.4, 0.5) 0.3 (−0.1, 0.7)	Intervention: 2 deaths unrelated to surgery, 2 revisions (1 of trochlear component, 1 to TKA), 2 other surgical procedures; control: 5 other surgical procedures
Pagenstert 2012 [50]	Open lateral retinacular lengthening	Open lateral retinacular release	AKPS Quadriceps atrophy		0.5 (−0.2, 1.3) **1.2 (0.4, 2.0)**	**0.8 (0.1, 1.6)** **1.6 (0.7, 2.5)**	No surgical complications. At 2 years, recurrence of symptoms in 1 lengthening and 2 release; and over-release in 5 release arm
Kettunen 2007 and 2012 [41,42]	Arthroscopy plus exercise	Exercise alone	AKPS Pain down stairs Pain up stairs Pain sit to stand			−0.0 (−0.6, 0.5) −0.1 (−0.7, 0.4) −0.0 (−0.5, 0.5) 0.2 (−0.3, 0.7)	Intervention: 1 delayed exercise due to pain, 1 refused exercise after surgery; 3 controls had arthroscopic surgery

Table 3. *Cont.*

First Author Year	Intervention	Control	Outcome	Short Term * <6 Weeks	Medium Term 7 Weeks–6 Months	Long Term, >6 Months	Adverse Events
Owens 2002 [51]	Debridement, radiofrequency probe	Debridement, standard shaver	Fulkerson-Shea			**1.7 (0.9, 2.4)**	NR
O'Neill 1997 [61]	Open lateral retinacular lengthening	Arthroscopic lateral retinacular release	Tegner Quadriceps open chain strength deficit Quadriceps atrophy Knee score (modified Lysholm)			RR 1.1 (1.0, 1.2) RR 1.1 (0.8, 1.6) RR 1.2 (0.9, 1.5) **NR**	Intervention: 1 stitch abscess; 1 skin incision keloid; 1 unable to achieve 90° knee flexion; control: 1 iliotibial band contracture, 1 hematoma, 1 superficial infection, 1 hematoma with infection after self-draining
Hejgaard 1982 [52]	Anterior displacement of tibial tuberosity plus debridement	Debridement	Surgeon, Good to Excellent			**RR 2.6 (1.2, 5.4)**	Intervention: 2 (10%) effusion, 2 (10%) thromboembolism; control: 5 effusion (23%)

Interventions for pharmaceutical/nutraceutical are taken orally unless otherwise stated. Outcomes reported are for pain, function, and global changes only—other outcomes (e.g., balance, biomechanics) were reported so rarely that comparisons are not possible. * Where multiple evaluations occur within a given follow-up time category, we report only the longest follow-up within that time period. † Approximate estimate only—numbers reported are knee-level comparisons, not person-level, with no indication of how many participants had both knees assessed. ‡ Values were extracted from figures; however, we assumed that the error bars represent 95% CI and NOT standard error of the mean as is reported in the publication figures—this is more congruent with the publication's reported results. SMD = standardized mean difference; RR = relative risk; NR = not reported or inadequate detail to extract data; Int = intervention arm; GAGPS = glycosaminoglycan polysulphate; PRP = platelet rich plasma; inj = injection; IA = intra-articular injection; IM = intra-muscular injection; SF-36 = Medical Outcomes Score short form 36; KOOS = Knee injury and Osteoarthritis Outcome Score; ADL= activities of daily living; SPR = sports and recreation; QOL = quality of life; AKPS = anterior knee pain scale; WOMAC = Western Ontario McMaster Osteoarthritis Index; AKSS = American Knee Society Score; UCLA = University of California Los Angeles Physical Activity Questionnaire; EQ5D3L = EuroQol five-dimension, three-level questionnaire; OKS: Oxford Knee Score. Bold indicates statistical significance ($p < 0.05$).

3.1. Pharmaceutical or Nutraceutical Studies

In total, 14 publications (13 studies, 534 participants analyzed) investigated pharmaceutical or nutraceutical interventions [44–49,53–60]. One study enrolled participants with patellofemoral degeneration including OA [56], and the remaining studies mostly enrolled participants with PFP (Table S1). Thirteen publications reported sex and age, approximately two/thirds of participants were women (n = 315), and the mean age was 28 years (Table 2). Only four studies reported body mass characteristics [44,47,48,60]. Three classes of interventions were rated as low certainty of evidence using GRADE, and seven were rated as very low certainty of evidence (Table S2).

3.2. Oral Administration

3.2.1. Nonsteroidal antiinflammatories, NSAIDs—Low Certainty of Evidence

One moderate [54] and two high RoB RCTs [49,55] investigated NSAIDs. While pain or global change improved with most interventions, aspirin (short-term) and naproxen (medium-term) were no different than placebo, and naproxen (short-term) was no different than diflunisal (Table 3). One other high RoB RCT found no within-arm change in pain (medium-term) using acetaminophen in a sample with patellofemoral degeneration [56].

3.2.2. Chloroquine—Very Low Certainty of Evidence

One high RoB RCT of chloroquine compared to placebo reported medium-term improvement in "spontaneous pain" [46], a combined measure of pain at rest, during walking and when first waking. There was inadequate detail to report SMDs.

