Determinants of progression of hip osteoarthritis

Max Reijman

Acknowledgements

Determinants of progression of hip osteoarthritis

Determinanten van progressie van heup artrose

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de Rector Magnificus
Prof.dr. S.W.J Lamberts
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag 1 december 2004 om 15.45 uur

door

Maxentius Reijman
geboren te 's Gravenhage

Promotiecommissie

Promotoren   : Prof. dr. B.W. Koes
                  Prof. dr. J.M.W. Hazes

Overige Leden : Prof. dr. H.A.P. Pols
                  Prof. dr. J.A.N. Verhaar
                  Prof. dr. W.J.J. Assendelft

Copromotor   : Dr. S.M.A. Bierma-Zeinstra

Voor Janine, Tim, Amber en Mats

Contents

Manuscripts based on studies described in this thesis

Chapter 2

Reijman M, Hazes JMW, Koes BW, Verhagen AP, Bierma-Zeinstra SMA. Validity, reliability and applicability of seven definitions of hip osteoarthritis used in epidemiological studies: a systematic appraisal. Ann Rheum Dis 2004;63(3):226-32.

Chapter 3

Reijman M, Hazes JMW, Pols HAP, Bernsen RMD, Koes BW, Bierma-Zeinstra SMA. Validity and reliability of three definitions of hip osteoarthritis. Ann Rheum Dis 2004; in press.

Chapter 4

Reijman M, Hazes JMW, Bierma-Zeinstra SMA, Koes BW, Christgau S, Christiansen C, Uitterlinden AG, Pols HAP. A new marker for osteoarthritis: cross-sectional and longitudinal approach. Arthritis Rheum 2004;50(8):2471-8.

Chapter 5

Reijman M, Hazes JMW, Pols HAP, Bernsen RMD, Koes BW, Bierma-Zeinstra SMA. X-ray findings strongly predict progression of osteoarthritis of the hip. Submitted

Chapter 6

Reijman M, Hazes JMW, Pols HAP, Koes BW, Bierma-Zeinstra SMA. Acetabular dysplasia predicts incident osteoarthritis of the hip: the Rotterdam Study. Submitted
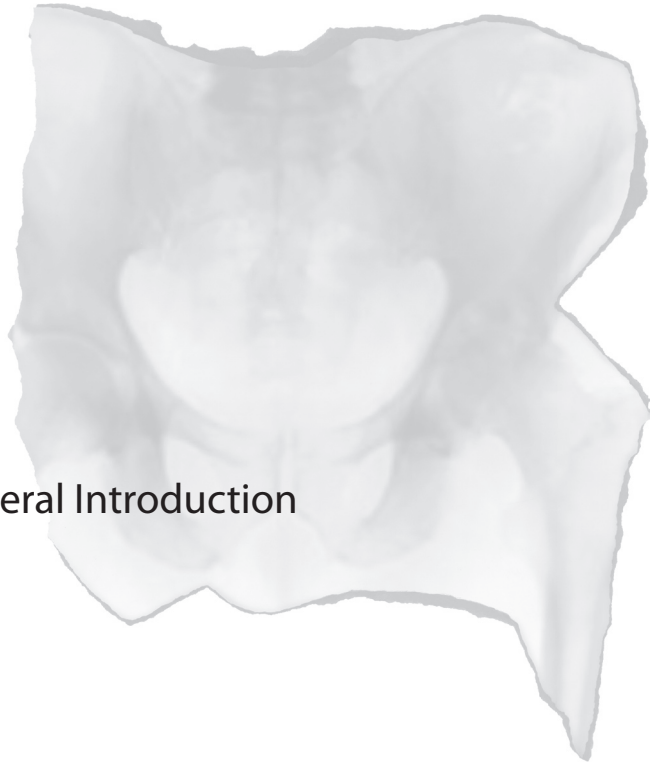
Chapter 7

Reijman M, Bierma-Zeinstra SMA, Pols HAP, Koes BW, Stricker BHC, Hazes JMW. Is there an association between the use of different types of NSAIDs and radiologic progression of osteoarthritis? The Rotterdam Study.

1

General Introduction

Osteoarthritis (OA), also called degenerative joint disease, is one of the most frequently occurring disorders of the locomotor system. It is a disease of the whole joint in which all articular structures are affected, including articular cartilage loss, development of marginal outgrowths, osteophytes, and increased thickness of the bony envelope (subchondral sclerosis). Soft-tissue structures in and around the joint are also affected, including synovium, ligaments and bridging muscles. Patients with OA of the hip or knee have pain that typically worsens with weight bearing and activity and improves with rest, as well as morning stiffness and gelling of the involved joint after periods of inactivity (1). Especially when the hip or knee is involved, it accounts for more difficulty in walking, stair climbing, and other lower extremity tasks (2). On physical examination, there is often tenderness on palpation, bony enlargement, crepitus on motion, and/or limitation of joint motion (1). Because of the longevity of working careers and the substantial prevalence of OA in middle-aged persons, OA causes a considerable burden in lost time at work and early retirement (3). OA is expected to be the fourth leading cause of disability by the year 2020 (4). Recent estimates suggest that total costs for arthritis, including OA, may exceed 2% of the gross domestic product (3).

Based on the baseline results of the Rotterdam Study (5) and standardized to the Dutch society (6), the estimated number of persons with radiologic OA (ROA) of the hip in 2000 was 257,400. This number of persons with OA of the hip is somewhat higher than the number known by the GPs (6). In a large follow-up study (EPOZ) of men and women aged between 45 and 65 years, it was found that the nine-year incidence of ROA of the hip was 9% for men and 12% for women (7). Because the prevalence of OA will increase with the aging of the Western society, it is expected that the percentage of persons with OA will increase by 37.7% between 2000 and 2020 (6). Over the last two decades many epidemiological studies have investigated the different determinants of the occurrence of hip OA. OA of the hip is assumed to be a multifactorial disease involving both systemic factors, such as metabolic, hormonal, genetic, age and gender, and local biomechanical factors, such as mechanical workload, mechanical load of sport activity, obesitas and acetabular dysplasia. Dieppe introduced the model proposing that the joint becomes susceptible for OA by systemic factors and that local biomechanical factors play the final role in determining site and severity of OA (8) (see Figure 1). This interplay between systemic and local factors was investigated among families in whom inherited OA occurs at an early age, and in each member of the susceptible families, different joints were affected, suggesting that local biomechanical factors had acted on a predisposed joint (9). To date, many more studies have investigated the epidemiology of knee OA than of hip OA. However, the influence of local biomechanical factors on the occurrence of hip OA has not been well explored. Furthermore, the independent influence of the different determinants on the occurrence of hip OA and also the interaction between different determinants is not clear. This is especially important because the identification of modifiable determinants of hip OA may help to develop preventive strategies. Furthermore,
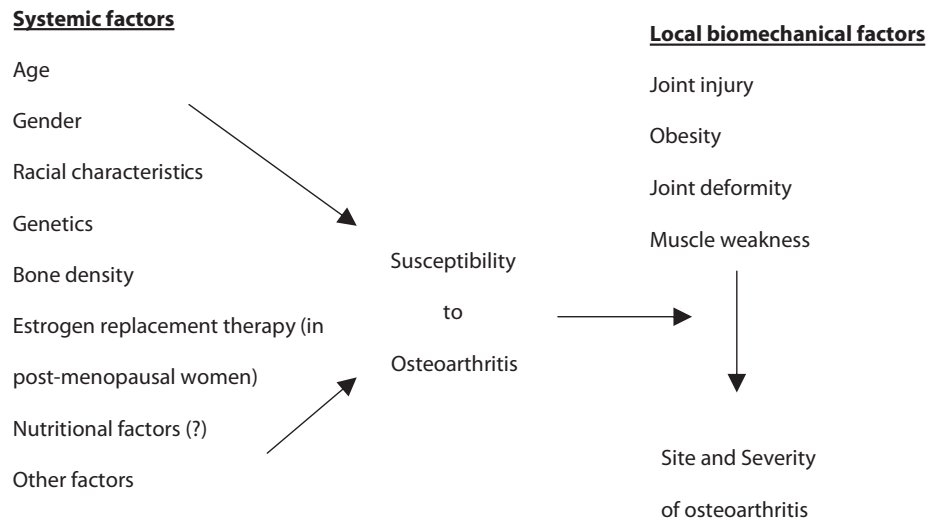
Systemic factors

Age

Gender

Racial characteristics

Genetics

Bone density

Estrogen replacement therapy (in

post-menopausal women)

Nutritional factors (?)