3.2.3. Glucosamine—Very Low Certainty of Evidence

One high RoB RCT compared oral glucosamine to NSAID (acetaminophen) in patellofemoral degeneration [56]. Medium-term results favored glucosamine in pain (SMD 2.5 [95% CI 1.1, 3.9]) and function outcomes (American Knee Society score, SMD 4.1 [2.2, 6.0]).

3.3. Injections

3.3.1. Glycosaminoglycan Polysulphate (GAGPS)—Very Low Certainty of Evidence

Two high RoB RCTs (three publications) investigated GAGPS injections [48,58,59]. Five intra-articular injections were no better than placebo for pain and function for up to five years after treatment [58,59]. Twelve intra-muscular injections were no better than placebo at short- and medium-term for pain, function, and physician-determined overall effect. However, long-term results (one year) favored GAGPS injections for pain descending stairs (SMD 1.6 [0.8, 2.5]), pain squatting (SMD 1.3 [0.5, 2.2]), hindrance in sport (SMD 0.8 [0.1, 1.6]), and physician's evaluation of being "improved" (relative risk, RR 3.6 [1.2, 10.4]).

3.3.2. Hyaluronic Acid—Low Certainty of Evidence

One low RoB RCT compared a single intra-articular injection of hyaluronic acid to a sham injection [44]. Investigators reported no clinically meaningful between-group differences for up to six months of follow-up. Our SMD calculations, based on complete case analysis, showed no differences in pain, Tegner scores, or strength, but the group receiving hyaluronic acid had worse function (Knee injury and Osteoarthritis Outcome Score [KOOS], SMD −0.6 [−1.0,−0.2]; Anterior Knee Pain Scale [AKPS], SMD-0.5 [−1.0, −0.1].

3.3.3. Platelet Rich Plasma (PRP)—Very Low Certainty of Evidence

One high RoB RCT compared three intra-articular PRP injections to one injection [47]. Both arms demonstrated medium-term within-group improvements in self-reported pain/function, balance, coordination, and endurance. However, only pain/function showed a between-group difference favoring three injections over one (AKPS SMD 1.1 [0.3, 2.0]).

3.3.4. Botox—Low Certainty of Evidence

One moderate RoB RCT evaluated Botox injected into vastus lateralis compared to placebo [57]. Medium-term results favored Botox for self-reported pain/function (AKPS SMD 0.9 [0.1, 1.8]), and pain in squatting (SMD 1.1 [0.2, 1.9]) and kneeling (SMD 1.4 [0.5, 2.3]), but no difference with pain in stairs or walking.

3.3.5. Anabolic Steroids—Very Low Certainty of Evidence

One high RoB RCT compared nandrolone phenylpropionate injection to placebo [53]. Short-term results favored the treatment arm in terms of the number of participants becoming moderately improved or resolved (RR 8.7 [2.3, 32.7]).

3.4. Transdermal

3.4.1. NSAIDs (Phonophoresis)—Very Low Certainty of Evidence

One moderate RoB RCT compared olive oil (phonophoresis) to NSAID (phonophoresis) and placebo (ultrasound alone) [60]. At one week, within-arm improvements were seen for olive oil (pain and function) and piroxicam (function). By two weeks, all three arms improved (pain and function), with olive oil improving more than placebo (function) but not more than piroxicam (values not reported), and piroxicam no different than placebo.

3.4.2. Corticosteroid (Iontophoresis or Phonophoresis)—Very Low Certainty of Evidence

One high RoB RCT compared Hexadrol with topical Xylocaine (phonophoresis), Hexadrol with topical Xylocaine (iontophoresis), ice, and ultrasound with ice [45]. Short term, more participants reported subjective improvement using ultrasound with ice than either corticosteroid arm, though this does not appear to be significant (inadequate reporting).

3.5. Surgical Studies

Eight publications (seven studies, 411 participants analyzed) investigated surgical interventions [40–43,50–52,61]. The source population was more heterogeneous compared to pharmaceutical and nutraceutical studies (Table 2, Table S1): two studies enrolled participants with isolated PFOA [40,43], one with isolated patellar cartilage lesions [51], and the remaining mostly PFP. Approximately 72% of participants were women (n = 294), with a mean age of 41 years (mean age range 28 to 64). Only two studies reported body mass characteristics [41–43]. Four classes of interventions were rated as very low certainty of evidence using GRADE, and one was rated as moderate certainty of evidence (Table S2).

3.6. Arthroscopy

3.6.1. Standard Arthroscopy—Very Low Certainty of Evidence

One high RoB RCT (two publications) compared arthroscopy plus exercise to exercise alone [41,42]. Specific procedures were as indicated at surgery, (e.g., plica resection, abrasion of chondral lesions, shaving of synovium). At medium- and long-term (two and five year) follow-ups, surgery was no better than exercise alone in pain or function (Table 3).

3.6.2. Radiofrequency Debridement—Very Low Certainty of Evidence

One high RoB RCT compared arthroscopy with a bipolar radiofrequency probe compared to arthroscopy with a standard mechanical shaver for treating isolated patellar cartilage lesions [51]. Long-term results favored the radiofrequency probe in terms of self-reported pain/function (SMD 1.7 [0.9, 2.4]).

3.7. Open Surgeries

3.7.1. Lateral Retinacular Surgery—Very Low Certainty of Evidence

One high RoB RCT compared open lateral retinacular lengthening to open lateral retinacular release [50]. There was no difference three months after surgery, but at six months, results favored lengthening in terms of less quadriceps atrophy (SMD 1.2 [0.4, 2.0]). Long-term results favored lengthening for quadriceps atrophy (SMD 1.6 [0.7, 2.5]) and pain/function (AKPS SMD 0.8 [0.1, 1.6]).

One high RoB RCT compared open lateral retinacular lengthening to arthroscopic lateral retinacular release [61]. Long-term results (up to six years) favored open lengthening for pain/function (modified Lysholm, inadequate detail to report values) but no difference for return to sport level or strength.

3.7.2. Anterior Tibial Tuberosity Displacement—Very Low Certainty of Evidence

One high RoB RCT compared anterior tibial tuberosity displacement plus debridement to debridement alone [52]. Long-term results favored anterior tibial displacement in terms of the surgeon-determined proportion rated as good or excellent (RR 2.6 [1.2, 5.4]).

3.7.3. Patellofemoral Arthroplasty—Moderate Certainty of Evidence

Two RCTs (one low and one moderate RoB) compared patellofemoral arthroplasty to total knee arthroplasty in isolated PFOA [40,43]. In one study [40], short- and medium-term results favored patellofemoral arthroplasty for two pain subscales (KOOS Symptoms, SMD 0.6 [0.2, 1.0]; and SF-36 Bodily Pain, SMD 0.5 [0.1, 0.9]), but results did not differ for remaining KOOS and SF-36 subscales. In the same study, long-term results (2 years post-surgical) only remained in favor of patellofemoral arthroplasty for KOOS Symptoms (SMD 0.5 [0.1, 0.9]). The other study reported no differences in any outcome at 1 year after surgery, including several outcomes measured five years after surgery [43].

4. Discussion

Medical interventions are common for individuals with PFP or PFOA [12,22,27,29,30]. Given the high prevalence and potential for chronicity in these patellofemoral conditions, it is clinically imperative to make evidence-informed decisions when managing these patients. Our results suggest that most pharmaceutical, nutraceutical, or surgical interventions for PFP and PFOA have not undergone sufficiently rigorous investigation to warrant their use. Rather than placebo, usual care, or a control of known efficacy, many interventions were compared to other interventions that have not yet themselves undergone rigorous evaluation. Such studies add little to the overall evidence. Most studies included in this review were of high RoB, and results should thus be interpreted with caution. Replication studies with lower RoB would likely reduce the number of positive results in this review and shift larger effect sizes towards the null [62]. Studies with negative findings reported in this review, however, are less likely to change meaningfully in the presence of less bias [62].

There is almost no robust evidence regarding pharmaceutical or nutraceutical interventions for PFP or PFOA. Highlighting interventions more commonly prescribed in general practice: short-term, oral NSAIDs were no better than placebo (low certainty of evidence); medium-term, large improvements were reported with oral glucosamine compared to NSAID in a sample with patellofemoral degeneration including OA (very low certainty of evidence); and long-term, large improvements were reported with intra-muscular GAGPS injections compared to placebo (very low certainty of evidence). Regarding

GAGPS, the difference in results between the two studies could be due to the different delivery site (intra-articular vs. intra-muscular), dose-response (five injections in the study with negative findings vs. 12 in the study favoring GAGPS), or other factors such as sample selection (different proportion of women, for example). Having said this, long-term results were only measured in one of the two studies, and it is unclear how to explain why results would differ at one year but not at earlier time points.

There is very low to moderate certainty of evidence regarding surgical interventions. Highlighting more common surgical interventions: standard arthroscopy plus exercise was no better than exercise alone for PFP. The only intervention in the present study with moderate certainty of evidence was for patellofemoral arthroplasty. This evidence suggests possible early benefits in some symptoms compared to total knee arthroplasty for treating isolated PFOA, but likely no differences between interventions one to five years following surgery, suggesting that both approaches may be equally effective during this period of follow-up. One of these studies was rated as low RoB, and the other study introduced some bias through its pragmatic design, with surgeons deviating from protocol in several cases based on clinical judgement. Rates of follow-up procedures were similar between groups in one study, and higher in the total knee group in the other study. Additional benefits of patellofemoral arthroplasty may include it being a less invasive procedure with quicker post-operative recovery and preservation of bone stock [63]. While longer follow-ups are still required, as are studies comparing patellofemoral arthroplasty to non-surgical interventions, patellofemoral arthroplasty appears to be as effective as total knee arthroplasty for treating isolated PFOA, and may thus be a valid surgical option [29,30,64].

Pharmaceutical, nutraceutical, and surgical interventions for the treatment of PFP and PFOA are generally administered on the premise that they address underlying inflammation or structural abnormalities that cause the pain [12,22]. However, the source of pain is highly debated [1], and much like in knee OA [65], pain and structure are not strongly correlated [66,67], and pain is highly variable and subjective. This may explain why some interventions included in this review offer no benefit over placebo. It is possible that a subgroup of individuals with PFP or PFOA may represent "responders" to disease-modifying treatment approaches, and future research may be able to identify these individuals and target treatment appropriately. On the other hand, a growing body of literature suggests that moving beyond localized physical interventions and treating individuals more holistically may hold promise for better outcomes [68]. Examples of such interventions may include improving patient education, weight management, treatments targeting pain perception and neuroplasticity, or psychological support [63]. Non-physical approaches such as these will likely be most effective in combination with an exercise-based approach, which is recommended as a core treatment for PFP and PFOA [23–26].