Other factors

Local biomechanical factors

Joint injury

Obesity

Joint deformity

Muscle weakness

Susceptibility

to

Osteoarthritis

Site and Severity

of osteoarthritis

Figure 1. Diagram showing the pathogenesis of osteoarthritis with putative risk factors adapted from Dieppe (8).

a pathogenetic approach may help elucidate the role of the different determinants in the occurrence of OA. Based on these arguments, we conclude that hip OA is a particularly important topic to investigate more thoroughly.

Case definition

A major problem in studying hip OA is the absence of consensus in defining hip OA for epidemiological studies. In most epidemiological studies OA is assessed by means of radiologic evaluation, because radiographs are easily obtainable and relatively cheap for large epidemiological studies. The most commonly used radiological definition of hip OA – the Kellgren & Lawrence grade – is also the one that is most criticized. Therefore several other definitions of hip OA have been proposed during the last decades. One limitation of a radiological definition of hip OA is that the majority of the subjects with radiographic evidence of OA have no symptoms (17–33% of the persons with ROA have joint pain). Moreover, not all people with symptomatic OA have radiographic evidence of OA. Another limitation of radiographs is that significant osteoarthritic changes must already have occurred in order to be visible on a radiograph. To overcome this, biochemical markers aiming to detect OA in an early stage have been developed. Such a biochemical marker might also be useful for identification of patients at high risk for progression, and for a faster assessment of therapeutic response in OA. Moreover, all imaging techniques provide a historical view

of damage that has already occurred, rather than assessing the current rate of disease progression (10).

Progression

The identification of patients at high risk for progression of hip OA is important for at least two reasons. Firstly, well-characterized 'high risk' groups may be useful in clinical trials and, secondly, assuming that disease-modifying OA drugs might become available in the future, to identify primary target groups in need of such therapy. Up till now the prognostic factors of progression of hip OA have been investigated in small studies, with a short follow-up time and in a hospital setting only. Hence, the conclusion that until now the main predictive factors showed to be radiological features (11) should be investigated in an open population and in a primary care setting.

NSAIDs and progression

There is no known cure for OA, and currently the main goals of medical management of patients with OA are symptomatic relief and preservation of function.
In about 80% of the patients with symptomatic hip OA the clinician will prescribe a non-steroidal anti-inflammatory drug (NSAID) (12). Efficacy and side-effects of NSAIDs are well understood. However, it remains controversial as to what effects these agents have on the progression of OA. Several in-vitro studies of human cartilage and animal studies suggest that some NSAIDs inhibit the synthesis of glycosaminoglycans and other aspects of articular cartilage metabolism, while others are supposed to have a neutral effect on cartilage metabolism (13). Whether the rate of progression is also increased in patients receiving such NSAIDs remains an open question.

Aim of this thesis

The overall objective of the studies described in this thesis is to determine the prognostic factors of osteoarthritis of the hip in a large open population. Based on the model of Dieppe the prognostic factors were divided in systemic factors (e.g. age, gender, genetics, hormonal influence) and metabolic factors; and in local biomechanical factors such as mechanical load by work or sport activity, weight, and acetabular dysplasia.

Description of chapters

In **chapter 2** we summarize and review articles addressing the quality, in terms of validity, reliability, applicability, of seven commonly used definitions of hip OA for epidemiological studies, primarily used as classification criteria.

In **chapter 3** we directly compare the reliability and validity of three frequently used radiological definitions of hip OA namely, Kellgren & Lawrence grade, Minimal Joint Space and Croft's grade in a large open population aged 55 years and over. Additionally, we investigate whether the validity of the three definitions of hip OA is gender dependent.

In **chapter 4** we investigate the association between urinary concentrations of C-telopeptide fragments of collagen type II (CTX-II) and the prevalence and progression of radiologic OA of the hip and knee in a large open population aged 55 years and over with a long-term follow-up. Additionally, we repeated the analyses in those subjects with pain at baseline (hip or knee).

In **chapter 5** we investigate which determinants will best identify those persons who are at high risk for progression of hip OA in a large open population aged 55 years and over, with a long-term follow-up period.

In **chapter 6** we investigate the association between radiographic evidence of acetabular dysplasia in participants without ROA of the hip at baseline, and an incident hip ROA, in a large open population aged 55 years and over, with a long-term follow-up period. Additionally we investigate whether the association between acetabular dysplasia and incident hip ROA is modified by other determinants of hip OA.

In **chapter 7** we investigate the associations between two groups of NSAIDs, those (indomethacin, naproxen and ibuprofen) that are supposed to have a deleterious effect on joint cartilage and those NSAIDs (such as diclofenac and piroxicam) that are supposed to have a neutral effect on cartilage metabolism, and progression of OA of the hip and knee in a large open population aged 55 years and over with a long-term follow-up period. Additionally, we investigated the associations between each of the NSAIDs and progression of OA of the hip and knee.

Finally, in **chapter 8**, the most important results of these studies, as well as their limitations and implications are discussed.

References

1.   Altman RD, Hochberg MC, Moskowitz RW, Schnitzer TJ. Recommendations for the medical management of osteoarthritis of the hip and knee. Arthritis Rheum 2000;43:1905–15.
2.   Guccione AA, Felson DT, Anderson JJ, Zhang Y, Wilson PW. The effects of specific medical conditions on functional limitations of elders in the Framingham Study. Am J Public Health 1994;84:351–8.

3.   Yelin E. The economics of osteoarthritis. Brandt KD, Doherty M, Lohmander LS, editors. Osteoarthritis. New York: Oxford University Press; 1998:23–30.

4.   Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. Bull World Health Organ 2003;81(9):646–56.

5.   Odding E, Valkenburg HA, Algra D, VandenOuweland FA, Grobbee DE, Hofman A. Associations of radiological osteoarthritis of the hip and knee with locomotor disability in the Rotterdam Study. Ann Rheum Dis 1998;57(4):203–8.

6.   Schouten JSAG, Gijsen R, Poos MJJC. Artrose. Omvang van het probleem. Research for man and environment (RIVM) www.rivm.nl/vtv/kompas/gezondheidstoestand/ziekte/artrose; (access date)26–04–2004.

7.   Schouten JS, VandenOuweland FA, Valkenburg HA. A 12-year follow-up study in the general population on prognostic factors of cartilage loss in osteoarthritis of the knee. Ann Rheum Dis 1992;51(8):932–7.

8.   Dieppe PA. Recommended methodology for assessing the progression of osteoarthritis of the hip and knee joints. Osteoarthritis Cartilage 1995;3:73–7.

9.   Felson DT, Zhang Y. An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. Arthritis Rheum 1998;41(8):1343–55.

10. DeGroot J, Bank RA, Tchetverikov I, Verzijl N, TeKoppele JM. Molecular markers for osteoarthritis: the road ahead. Curr Opin Rheumatol 2002;14:585–9.

11. Lievense AM, Bierma-Zeinstra SMA, Verhagen AP, Verhaar JAN, Koes BW. Prognostic factors of progress of osteoarthritis of the hip: A systematic review. Arthritis Rheum 2002;47(5):556–62.

12. Bierma-Zeinstra SMA, Lipschart S, Njoo KH, Bernsen R, Verhaar JAN, Prins A, Bohnen AM. How do general practitioners manage hip problems in adults? Scan J Prim Health Care 2000;18:159–64.