Limitations to the present study include broad eligibility criteria, which resulted in study heterogeneity. For example, we included both structural diagnoses and clinical definitions of PFP, including PFOA. However, these diagnoses fall along a spectrum of patellofemoral disease, so presenting interventions in a single review provides evidence across this spectrum. In addition, we included any pharmaceutical, nutraceutical, or surgical intervention, and did not limit control interventions. These broad inclusion criteria enabled us to capture more studies relating to our research question but should be considered in the context of synthesizing and interpreting of results, with very few studies including a control reflecting treatments of efficacy, usual care, or wait and see. Sample sizes were generally small in our included studies (11 studies had fewer than 50 participants). It was not possible to formally evaluate publication bias on account of the limited number of studies published for each interventions [69]. Finally, while limiting our inclusion to RCTs is arguably a strength of our study, one of the possible limitations is that we may have missed observational studies that reported poor or even harmful outcomes of some interventions. Thus, further evaluation of some procedures, such as anterior tibial tuberosity displacement, may be both unnecessary and inappropriate [70].

The results of the present systematic review suggest that clinical decisions regarding pharmaceutical, nutraceutical, and surgical interventions are currently being made with very little supporting evidence. Limited efficacy of these treatments in the target population(s) should be considered within the broader context of therapeutic options for PFP and PFOA. In particular, therapeutic exercise consistently demonstrates large effect sizes on pain and function in short and long-term studies [71], and evidence continues to emerge to guide exercise types and dose for improving these outcomes [32,72,73]. The present review suggests that rigorous research is urgently needed to support current best-practices in primary care for the management of PFP and PFOA. This is particularly relevant given the chronic nature of PFP, its association with PFOA [7,8,14–18], and the likelihood of PFOA progressing further towards whole knee OA [21].

In conclusion, our systematic review found very limited to moderate certainty of evidence, primarily based on single studies, and often comparing interventions to other interventions with no established efficacy. This limits our ability to offer robust recommendations as to which medical interventions may be effective in the management of PFP and PFOA.

Supplementary Materials: The following are available online at http://www.mdpi.com/2077-0383/9/11/3397/s1, Table S1: Individual study population and quality assessment of eligibility criteria (1 indicates quality indicator present, 0 not present), Table S2: GRADE assessment for pain outcomes.

Author Contributions: Conceptualization: E.M.M., H.F.H., D.T., C.J.B, K.M.C., S.M.A.B.-Z., M.v.M.; methodology, E.M.M., H.F.H., D.T., C.J.B, K.M.C., S.M.A.B.-Z., M.v.M.; formal analysis, E.M.M., H.F.H., D.T., and C.J.B.; writing—original draft preparation, E.M.M.; writing—review and editing, E.M.M., H.F.H., D.T., C.J.B, K.M.C., S.M.A.B.-Z., M.v.M.; supervision S.M.A.B.-Z. and M.v.M. All authors have read and agreed to the published version of the manuscript.

Funding: This study received no external funding.

Acknowledgments: We thank Wichor Bramer at Erasmus MC for assisting us with developing our search terms and managing our search.

Conflicts of Interest: Dr. Macri is funded by a Banting Postdoctoral Fellowship (Canadian Institutes of Health Research). Dr. Hart is supported in part by a Transdisciplinary Bone and Joint Training Award from the Collaborative Training Program in Musculoskeletal Health Research at Western University. Dr. Barton is supported by a Translating Research Into Practice Fellowship (Medical Research Future Fund, Australia). Dr. Bierma-Zeinstra reports grants from The Netherlands Organisation for Health Research and Development, CZ, European Union, Foreum, Dutch Arthritis Association, personal fees from Osteoarthritis Research Society International (OARSI), personal fees from Infirst Healthcare, and personal fees from Pfizer, all outside the submitted work.

References

1. Crossley, K.M.; Stefanik, J.J.; Selfe, J.; Collins, N.J.; Davis, I.S.; Powers, C.M.; McConnell, J.; Vicenzino, B.; Bazett-Jones, D.M.; Esculier, J.F.; et al. 2016 patellofemoral pain consensus statement from the 4th international patellofemoral pain research retreat, manchester. Part 1: Terminology, definitions, clinical examination, natural history, patellofemoral osteoarthritis and patient-reported outcome measures. *Br. J. Sports Med.* **2016**, *50*, 839–843. [PubMed]
2. Glaviano, N.R.; Baellow, A.; Saliba, S. Physical activity levels in individuals with and without patellofemoral pain. *Phys. Ther. Sport* **2017**, *27*, 12–16. [CrossRef]
3. Rathleff, M.S.; Rathleff, C.R.; Olesen, J.L.; Rasmussen, S.; Roos, E.M. Is knee pain during adolescence a self-limiting condition? Prognosis of patellofemoral pain and other types of knee pain. *Am. J. Sports Med.* **2016**, *44*, 1165–1171. [CrossRef]
4. Coburn, S.L.; Barton, C.J.; Filbay, S.R.; Hart, H.F.; Rathleff, M.S.; Crossley, K.M. Quality of life in individuals with patellofemoral pain: A systematic review including meta-analysis. *Phys. Ther. Sport* **2018**, *33*, 96–108. [CrossRef] [PubMed]
5. Collins, N.J.; Crossley, K.M. Young to middle-aged adults with persistent patellofemoral pain demonstrate similar pain, disability and knee-related quality of life as older adults with patellofemoral osteoarthritis. *Osteoarthr. Cartil.* **2019**, *27*, S220–S221. [CrossRef]
6. Maclachlan, L.R.; Collins, N.J.; Matthews, M.L.G.; Hodges, P.W.; Vicenzino, B. The psychological features of patellofemoral pain: A systematic review. *Br. J. Sports Med.* **2017**, *51*, 732–742. [CrossRef]

7. Collins, N.; Vicenzino, B.; Macri, E.; Crossley, K. Prevalence and factors associated with radiographic PFJ OA in young to middle-aged adults with chronic patellofemoral pain. *J. Sci. Med. Sport* **2015**, *19*, e85. [CrossRef]
8. Hinman, R.S.; Lentzos, J.; Vicenzino, B.; Crossley, K.M. Is patellofemoral osteoarthritis common in middle-aged people with chronic patellofemoral pain? *Arthritis Care Res.* **2014**, *66*, 1252–1257. [CrossRef]
9. Macri, E.M.; Neogi, T.; Tolstykh, I.; Widjajahakim, R.; Lewis, C.E.; Torner, J.C.; Nevitt, M.C.; Roux, M.; Stefanik, J.J. Relation of patellofemoral joint alignment, morphology, and radiographic osteoarthritis to frequent anterior knee pain: TheMOST study. *Arthritis Care Res.* **2019**, *72*, 1066–1073. [CrossRef]
10. Smith, B.E.; Selfe, J.; Thacker, D.; Hendrick, P.; Bateman, M.; Moffatt, F.; Rathleff, M.S.; Smith, T.O.; Logan, P. Incidence and prevalence of patellofemoral pain: A systematic review and meta-analysis. *PLoS ONE* **2018**, *13*, e0190892. [CrossRef]
11. Hart, H.F.; Stefanik, J.J.; Wyndow, N.; Machotka, Z.; Crossley, K.M. The prevalence of radiographic and mri-defined patellofemoral osteoarthritis and structural pathology: A systematic review and meta-analysis. *Br. J. Sports Med.* **2017**, *51*, 1195–1208. [CrossRef]
12. Van Middelkoop, M.; Van Linschoten, R.; Berger, M.Y.; Koes, B.W.; Bierma-Zeinstra, S.M. Knee complaints seen in general practice: Active sport participants versus non-sport participants. *BMC Musculoskelet. Disord.* **2008**, *9*, 36. [CrossRef]
13. Glaviano, N.R.; Kew, M.; Hart, J.M.; Saliba, S. Demographic and epidemiological trends in patellofemoral pain. *Int. J. Sports Phys.* **2015**, *10*, 281–290.
14. Kastelein, M.; Luijsterburg, P.A.J.; Heintjes, E.M.; van Middelkoop, M.; Verhaar, J.A.N.; Koes, B.W.; Bierma-Zeinstra, S.M.A. The 6-year trajectory of non-traumatic knee symptoms (including patellofemoral pain) in adolescents and young adults in general practice: A study of clinical predictors. *Br. J. Sports Med.* **2015**, *49*, 400–405. [CrossRef] [PubMed]
15. Nimon, G.; Murray, D.; Sandow, M.; Goodfellow, J. Natural history of anterior knee pain: A 14-to 20-year follow-up of nonoperative management. *J. Pediatric Orthop.* **1998**, *18*, 118–122. [CrossRef]
16. Thomas, M.J.; Wood, L.; Selfe, J.; Peat, G. Anterior knee pain in younger adults as a precursor to subsequent patellofemoral osteoarthritis: A systematic review. *BMC Musculoskelet. Disord.* **2010**, *11*, 201. [CrossRef] [PubMed]
17. Utting, M.; Davies, G.; Newman, J. Is anterior knee pain a predisposing factor to patellofemoral osteoarthritis? *Knee* **2005**, *12*, 362–365. [CrossRef] [PubMed]
18. Wyndow, N.; Collins, N.; Vicenzino, B.; Tucker, K.; Crossley, K. Is there a biomechanical link between patellofemoral pain and osteoarthritis? A narrative review. *Sports Med.* **2016**, *46*, 1797–1808. [CrossRef] [PubMed]
19. Duncan, R.; Peat, G.; Thomas, E.; Wood, L.; Hay, E.; Croft, P. How do pain and function vary with compartmental distribution and severity of radiographic knee osteoarthritis? *Rheumatology* **2008**, *47*, 1704–1707. [CrossRef]
20. Hart, H.F.; Filbay, S.R.; Coburn, S.; Charlton, J.M.; Sritharan, P.; Crossley, K.M. Is quality of life reduced in people with patellofemoral osteoarthritis and does it improve with treatment? A systematic review, meta-analysis and regression. *Disabil. Rehabil.* **2019**, *41*, 2979–2993. [CrossRef]
21. Lankhorst, N.; Damen, J.; Oei, E.; Verhaar, J.; Kloppenburg, M.; Bierma-Zeinstra, S.; van Middelkoop, M. Incidence, prevalence, natural course and prognosis of patellofemoral osteoarthritis: The cohort hip and cohort knee study. *Osteoarthr. Cartil.* **2017**, *25*, 647–653. [CrossRef] [PubMed]
22. Brand, C.A.; Harrison, C.; Tropea, J.; Hinman, R.S.; Britt, H.; Bennell, K. Management of osteoarthritis in general practice in australia. *Arthritis Care Res.* **2014**, *66*, 551–558. [CrossRef] [PubMed]
23. Lack, S.; Neal, B.; De Oliveira Silva, D.; Barton, C. How to manage patellofemoral pain—Understanding the multifactorial nature and treatment options. *Phys. Ther. Sport* **2018**, *32*, 155–166. [CrossRef]
24. Crossley, K.M.; Vicenzino, B.; Lentzos, J.; Schache, A.G.; Pandy, M.G.; Ozturk, H.; Hinman, R.S. Exercise, education, manual-therapy and taping compared to education for patellofemoral osteoarthritis: A blinded, randomised clinical trial. *Osteoarthr. Cartil.* **2015**, *23*, 1457–1464. [CrossRef] [PubMed]
25. Bannuru, R.R.; Osani, M.C.; Vaysbrot, E.E.; Arden, N.K.; Bennell, K.; Bierma-Zeinstra, S.M.A.; Kraus, V.B.; Lohmander, L.S.; Abbott, J.H.; Bhandari, M.; et al. Oarsi guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. *Osteoarthr. Cartil.* **2019**, *27*, 1578–1589. [CrossRef]