13. Ding C. Do NSAIDs affect the progression of osteoarthritis? Inflammation 2002;26(3):139–42.

2

Validity, reliability and
applicability of seven definitions
of hip osteoarthritis used in
epidemiological studies: a
systematic appraisal.
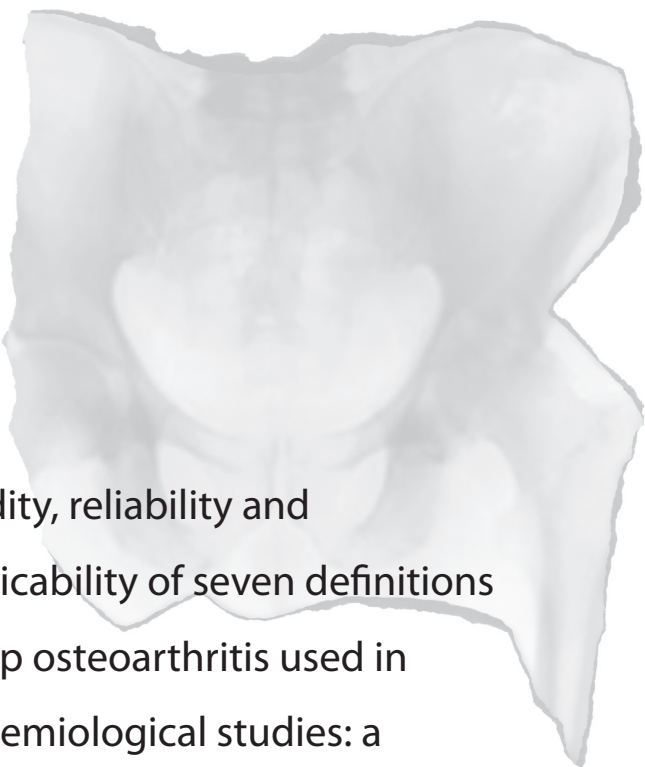
Abstract

Objective: To summarise and review articles addressing quality (in terms of validity, reliability, applicability) of seven commonly used definitions of hip osteoarthritis (OA) for epidemiological studies, in order to use it primary as classification criteria.

Methods: Relevant articles were identified based on a search in Medline and Embase. Articles with the aim to study the validity, reliability or applicability of the definitions of hip OA were selected. Two reviewers independently performed data extraction of the quality of the 7 definitions of hip OA.

Results: Review of the literature reveals that particularly the validity of the various definitions of hip OA has barely been investigated. Minimal joint space (MJS) demonstrated the highest (intra- and inter-rater) reliability, and showed the highest association with hip pain and restricted internal rotation compared to the other definitions of hip OA. The reliability of the Kellgren & Lawrence grade and the index according to Lane is comparable to that of MJS, but the construct validity should be investigated more thoroughly. The Croft grade, appeared to be inferior to the MJS, the Kellgren & Lawrence grade and the index according to Lane, regarding reliability and validity. Despite a precise and extensive method of development, the ACR criteria showed poor reliability and poor cross-validity (agreement between 3 ACR criteria sets) in a primary care setting.

Conclusions: Summarising the literature, it showed that the reliability of MJS, Kellgren & Lawrence and index according to Lane was comparable, but MJS had the highest relationship with hip pain in a male population. Considering how frequently the definitions of hip OA are used, it is surprising that the validity has been so poorly investigated, therefore we recommend that the validity be studied more thoroughly.

Introduction

Osteoarthritis (OA) is the most common joint disorder [1] and represent a considerable burden to society. Depending on the definition of hip OA used, the prevalence ranges from 7 to 25% in persons aged 55 years and over [2]. The hip is particularly interesting because it is often the sole joint affected by OA, suggesting an important role of local biomechanical risk factors. In addition the hip is crucial to independent function [3].

A major problem in studying hip OA, is the absence of consensus in defining hip OA for epidemiological and clinical studies [4]. Most epidemiological studies have used a single hallmark of hip OA (namely radiological changes) to define hip OA [5, 6].

To investigate (potential) risk factors, a valid and reliable definition of hip OA is required. Therefore we appraised the quality (in terms of validity, reliability and applicability) of seven definitions of hip OA commonly used for epidemiological studies:

1. The radiological grading system of Kellgren & Lawrence [7];
2. Croft's radiological grading system (a modification of the Kellgren & Lawrence grading system) [8];
3. Minimal Joint Space (MJS) according to Croft et al. (a measurement of the narrowing of the joint space) [8];
4. Measurement of the joint space according to Resnick & Niwayama [9];
5. Three sets of criteria (1 clinical, and 2 combined sets of clinical and radiographic criteria) of the American College of Rheumatology (ACR) [10];
6. Clinical definition of hip OA: radiological OA combined with pain in the hip region [11, 12];
7. Radiographic index grade according to Lane [13, 14].

The objective of the present study was to review the quality (reliability, validity, applicability) of these seven definitions of hip OA commonly used epidemiological studies, in order to use it primary as classification criteria [15, 16].

Methods

The literature was searched for all relevant papers containing one of the seven definitions of hip OA. Studies, which fulfilled predefined inclusion criteria were identified and subsequently assessed on aspects of reliability, validity, and applicability of the definition of hip OA used in each particular study.

Identification of the literature

To identify the studies a search was made in the following databases: Medline / Pubmed (1966 – March 2002), Cochrane Library and Embase (1990 – March, 2002). The specific keywords were: "osteoarthritis, hip" or "osteoarthritis" and "hip" and "clinical definition", "radiological definition", "case definition", "radiographic grading", "diagnosis", "severity", "index of severity", "classification criteria", "radiographic change", "minimal joint space", "Kellgren", "Kellgren and Lawrence", "reliability", "reproducibility of results", "epidemiologic studies" or "feasibility studies". The search was extended by screening the reference lists of all relevant articles identified. We repeated the search using the keywords of all selected articles.

Criteria for studies considered for inclusion

A study was included in this review if it fulfilled all of the following criteria: 1) the study population contained persons with and persons without hip OA, 2) it was an original article or a systematic review, 3) at least one of the seven definitions of hip OA investigated here was used, 4a) the study described the design, or the reliability, or the validity, or the applicability of at least one of the above mentioned definitions, **or** 4b) the study investigated the risk factors (or determinants) of hip OA, and used at least two of the above mentioned definitions.

Critical assessment of OA definitions

Using the information from the criteria of Buchbinder et al. [15], Felson et al. [16] and Bierma-Zeinstra et al. [17], we compiled a list of criteria to evaluate the definitions of hip OA (Appendix 1). These criteria relate to the reliability, the validity and the applicability of the definition of hip OA:

1. The reliability of the definition expressed in intra- and inter-rater reliability
2. The validity of the definition expressed in
    Criterion validity
    –   expert validity
    The expert validity evaluates the sensitivity and specificity of the classification criteria with the use of a predefined "gold standard" by expert's opinion in a trans-sectional study design [15, 16].

– predictive validity

The predictive validity evaluates the sensitivity and specificity of the classification criteria with the use of a predefined "gold standard" by an "obvious hip OA" (for example a total hip replacement) after a certain period of follow-up [15, 16].

Construct validity.

The construct validity evaluates whether the definition correlates with the external variables it should correlate with [16, 17]. In case of radiological hip OA, the definition should correlate with known symptoms (hip pain, disability, limited ROM, morning stiffness < 1 hour) of hip OA, or with known risk factors of hip OA. If the definition is based on clinical signs, it should correlate with radiological signs of hip OA.

3. The applicability of the definition of hip OA expressed in three issues, namely:
   – the ability to discriminate between hip OA and no hip OA,
   – the ability to categorise the severity of hip OA,
   – the tools and skills needed to define persons with hip OA.
4. A description of which method has been used to develop the definition of hip OA (content validity).

Two reviewers (MR and SMABZ) independently evaluated the definition of hip OA used in the included articles according to the above criteria. In case of disagreement, both reviewers tried to achieve consensus. If disagreement was not resolved, a third reviewer (BWK) was consulted to achieve a final judgement.