26. Willy, R.W.; Hoglund, L.T.; Barton, C.J.; Bolgla, L.A.; Scalzitti, D.A.; Logerstedt, D.S.; Lynch, A.D.; Snyder-Mackler, L.; McDonough, C.M.; Altman, R. Patellofemoral pain: Clinical practice guidelines linked to the international classification of functioning, disability and health from the academy of orthopaedic physical therapy of the american physical therapy association. *J. Orthop. Sports Phys. Ther.* **2019**, *49*, CPG1–CPG95. [CrossRef]
27. Rodriguez-Merchan, E.C. Evidence based conservative management of patello-femoral syndrome. *Arch. Bone Jt. Surg.* **2014**, *2*, 4–6.
28. Castrogiovanni, P.; Trovato, F.M.; Loreto, C.; Nsir, H.; Szychlinska, M.A.; Musumeci, G. Nutraceutical supplements in the management and prevention of osteoarthritis. *Int. J. Mol. Sci.* **2016**, *17*, 2042. [CrossRef]
29. Rodriguez-Merchan, E.C. Surgical treatment of isolated patellofemoral osteoarthritis. *HSS J.* **2014**, *10*, 79–82. [CrossRef]
30. Rodriguez-Merchan, E.C. The Present Situation of Patellofemoral Arthroplasty in the Management of Solitary Patellofemoral Osteoarthritis. *Arch. Bone Jt. Surg.* **2020**, *8*, 325–331.
31. Barton, C.J.; Lack, S.; Hemmings, S.; Tufail, S.; Morrissey, D. The 'best practice guide to conservative management of patellofemoral pain': Incorporating level 1 evidence with expert clinical reasoning. *Br. J. Sports Med.* **2015**, *49*, 923–934. [CrossRef] [PubMed]
32. Collins, N.J.; Barton, C.J.; van Middelkoop, M.; Callaghan, M.J.; Rathleff, M.S.; Vicenzino, B.T.; Davis, I.S.; Powers, C.M.; Macri, E.M.; Hart, H.F.; et al. 2018 consensus statement on exercise therapy and physical interventions (orthoses, taping and manual therapy) to treat patellofemoral pain: Recommendations from the 5th international patellofemoral pain research retreat, gold coast, australia, 2017. *Br. J. Sports Med.* **2018**, *52*, 1170–1178. [CrossRef] [PubMed]
33. Heintjes, E.M.; Berger, M.; Bierma-Zeinstra, S.M.A.; Bernsen, R.M.D.; Verhaar, J.A.N.; Koes, B.W. Pharmacotherapy for patellofemoral pain syndrome. *Cochrane Database Syst. Rev.* **2004**. [CrossRef] [PubMed]
34. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; Group, P. Preferred reporting items for systematic reviews and meta-analyses: The prisma statement. *Open Med.* **2009**, *3*, e123–e130. [PubMed]
35. Prospero International Prospective Register of Systematic Reviews. Available online: http://www.crd.york.ac.uk/prospero (accessed on 8 December 2017).
36. Grelsamer, R.P. Patellar nomenclature: The Tower of Babel revisited. *Clin. Orthop. Relat. Res.* **2005**, *436*, 60–65. [CrossRef]
37. Higgins, J.P.T.; Sterne, J.A.C.; Savović, J.; Page, M.J.; Hrobjartsson, A.; Boutron, I.; Reeves, B.; Eldridge, S.A. A revised tool for assessing risk of bias in randomized trials. *Cochrane Database Syst. Rev.* **2016**, *10* (Suppl 1), 29–31.
38. Collins, N.J.; Bisset, L.M.; Crossley, K.M.; Vicenzino, B. Efficacy of nonsurgical interventions for anterior knee pain: Systematic review and meta-analysis of randomized trials. *Sports Med.* **2012**, *42*, 31–49. [CrossRef]
39. Atkins, D.; Best, D.; Briss, P.A.; Eccles, M.; Falck-Ytter, Y.; Flottorp, S.; Guyatt, G.H.; Harbour, R.T.; Haugh, M.C.; Henry, D.; et al. Grading quality of evidence and strength of recommendations. *BMJ* **2004**, *328*, 1490.
40. Odgaard, A.; Madsen, F.; Kristensen, P.W.; Kappel, A.; Fabrin, J. The mark coventry award: Patellofemoral arthroplasty results in better range of movement and early patient-reported outcomes than TKA. *Clin. Orthop. Relat. Res.* **2018**, *476*, 87. [CrossRef]
41. Kettunen, J.A.; Harilainen, A.; Sandelin, J.; Schlenzka, D.; Hietaniemi, K.; Seitsalo, S.; Malmivaara, A.; Kujala, U.M. Knee arthroscopy and exercise versus exercise only for chronic patellofemoral pain syndrome: A randomized controlled trial. *BMC Med.* **2007**, *5*, 38. [CrossRef]
42. Kettunen, J.A.; Harilainen, A.; Sandelin, J.; Schlenzka, D.; Hietaniemi, K.; Seitsalo, S.; Malmivaara, A.; Kujala, U.M. Knee arthroscopy and exercise versus exercise only for chronic patellofemoral pain syndrome: 5-year follow-up. *Br. J. Sports Med.* **2012**, *46*, 243–246. [CrossRef] [PubMed]
43. Joseph, M.; Achten, J.; Parsons, N.; Costa, M. The PAT randomised clinical trial: Total knee arthroplasty versus patellofemoral arthroplasty in patients with severe arthritis of the patellofemoral joint. *Bone Joint J.* **2020**, *102*, 310–318. [CrossRef]
44. Hart, J.M.; Kuenze, C.; Norte, G.; Bodkin, S.; Patrie, J.; Denny, C.; Hart, J.; Diduch, D.R. Prospective, randomized, double-blind evaluation of the efficacy of a single-dose hyaluronic acid for the treatment of patellofemoral chondromalacia. *Orthop. J. Sport Med.* **2019**, *7*, 2325967119854192. [CrossRef] [PubMed]

45. Antich, T.J.; Randall, C.C.; Westbrook, R.A.; Morrissey, M.C.; Brewster, C.E. Physical therapy treatment of knee extensor mechanism disorders: Comparison of four treatment modalities. *J. Orthop. Sports Phys.* **1986**, *8*, 255–259. [CrossRef]
46. Matoso, L.; Schelling, J.L.; Radi, I. Essai clinique contrôlé de la chloroquine dans la chondromalacie rotulienne. *Med. Hyg.* **1980**, *38*, 1205–1209.
47. Örsçelik, A.; Yildiz, Y. Comparison of single and triple platelet rich plasma injections in the treatment of patellofemoral pain syndrome. *Turk. Klin. J. Med Sci.* **2015**, *35*, 78–87. [CrossRef]
48. Raatikainen, T.; Väänänen, K.; Tamelander, G. Effect of glycosaminoglycan polysulfate on chondromalacia patellae: A placebo-controlled 1-year study. *Acta Orthop. Scand.* **1990**, *61*, 443–448. [CrossRef]
49. Suter, E.; Herzog, W.; De Souza, K.; Bray, R. Inhibition of the quadriceps muscles in patients with anterior knee pain. *J. Appl. Biomech.* **1998**, *14*, 360–373. [CrossRef]
50. Pagenstert, G.; Wolf, N.; Bachmann, M.; Gravius, S.; Barg, A.; Hintermann, B.; Wirtz, D.C.; Valderrabano, V.; Leumann, A.G. Open lateral patellar retinacular lengthening versus open retinacular release in lateral patellar hypercompression syndrome: A prospective double-blinded comparative study on complications and outcome. *Arthrosc. J. Arthrosc. Relat. Surg.* **2012**, *28*, 788–797. [CrossRef]
51. Owens, B.D.; Stickles, B.J.; Balikian, P.; Busconi, B.D. Prospective analysis of radiofrequency versus mechanical debridement of isolated patellar chondral lesions. *Arthrosc. J. Arthrosc. Relat. Surg.* **2002**, *18*, 151–155. [CrossRef]
52. Hejgaard, N.; Watt-Boolsen, S. The effect of anterior displacement of the tibial tuberosity in idiopathic chondromalacia patellae: A prospective randomized study. *Acta Orthop. Scand.* **1982**, *53*, 135–139. [CrossRef]
53. Darracott, J. Treatment of the painful knee fulfilling diagnostic criteria for 'chondromalacia patellae'. *Curr. Med. Res. Opin.* **1973**, *1*, 412–422. [CrossRef]
54. Bentley, G.; Leslie, I.J.; Fischer, D. Effect of aspirin treatment on chondromalacia patellae. *Ann. Rheum. Dis.* **1981**, *40*, 37–41. [CrossRef]
55. Fulkerson, J.P.; Folcik, M.A. Comparison of diflunisal and naproxen for relief of anterior knee pain. *Clin. Ther.* **1986**, *9*, 59–61. [PubMed]
56. Martí-Bonmatí, L.; Sanz-Requena, R.; Rodrigo, J.L.; Alberich-Bayarri, Á.; Carot, J.M. Glucosamine sulfate effect on the degenerated patellar cartilage: Preliminary findings by pharmacokinetic magnetic resonance modeling. *Eur. Radiol.* **2009**, *19*, 1512–1518. [CrossRef]
57. Singer, B.J.; Silbert, P.L.; Song, S.; Dunne, J.W.; Singer, K.P. Treatment of refractory anterior knee pain using botulinum toxin type a (dysport) injection to the distal vastus lateralis muscle: A randomised placebo controlled crossover trial. *Br. J. Sports Med.* **2011**, *45*, 640–645. [CrossRef]
58. Kannus, P.; Natri, A.; Niittymäki, S.; Järvinen, M. Effect of intraarticular glycosaminoglycan polysulfate treatment on patellofemoral pain syndrome. A prospective, randomized double-blind trial comparing glycosaminoglycan polysulfate with placebo and quadriceps muscle exercises. *Arthritis Rheum. Offic. J. Am. Coll. Rheumatol.* **1992**, *35*, 1053–1061. [CrossRef]
59. Kannus, P.; Natri, A.; Paakkala, T.; JÄrvinen, M. An outcome study of chronic patellofemoral pain syndrome. Seven-year follow-up of patients in a randomized, controlled trial. *J. Bone Joint Surg.* **1999**, *81*, 355–363. [CrossRef] [PubMed]
60. Nakhostin-Roohi, B.; Khoshkhahesh, F.; Bohlooli, S. Effect of virgin olive oil versus piroxicam phonophoresis on exercise-induced anterior knee pain. *Avicenna J. Phytomedicine* **2016**, *6*, 535.
61. O'Neill, D.B. Open lateral retinacular lengthening compared with arthroscopic release. A prospective, randomized outcome study. *J. Bone Joint Surg.* **1997**, *79*, 1759–1769.
62. Nüesch, E.; Trelle, S.; Reichenbach, S.; Rutjes, A.W.S.; Tschannen, B.; Altman, D.G.; Egger, M.; Jüni, P. Small study effects in meta-analyses of osteoarthritis trials: Meta-epidemiological study. *BMJ* **2010**, *341*, c3515. [CrossRef]
63. Tarassoli, P.; Punwar, S.; Khan, W.; Johnstone, D. Suppl 2: Patellofemoral arthroplasty: A systematic review of the literature. *Open Orthop. J.* **2012**, *6*, 340. [CrossRef]
64. Parratte, S.; Ollivier, M.; Lunebourg, A.; Abdel, M.P.; Argenson, J.N. Long-term results of compartmental arthroplasties of the knee: Long term results of partial knee arthroplasty. *Bone Jt. J.* **2015**, *97*, 9–15. [CrossRef]
65. Bedson, J.; Croft, P.R. The discordance between clinical and radiographic knee osteoarthritis: A systematic search and summary of the literature. *BMC Musculoskelet. Disord.* **2008**, *9*, 116. [CrossRef] [PubMed]