Data extraction

In the included studies, data on reliability (various measures of intra- and inter-rater reliability), construct validity (association measures) as well as information on the applicability of the seven definitions used for hip OA were collected by two reviewers independently of each other and summarised (descriptive analysis) according to each definition separately.

Results

Identification / selection of the literature

The initial searches resulted in 1,170 potentially relevant articles [18]. Of these, 12 articles fulfilled the predefined inclusion criteria. After screening the reference lists of the 12 articles, another two articles were included. Finally, 14 publications were used to extract data regarding the reliability, validity, or applicability of the definitions for hip OA.

Description of included studies

Of the 14 articles, 13 studied the reliability and 7 the validity of one (or more) of the definitions. Table 1 lists the characteristics of the studies. As can be seen, there is a large difference in the reported prevalence of hip OA, probably due to the large difference in the percentage of males and the different classifications of hip OA used. All studies used a relatively young population (mean age < 66 years).

The 14 studies defined hip OA according to one (or more) of the following seven definitions (Appendix 2): Kellgren & Lawrence grade = 5 studies, Croft grade = 6, MJS (according to Croft et al.) = 8, MJS (according to Resnick & Niwayama) = 1, the ACR criteria = 3, hip pain and joint space narrowing (JSN) = 1 study, and the index grade according to Lane = 2.

Results of the included studies

Reliability

Of the 14 studies, 13 investigated the reliability of 5 of the 7 definitions of hip OA (Table 2).

The four studies that investigated the reliability of the Kellgren & Lawrence grade reported an intra-rater reliability with Kappa statistics of 0.76 [19], Pearson Correlation Coefficient of 0.66–0.89 [5], an inter-rater reliability with Kappa statistics of 0.60–0.75 [12, 19], and an Intraclass Correlation Coefficient (ICC) of 0.63 [5]. In contrast to more recent studies, the original study of Kellgren & Lawrence showed a relatively lower inter-rater reliability (Correlation Coefficient of 0.40) [7].

In five studies the overall grade of Croft (a modification of the Kellgren & Lawrence grade) had an intra-rater reliability with Kappa statistics of 0.49–0.93 [5, 8, 20, 21] but a relatively lower inter-rater reliability with Kappa statistics of 0.37–0.79 [5, 8, 20]. The wide range of intra- and inter-rater reliability between the studies is mainly explained by the different cut-off levels used.

In seven studies the MJS according to Croft et al. showed the highest intra- and inter-rater reliability compared to the other definitions of hip OA. The MJS according to Croft et al. showed an intra-rater reliability with Kappa statistics of 0.81–0.85 [5, 8, 21] and an ICC of 0.83–0.94 [19, 20], an inter-rater reliability with Kappa statistics of 0.42–0.84 [5, 8, 22] and an ICC of 0.75–0.96 [19, 20]. Only the study by Hirsch et al. [5], described a relatively low inter-rater reliability with Kappa statistic of 0.42.

Only one study investigated the inter-rater reliability of the ACR classification(s) [23] and reported a wide range for the clinical set with Kappa statistics of 0.0–0.65 and the combined clinical, radiological and lab signs set with Kappa statistics of 0.31–0.85.

Two studies investigated the reliability of the index according to Lane. These studies reported an intra-rater reliability with Kappa statistics of 0.83 (≥ grade 2) and an ICC of

Table 1: Characteristics of the included studies (N = 14)

Reference	Definition used in study	Study population setting	size	male (%)	age	Qualification of observers	Radiograph	Prevalence (%)	Number of observers	Number of readings
Altman (1991)	ACR	Rheumatologic setting	227	63.7	61 (mean)	Musculo-skeletal radiologist	AP	NA		
Antoniades (2000)	Croft's grade MJS (Croft)	St. Thomas' UK Twin register	1148	0	53 (median) (range: 40-70)	Trained	Pelvic AP (supine)	8 (≥ 2) 5 (≥ 3) 8 (≤ 2.5 mm)	1	2
Bellamy (1999)	ACR	Australian Twin Register	159	28.9	58.79 (mean)	Rheumatologist	AP	NP	3	1
Bierma-Zeinstra (1999)	Kellgren & Lawrence JSN	General practice	227	27	66 (mean) (SD: 9.6)	Radiologist and 1 trained	Pelvic AP and axial (frog-leg-position)	NA*	2	1
Birrell (2001)	Croft's grade MJS (Croft)	General practice	195	32.3	63 (median) (inter-quartile range: 54-71)	NP	Pelvic AP	NA*	2	2
Croft (1990)	Croft's grade MJS (Croft)	Intravenous urogram	1315	100	60-75 (range)	Trained	Intravenous urogram	24.3 (≥ 2) 11.0 (≥ 3) 14.4 (≤ 2.5) 2.0 (≤ 1.5)	2	2
Hirsch (1998)	Kellgren & Lawrence MJS (Croft)	Population based	755	37.9	58.3 (mean)	Trained	Pelvic AP	3.6 10.8 (<2.5)	2	2
Ingvarsson (2000)	Kellgren & Lawrence MJS (Croft)	Colon radiographs	1501	42.5	Not given	NP	Colon AP (supine)	9.2 10.0 (≤ 2.5)	3	2
Kellgren (1957)	Kellgren & Lawrence	Population based	85	NP	55-64 (range)	Trained	NP	NP	2	2

Reference	Definition used in study	Study population (setting)	size	male (%)	age	Qualification of observers	Radiograph	Prevalence (%)	Number of observers	Number of readings
Lane (1993)	Index according to Lane	Study of Osteoporotic Fractures	31	0	70 (median) (range: 65-86)	Trained	Pelvic AP	55.5 (right hip ≥ 2) 36.1 (left hip ≥ 2)	3	2
MacGregor (2000)	Croft's grade MJS (Croft)	St. Thomas UK Adult Twin register	824	0	57.4 (mean) (range: 50-72)	Trained	Pelvic AP (supine)	16.3 (≥ 1) 11.7 (≥ 2) 10.9	1	2
Nevitt (1995)	Index according to Lane	Study of Osteoporotic Fractures	4090	0	71.1 (mean) (SD 4.9)	Trained	Pelvic AP	7.1 (≥ 2) 4.7 (≥ 3)	3	1
Smith (1995)	Croft's grade MJS (Croft)	Intravenous urogram	203	0	60-75 (range)	NP	Intravenous urogram	NP	2	2
Sun (1997)	Kellgren & Lawrence Croft's grade MJS (Croft)	Systematic review	-	-	-	-	-	-		

* population includes patients with hip pain who consulted the general practitioner