66. Van der Heijden, R.A.; de Kanter, J.L.; Bierma-Zeinstra, S.M.; Verhaar, J.A.; van Veldhoven, P.L.; Krestin, G.P.; Oei, E.H.; van Middelkoop, M. Structural abnormalities on magnetic resonance imaging in patients with patellofemoral pain a cross-sectional case-control study. *Am. J. Sports Med.* **2016**, *44*, 2339–2346. [CrossRef]
67. Stefanik, J.J.; Neogi, T.; Niu, J.; Roemer, F.W.; Segal, N.A.; Lewis, C.E.; Nevitt, M.; Guermazi, A.; Felson, D.T. The diagnostic performance of anterior knee pain and activity-related pain in identifying knees with structural damage in the patellofemoral joint: The multicenter osteoarthritis study. *J. Rheumatol.* **2014**, *41*, 1695–1702. [CrossRef]
68. Barton, C.J.; Crossley, K.M.; Macri, E.M. Should we consider changing traditional physiotherapy treatment of patellofemoral pain based on recent insights from the literature? *Br. J. Sports Med.* **2018**, *52*, 1546–1547. [CrossRef] [PubMed]
69. Sterne, J.A.C.; Egger, M. Regression methods to detect publication and other bias in meta-analysis. In *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*; Rothstein, H.R., Sutton, A.J., Borenstein, M., Eds.; John Wiley & Sons, Ltd.: West Sussex, UK, 2005; pp. 99–110.
70. Karlsson, J.; Sward, L.; Lansinger, O. Bad results after anterior advancement of the tibial tubercle for patello-femoral pain syndrome. *Arch. Orthop. Trauma Surg.* **1992**, *111*, 195–197. [CrossRef]
71. Van der Heijden, R.A.; Lankhorst, N.E.; Van Linschoten, R.; Bierma-Zeinstra, S.; Van Middelkoop, M. Exercise for treating patellofemoral pain syndrome. An abridged version of cochrane systematic review. *Eur. J. Phys. Rehabil. Med.* **2015**, *52*, 110–133.
72. Thomson, C.; Krouwel, O.; Kuisma, R.; Hebron, C. The outcome of hip exercise in patellofemoral pain: A systematic review. *Man. Ther.* **2016**, *26*, 1–30. [CrossRef]
73. Lack, S.; Barton, C.; Sohan, O.; Crossley, K.; Morrissey, D. Proximal muscle rehabilitation is effective for patellofemoral pain: A systematic review with meta-analysis. *Br. J. Sports Med.* **2015**, *49*, 1365–1376. [CrossRef] [PubMed]

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 978-3-0365-2661-4