NA = not applicable

NP = not provided

AP = anterior posterior radiograph

27

Table 2: Reliability of the definitions of hip osteoarthritis

	Intra-rater	Inter-rater	Cut-off level	Statistic	Size	Prevalence (%)	Reference
Kellgren & Lawrence (0–4)		0.75	≥ 2	Kappa (dichotomous)	227	NA	Bierma-Zeinstra (1999)
	0.66 / 0.89	0.63 (0.59-0.67)	-	Intra: Pearson Inter: ICC	775	3.6	Hirsch (1998)
	0.76	0.60 / 0.65	≥ 2	Kappa (dichotomous)	147	9.2	Ingvarsson (2000)
	0.75	0.40	-	Correlation coefficient	85	NP	Kellgren (1957)
Croft Grade (0–5)	0.81	0.79	NP	Kappa		NA	Birrell (2001)
	0.49 0.93	0.41 0.63	≥ 3 ≥ 4	Kappa (dichotomous)	50	24.3 (≥2) 11.0 (≥3)	Croft (1990)
	0.61	0.37	NP	Kappa			Hirsch (1998)
	081		NP	Kappa	350	16.3 (≥1) 11.7 (≥2)	MacGregor (2000)
	0.61	0.37	≥ 3	Kappa	40	NP	Smith (1995)
MJS According to Croft		0.81 / 0.84	NP	Kappa	350	8[a]	Antoniades (2000)
	0.83	0.75	-	ICC	195	NA	Birrell (2001)
	0.81	0.70	≤ 2.5 mm	Kappa (dichotomous)	50	14.4	Croft (1990)
	0.83	0.79	≤ 1.5 mm			2.0	
	0.85	0.42	NP	Kappa			Hirsch (1998)
	0.94	0.81 / 0.96	-	ICC	147	10	Ingvarsson (2000)
	0.84		NP	Kappa	350	10.9	MacGregor (2000)
	0.85	0.42	≤ 2.5 mm	Kappa	40	NP	Smith (1995)
ACR criteria **– Clinical**		0 / 0.65 (0.93/0.95 [b])		Kappa (dichotomous)	159	NP	Bellamy (1999)
– Clin/rad/lab		0.31 / 0.85 (0.91/0.97 [b])		Kappa (dichotomous)	159	NP	Bellamy (1999)
Index according to Lane	0.70/ 0.88 0.36 0.83	0.76 / 0.85 0.40 / 0.47 0.72 / 0.92	0-3 ≥ 1 ≥ 2	ICC Kappa Kappa	31	55.5 (right hip ≥ 2) 36.1 (left hip ≥ 2)	Lane (1993)
		0.87 0.71	0-4 ≥ 2	ICC Kappa	4090	7.1 (≥ 2) 4.7 (≥ 3)	Nevitt (1995)

[a] < 2.5 mm used as cut-off level

[b] Kappa adjusted for prevalence and bias (24)

NP = not provided

NA = not applicable (population includes patients with hip pain who consulted the general practitioner)

0.70–0.88[13], an inter-rater reliability with Kappa statistics of 0.72–0.92 (≥ grade 2) and an ICC of 0.76–0.87 [13, 14].

Validity

None of the screened studies investigated the criterion (expert or predictive) validity of the seven definitions of hip OA. In the 14 studies the construct validity was evaluated by considering two questions: Does the radiological definition correlate with known symptoms of hip OA? and Does the definition correlate with other definitions of hip OA? Of the 14 studies, 7 evaluated the construct validity of two of the definitions of hip OA (MJS and the overall grade of Croft) (Table 3). The association between the radiological definition and (known) symptoms of hip OA (hip pain, restricted ROM) was used as a measure of construct validity. The highest association was described between severe radiological hip OA and hip pain, and between severe radiological hip OA and a restricted internal rotation of the hip [8, 20]. In their study, Birrell et al. [20] investigated the association between restricted ROM and mild to moderate radiological hip OA defined as grade ≥ 2 (Croft grade) and severe OA defined as MJS ≤ 1.5 mm. Internal rotation appeared to be the most discriminating movement for severe hip OA (OR of 46.8 (95% CI: 5.2–420.0) versus 3.6 (95% CI: 1.6–8.0) for moderate OA). In 1990 Croft et al. investigated the association between hip pain and radiologic hip OA [8]. Severe hip OA defined by MJS ≤ 1.5 mm, showed a stronger association with hip pain than defined by the Croft grade (prevalence of 56.0% versus 47.5% of those with hip pain). The association with pain and MJS ≤ 2.5 or Croft grade ≥ 3 is comparable (prevalence of 28.3% versus 28.8% of those with hip pain).

For the construct validity we also reported the correlation between the different definitions of hip OA. The relationship between the Kellgren & Lawrence definition and the three sets of ACR criteria is very low (Kappa of 0.03–0.16) [12]. There is a moderate agreement between the definition of Kellgren & Lawrence and "hip pain and JSN" (Kappa of 0.52) [12]. There was a high association between a severe hip OA defined by MJS ≤ 1.5 mm and grade

Table 3: Association of definitions with known symptoms of hip osteoarthritis

	MJS (Croft)	Croft grade			
	≤ 2.5 mm	≤ 1.5 mm	≥ 2	≥ 3	≥ 4
Restriction ROM (OR)					
– Flexion (≤ 94°)		2.6 (0.8-8.9)	1.5 (0.7-3.2)		
– External rotation (≤ 23°)		1.2 (0.3-3.9)	3.0 (1.4-6.2)		
– Internal rotation (≤ 23°)		46.8 (5.2-420.0)	3.6 (1.6-8.0)		
Prevalence of pain	28.3%	56.0%		28.8%	47.5%

References: 8, 20
OR = Odds Ratio (95% Confidence Interval)

≥ 4 (Croft grade) (OR of 153.5) [8]. None of the studies compared the association between two of more definitions with known risk factors.

The method of development of the seven definitions of hip OA also differs considerably. The Kellgren & Lawrence grade and the index according to Lane were developed based on the opinion of the researchers. The overall grade of Croft and the MJS were based on a study population, and were developed based on pain within the study population. The ACR criteria sets were also based on a study population, and were developed using regression analysis (classification tree) on the occurrence of hip OA defined by an expert team. The methods of development of the remaining two definitions were not given.

Applicability

The applicability of the definitions of hip OA in the present study was made operational as the ability to discriminate between hip OA and no hip OA, the ability to categorise the severity of hip OA, and the skills and tools needed to classify persons according to the respective definitions (Table 4). According to their own description, six definitions intend to discriminate between persons with and without hip OA, and all six are easy to apply for persons at MD level. The Kellgren & Lawrence grade, Croft grade, the MJS and the index grade according to Lane are also able to categorise the severity of hip OA.

All definitions include information from a radiograph (except the clinical set of the ACR criteria). The ACR also makes use of information of the clinical history and physical examination (restricted ROM).

Discussion

Reviewing the selected literature demonstrates that particularly the validity of the various definitions of hip OA has barely been investigated. The highest (intra- and inter-rater) reliability was reported for the MJS and the index according to Lane and the highest association with hip pain compared to the other definitions of hip OA for the MJS.

Despite putting much effort into identifying all relevant articles, some relevant articles may have been missed because e.g. they used other keywords, had unclear abstracts, or were not indexed in Pubmed or Embase. Although the sensitivity of our search action might not be optimal [24], [25], we nevertheless believe that we included the most appropriate studies that evaluated aspects of the quality of definition of hip OA, and assume that the data presented here gives a clear insight in the currently available studies on this topic. Only 14 of 1170 potentially relevant articles fulfilled the predefined inclusion criteria. The most restrictive inclusion criterion was, that the study population contained persons with and persons without hip OA.

The problems encountered when comparing the results of the included studies, were the differences in study populations (percentage of males), settings (open population, patients with hip pain who consulted their GP), different cut-off points for case definitions, and the different or not transparent statistics used in the studies. For example, the percentage of males in the different studies ranged from 0–100%; because gender is a known risk factor for hip OA this will obviously influence the prevalence of hip OA. The prevalence, in turn, will also affect the value of reliability [26]. One study [23] adjusted the Kappa (Cohen) they found for prevalence (Prevalence Adjusted Bias Adjusted Kappa / PABAK [26]); the adjusted Kappa was much higher than the crude Kappa.

In the absence of a gold standard for a definition of hip OA, we were particularly careful when evaluating the validity. Two potential solutions to define a "gold standard", by expert's opinion or by an "obvious hip OA" (such as total hip replacement) after a certain period of follow-up were not used in the screened studies. Summarising the available information, it was clear that very few studies investigated the construct validity of the definitions used for hip OA. Of the 14 articles, not one focused on the relationship between risk factors and radiologic hip OA, leaving us to evaluate the studies that reported the association between symptoms and radiological hip OA. Croft et al. [8] investigated the association between hip pain and radiological hip OA (2 definitions of Croft); in their study population of 1315 men, only 759 completed the questionnaire (243 men died, 152 men were too ill according to the GP). The men excluded were probably older, more disabled and had more co-morbidity compared to the men included, which may have led to a selection bias; the results of that study should therefore be interpreted with caution. Croft et al. [8] also investigated the association between individual radiological features and hip pain; they concluded that MJS ($\leq$ 1.5 mm) showed a stronger association with hip pain than osteophytes (56% versus 34.4%). Surprisingly, no articles were found that investigated the association between the overall Kellgren & Lawrence grade and hip pain. The validity of 3 sets of criteria of the ACR was investigated in only one study [12], which concluded that the clinical ACR criteria showed no cross-validity (agreement between 3 ACR criteria sets) with the two other ACR criteria sets, tested in primary care.

For reliability, the lack of comparability between the different studies is also an important confounder. Different standardisation of the X-rays between studies, or a possible difference in mean joint space between men and women, can influence the results of the reliability. Only one study [19] directly compared the reliability of the Kellgren & Lawrence grade with MJS (according to Croft); the MJS showed a better (intra- and inter-rater) reliability. Five studies [5, 8, 20, 21, 27] directly compared the overall grade of Croft and the MJS; all these studies showed a better reliability of the MJS. No studies compared the other definitions. Only three studies reported the time interval between repeated readings: Croft et al.: 3–5 months [8], Kellgren 1 month [7] and Lane et al. 1 month [13]. The length of this interval will probably influence the reliability (a longer time interval between repeat readings will give

a lower intra-rater reliability) [4]. Overall, we assume that the MJS and the index according to Lane definition for hip OA have the highest reliability for epidemiological and clinical studies.

The most commonly used definition of hip OA, the Kellgren & Lawrence grade, is also the most criticised one. Previous criticisms on the Kellgren & Lawrence grade include: inconsistencies in the description of radiographic features of OA [28–30], the prominence awarded to the osteophytes at all joint sites [1, 30], and a poor inter-rater and between-center reliability [1, 28–30]. According to the articles included in our review, the inter-rater reliability was poor only in the original study of Kellgren & Lawrence [7], but much better in 3 other much larger studies [5, 12, 19]. Notably, the same description of the Kellgren & Lawrence grade was used in all studies. Therefore in the present study we could not confirm the criticism of inconsistent grades and poor reliability of the Kellgren & Lawrence grade. The main criticism of the Kellgren & Lawrence grade is the importance of the presence of osteophytes. Although it is well known that the association between osteophytes and hip pain is poor [8], not one of the 14 articles investigated the association between the overall Kellgren & Lawrence grade and hip pain. Overall, we assume that the Kellgren & Lawrence grade for hip OA is a useful definition for epidemiological studies.

Summarising the properties of the definitions used for hip OA investigated in the present study, we conclude that:

1. The MJS showed a good intra- and inter-rater reliability, a good association with hip pain and restricted internal rotation, and a good applicability; however, the quality (validity, reliability) of this definition should be investigated in an open population.

2. The Kellgren & Lawrence grade has a reliability comparable to MJS, but the construct validity should be investigated more thoroughly.

3. The Croft grade appeared to be inferior to the MJS and the Kellgren & Lawrence grade for both reliability and validity.

4. The ACR criteria (despite their precise and extensive method of development) showed a poor reliability and a poor cross-validity in a primary care setting. Because these data are based on the results of only two studies, more research is needed on the ACR criteria (also in other settings).

5. The index according to Lane showed also a good intra- and inter-rater reliability, but no studies were included which investigated the construct validity of this index grading system.

Considering how frequently the definitions of hip OA are used, it is surprising that the validity has been so poorly investigated. Meanwhile, because of the lack of such validity studies, we recommend that only those definitions with the best construct validity and the best reliability be used in epidemiological studies. We also recommend that the validity, especially the expert or predictive validity, of the commonly used definitions be studied more thoroughly.

Acknowledgments

References

1. Croft P, Cooper C, Coggon D. Case definition of hip osteoarthritis in epidemiologic studies. J Rheumatol 1994;21(4):591–2.
2. Dougados M, et al. Radiological progression of hip osteoarthritis: definition, risk factors and correlations with clinical status. Ann Rheum Dis 1996;55: 356–62.
3. Birrell F, Croft P, Cooper C, Hosie G, Mac Farlane G, Silman A. Radiographic change is common in new presenters in primary care with hip pain. Rheumatology 2000;39:772–5.
4. Sun Y, Gunther KP, Brenner H. Reliability of radiographic grading of osteoarthritis of the hip and knee. Scan J Rheumatol 1997;26(3):155–65.
5. Hirsch R, et al. Hip osteoarthritis prevalence estimates by three radiographic scoring systems. Arthritis Rheum 1998;41(2):361–8.
6. Gunther KP, Sun Y. Reliability of radiographic assessment in hip and knee osteoarthritis. Osteoarthritis Cartilage 1999;7(2):239–46.
7. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494–502.
8. Croft P, Cooper C, Wickham C, Coggon D. Defining osteoarthritis of the hip for epidemiologic studies. Am J Epidemiol 1990;132(3):514–22.
9. Resnick D, Niwayama G. Degenerative disease of extraspinal locations. Diagnosis of Bone and Joint Disorders. 2nd ed. Philadelphia: WB Saunders; 1988:1426–42.
10. Altman RD. The American College of Rheumatology Criteria for the Classification and reporting of osteoarthritis of the hip. Arthritis Rheum 1991;34(5):505–14.
11. Odding E, Valkenburg HA, Algra D, VandenOuweland FA, Grobbee DE, Hofman A. Associations of radiological osteoarthritis of the hip and knee with locomotor disability in the Rotterdam Study. Ann Rheum Dis 1998;57(4):203–8.
12. Bierma-Zeinstra SMA, Bohnen AM, Ginai AZ, Prins A, Verhaar JAN. Validity of American College of Rheumatology criteria for diagnosing hip osteoarthritis in primary care research. J Rheumatol 1999;26(5):1129–33.
13. Lane NE, Nevitt MC, Genant HK, Hochberg MC. Reliability of new indices of radiographic osteoarthritis of the hand and hip and lumbar disc degeneration. J Rheumatol 1993;20:1911–8.
14. Nevitt MC, et al. Radiographic osteoarthritis of the hip and bone mineral density. Arthritis Rheum 1995;38(7):907–16.
15. Buchbinder R, Goel V, Bombardier C, Hogg-Johnson S. Classification systems of soft tissue disorders of the neck and upper limb: do they satisfy methodological guidelines? J Clin Epidemiol 1996;49(2):141–9.
16. Felson DT, Anderson JJ. Methodological and statistical approaches to criteria development in rheumatic diseases. Baillieres Clin Rheumatol 1995;9(2):253–66.
17. Bierma-Zeinstra SMA, Verhagen AP, Berger MY. Het ontwikkelen van classificatiesystemen bij aandoeningen van het bewegingsapparaat. Huisarts en wetenschap 2000;43(11):473–7.
18. Haynes RB, Wilczynski N, McKibbon A, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. J Am Med Informatics Assoc 1994;1(6):447–58.
19. Ingvarsson T, Hagglund G, Lindberg H, Lohmander LS. Assessment of primary hip osteoarthritis: Comparison of radiographic methods using colon radiographs. Ann Rheum Dis 2000;59(8):650–3.
20. Birrell F, Croft P, Cooper C, Hosie G, MacFarlane G, Silman A. Predicting radiographic hip osteoarthritis from range of movement. Rheumatology 2001;40(5):506–12.

21.   MacGregor AJ, Antoniades L, Matson M, Andrew T, Spector TD. The genetic contribution to radiographic hip osteoarthritis in women: Results of a classic twin study. Arthritis Rheum 2000;43(11):2410–6.

22.   Antoniades L, MacGregor AJ, Matson M, Spector TD. A cotwin control study of the relationship between hip osteoarthritis and bone mineral density. Arthritis Rheum 2000;43(7):1450–5.

23.   Bellamy N, et al. Perceptual variation in categorizing individuals according to American College of Rheumatology classification criteria for hand, knee, and hip osteoarthritis (OA): Observations based on an Australian Twin Registry study of OA. J Rheumatol 1999;26(12):2654–8.

24.   Dickersin K, Scherer R, Lefebre C. Systematic reviews: Identifying relevant studies for systematic reviews. BMJ 1994;309(12):1286–91.

25.   Dickersin K. Systematic reviews in epidemiology: why are we so far behind? Int J Epidemiol 2002;31:6–12.

26.   Byrt T, Bishop J, Carlin JB. Bias, Prevalence and Kappa. J Clin Epidemiol 1993;46(5):423–9.

27.   Smith RW, Egger P, Coggon D, Cawley MID, Cooper C. Osteoarthritis of the hip joint and acetabular dysplasia in women. Ann Rheum Dis 1995;54:179–81.

28.   Hart DJ, Spector TD. The classification and assessment of osteoarthritis. Baillieres Clin Rheumatol 1995;9(2):407–33.

29.   Hart DJ, Spector TD. Radiographic criteria for epidemiologic studies of osteoarthritis. J Rheumatol 1995; suppl 43(22):46–8.

30.   Spector TD, Cooper C. Radiographic assessment of osteoarthritis in population studies: whither Kellgren and Lawrence? Osteoarthritis Cartilage 1993;1:203–6.

Appendix 1: Criteria used in the present study to evaluate the definitions of hip osteo-arthritis used in the literature

Reliability

1. Does the definition provide consistent results when classifying the same conditions (e.g. split-half reliability)?
   Positive if the results are comparable when tested in the same setting, but in a new group (e.g. split-half reliability).

2. Is the intraobserver reliability described?
   – Results individual variables/features: Kappa or ICC or Pearson Product Moment Correlation Coefficient (range, CI)
   – Results case definition hip OA: Kappa or ICC or Pearson Product Moment Correlation Coefficient (range, CI)
   – Are the results specified for experienced observer, specialisation?

3. Is the interobserver reliability described?
   – Results individual variables/features: Kappa or ICC or Pearson Product Moment Correlation Coefficient (range, CI)
   – Results case definition hip OA: Kappa or ICC or Pearson Product Moment Correlation Coefficient (range, CI)
   – Are the results specified for experienced observer, specialisation?

Criterion validity

4. Did the study investigate the validity of the definition with a predefined "gold standard" by expert's opinion (expert validity), in a cross-sectional study design?
5. Did the study investigate the validity of the definition with a predefined "gold standard" by an "obvious hip OA" (for example a THR) after a certain period of follow-up (predictive validity), in a longitudinal study design?

Construct validity

6. Does the definition discriminate between entities that are thought to be different in a way appropriate for the purpose?
   – category is related to a different intervention, **or**
   – category is related to a different prognosis, **or**
   – category has a different underlying etiological process

7. Do the definition show adequate associations with known risk factors of hip OA?
   Positive if the definition showed an equal (positive) or higher association than the other definition of hip OA.
   Risk factors of hip OA:
   – Genetics
   – Bone Mineral density
   – Biomechanical workload
   – Sport activities
   – Acetabular dysplasia

8. Do the definition show adequate associations with other symptoms (or signs) of hip OA, than included in the definition?
   Symptoms of hip OA:
   – hip pain
   – limited physical function of the lower limb
   – limited ROM of the hip joint
   – morning stiffness
   For the radiological definitions 1–4 and 7:
   Positive if the definition performs a positive association with pain (of the hip) **and/or** limited physical function of the lower limb **and/or** limited ROM of the hip joint **and/or** morning stiffness
   For definitions 5 and 6:
   Positive if the definition performs a positive association with limited physical function of the lower limb **and/or** limited ROM of the hip joint **and/or** morning stiffness.
   For definition 5, the clinical definition of the ACR criteria of hip OA:
   Positive if the definition performs a positive association with radiological symptoms (joint space narrowing, osteophytes of femur head, cysts, subchondral sclerosis, and migration of the femur head) of hip OA.

Content validity

9. Is the method of development of the definition clearly specified?
   Which method is used?
   – Informal: opinion of researcher
   – Informal: opinion of (international) "experts"
   – Formal: the classification is based on a study population, and frequencies of symptoms are given
   – Formal: construction of the groups (classification) with help of clinical endpoints (effect of intervention or known progression)

– Mathematical method: cluster analysis, factor analysis, split-half analysis, classification tree (regression analysis)

– Other method, …

Applicability

10. Is the definition easy to perform (for persons at MD level) without special training? Positive if no special training for persons with MD level (specific skills needed) is required.

11. Which tests are necessary to perform the definition?
    – clinical history
    – physical examination/measurements (ROM)
    – radiographs
    – lab/blood samples

Appendix 2: Definitions of hip osteoarthritis used in the literature

1. Kellgren & Lawrence grading system (7)

Grade		Description
0	No osteoarthritis	
1	Doubtful	Possible narrowing of joint space medially and possible osteophytes around femoral head; or osteophytes alone
2	Mild	Definite narrowing of joint space inferiorly, definite osteophytes and slight sclerosis
3	Moderate	Marked narrowing of joint space, definite osteophytes, some sclerosis and cyst formation and deformity of femoral head and acetabulum
4	Severe	Gross loss of joint space with sclerosis and cysts, marked deformity of femoral head and acetabulum and large osteophytes

Hips classified as grade 2 or higher were defined as having OA

2. Croft's modification of the Kellgren & Lawrence grading system ("Croft grade") (8)

Grade	Description
0	No change
1	Definite osteophytes only
2	Joint space narrowing (JSN) only (defined as an MJS of ≤ 2.5 mm)
3	Presence of 2 of the following: JSN, osteophytosis, subchondral sclerosis (of ≥5 mm), and cyst formation
4	Presence of 3 of the following: JSN, osteophytosis, subchondral sclerosis (of ≥5 mm), and cyst formation
5	Same as grade 4, but with deformity of the femoral head or total hip replacement due to OA (verified by record view)

3. **Croft's measurement of the "minimal joint space" (lateral, superior, axial, medial) (8)**

Minimal joint space (MJS) is the shortest distance on the radiograph between the femoral head margin and the acetabular edge.

Grade	Description
0	MJS > 2.5 mm
1	MJS > 1.5 mm and ≤ 2.5 mm
2	MJS ≤ 1.5 mm

4. **Resnick and Niwayama measurement of the joint space (superior, axial and medial) (9)**

Grade	Description
0	MJS > 3.5 mm
1	MJS ≤ 3.5 mm

5. **ACR criteria (10)**

ACR 1 Clinical criteria (Classification tree format)	ACR 2 Combined Clinical and Radiographic Criteria (Traditional format)	ACR 3 Combined Clinical and Radiographic Criteria (Classification tree format)
Hip pain	Hip pain	Hip pain
+ Hip internal rotation < 15° and ESR ≤ 45 mm/h (if ESR not available, hip flexion ≤ 115°) OR	+ At least 2 of the following 3 features:	+ Radiographic femoral and/or acetabular osteophytes
		OR
+ Hip internal rotation ≥ 15° and pain on internal rotation and morning stiffness of the hip ≤ 60 min and age > 50 years	− ESR < 20 mm/h − Radiographic femoral or acetabular osteophytes − Radiographic joint space narrowing (superior, axial and/or medial)	+ ESR ≤ 20 mm/h and radiographic axial joint space narrowing

ESR: one-hour erythrocyte sedimentation rate.

6. **Clinical osteoarthritis of the hip: positive radiological osteoarthritis combined with pain in the hip region (19)**

Hip pain with joint space narrowing (JSN):
– Superior JSN < 3.5 mm **and/or**
– Axial JSN <2.5 mm

7. Radiographic index grade according to Lane

Grade	Description
0	Normal (no findings of OA)
1	Possible osteophytes (IRF grade 1) and / or narrowing (IRF grade 1), or isolated definite osteophytes or narrowing (IRF grade $\geq$ 2)
2	Definite osteophytes or narrowing (IRF grade $\geq$ 2) plus cysts or sclerosis
3	3 of the following: definite osteophytes or narrowing (IRF grade $\geq$ 2), cysts or sclerosis
4	Grade 3 (as above) plus femoral head deformity

Number of IRF (Individual Radiographic Features) scores that correspond to 0 = normal, 1 = mild, 2 = moderate, 3 = severe.

3

Validity and reliability of three definitions of hip osteoarthritis: cross-sectional and longitudinal approach
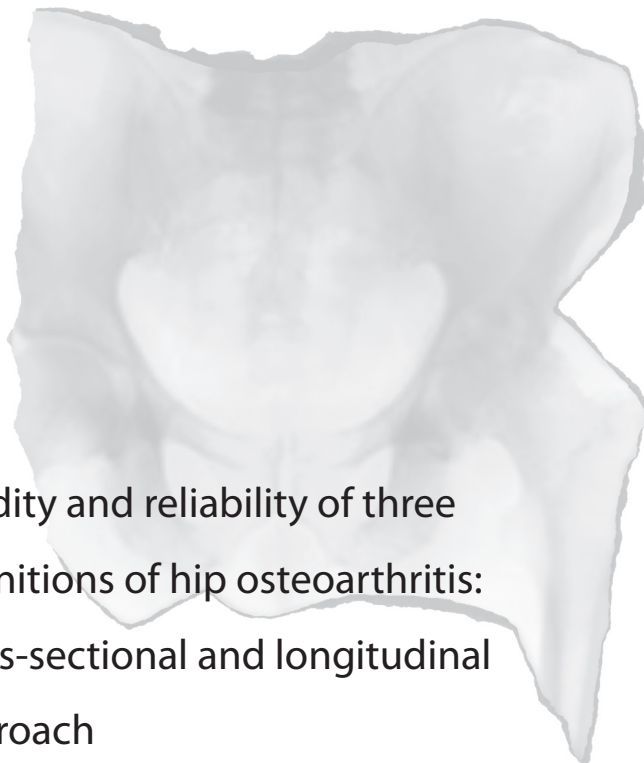
Abstract

Objectives: To compare the reliability and validity in a large open population of three frequently used radiological definitions of hip osteoarthritis (OA) namely, Kellgren & Lawrence grade, Minimal Joint Space (MJS) and Croft's grade. Additionally, to investigate whether the validity of the three definitions of hip OA is gender dependent.

Methods: From subjects of the Rotterdam study (elderly aged 55 years and over, N = 3,585 participants) all X rays were evaluated. The inter-rater reliability was tested in a random set of 148 X-rays. The validity was expressed as the ability to identify patients who show clinical symptoms of hip OA (construct validity) and as the ability to predict Total Hip Replacement (THR) at follow-up (predictive validity).

Results: The inter-rater reliability was similar for the Kellgren & Lawrence grade and MJS (Kappa statistics of 0.68 and 0.62, respectively) but somewhat lower for Croft's grade (Kappa statistics of 0.51). The Kellgren & Lawrence grade and MJS demonstrated both the strongest associations with clinical symptoms of hip OA.

Gender appears to be a significant effect modifier for Kellgren & Lawrence; women had a significantly stronger association with symptoms than men, and also for MJS; however, this gender dependency was attributed to differences in height between women and men. The Kellgren & Lawrence grade showed the highest predictive value for THR at follow-up compared to the other definitions.

Conclusion: Based on these findings, Kellgren & Lawrence still appears to be a useful definition for hip OA for epidemiological studies focusing on the presence of hip OA.

Introduction

Osteoarthritis (OA) of the hip is of particular interest since it is often the sole joint affected by OA suggesting an important role of local biomechanical risk factors. In addition, the prevalence of hip OA is expected to increase with the aging of the Western society (1) and the hip is crucial to independent function (2).

A problem in studying hip OA is the absence of consensus in defining hip OA for epidemiological research (3). To investigate occurrence and (potential) risk factors, a valid and reliable definition of hip OA is required. Most epidemiological studies have used a single hallmark of hip OA (namely radiological signs) to define hip OA (4, 5).

In a previous systematic appraisal, we summarised the validity, reliability and applicability of seven definitions of hip OA used in epidemiological studies (6). Considering the frequent use of the definitions of hip OA, it is noteworthy that the validity of these definitions has been so poorly investigated. Because of lack of comparability between the different studies and because most studies only investigated a single definition, it was difficult to compare the reliability and validity of different definitions of hip OA. Our appraisal also showed that the validity and reliability of Minimal Joint Space (MJS; according to Croft) and Croft's grade (a modification of Kellgren & Lawrence) (7) have only been studied in a male population.

The primary objective of the present study was to compare the reliability and validity of three most frequently used radiological definitions of hip OA, Kellgren & Lawrence grade, MJS (according to Croft) and Croft's grade, in a large open population of elderly people. The secondary objective was to investigate whether the validity of the three definitions of hip OA was gender dependent.

Subjects and Methods

The study population consisted of participants of the Rotterdam Study, a prospective cohort of men and women aged 55 years and over. The objective of the Rotterdam Study is to investigate the incidence of, and risk factors for, chronic disabling diseases. The rationale and study design have been described previously (8). The focus is on neurogeriatric, cardiovascular, ophthalmologic and locomotor diseases. All 10,275 inhabitants of Ommoord (a district in Rotterdam, the Netherlands) were invited to participate. The response rate was 78%, resulting in 7,983 subjects participating in the study. Written informed consent was obtained from each participant. The Medical Ethics Committee of the Erasmus University Medical Centre has approved the Rotterdam Study.

For the present study a sample of 3,585 subjects of the Rotterdam study was used. The selection was based on the availability of the radiographs of the hip at baseline and follow-up. The fact that subjects had to be mobile enough to visit the research centre at baseline

and follow-up and survive the follow-up period caused a health selection bias in our study population. Compared to the total Rotterdam study population, the present study population was significantly younger (70.6 years versus 66.0 years), had a lower prevalence of lower limb disability at baseline (≥ index score of 0.5: 35.5% versus 12.9%) and a somewhat lower prevalence of hip pain at baseline (12.7% versus 11.7%).

Subjects with bilateral Total Hip Replacement (THR) at baseline (N = 24) were excluded from analysis, which resulted in a study population of 3561 subjects. The baseline measurements were conducted between April 1990 and July 1993, and the follow-up measurements between 1996 and 1999, with a mean follow-up time of 6.6 years (standard deviation of ± 0.50).

Radiographic assessment

Weight bearing anteroposterior pelvic radiographs with both feet in 10° internal rotation were obtained at 70 KV, a focus of 1.8, and a focus to film distance of 120 cm, applying a Fuji High Resolution G 35 × 43 cm film (9). The X-ray beam was centred on the umbilicus. One independent trained reader (MR) evaluated the radiographs according to a standardized protocol, unaware of the clinical status of the patients.

At baseline radiographic osteoarthritis (ROA) of the hip was quantified by measurements of Kellgren & Lawrence grading system (atlas-based) (Appendix) (6, 10–13), Croft grading system (a modification of Kellgren & Lawrence) (Appendix) and MJS defined by Croft (6, 7, 12–14). For the Croft grading scale, we assessed the individual radiographic features of minimal joint space, presence of osteophytes, subchondral sclerosis and cysts formation. The presence of the individual radiographic features (of any grade) was examined, using an atlas of individual features (12, 13). Different cut off points to quantify hip ROA were used; for Kellgren & Lawrence ≥ grade 2 (moderate) and ≥ grade 3 (severe), for Croft grading system ≥ grade 3 (moderate) and ≥ grade 4 (severe), and for MJS ≤ 2.5 mm (moderate), ≤ 2.0 mm (intermediate) and ≤ 1.5 mm (severe).

The joint space width (lateral, superior, axial, medial and minimal) measurements were standardised using a 0.5-millimetre graduated magnifying glass laid directly over the radiograph (15–17).

The follow-up radiographs were evaluated for the presence of an incident THR (not present at baseline).

For all three grading systems and all measurements, inter-rater reliability (SMABZ and MR) was tested in a random set of 148 radiographs (18, 19).

45

Clinical assessment

At baseline, trained interviewers performed an extensive home interview on demographic characteristics, medical history, risk factors for chronic diseases and medication use.

For the present study we used information on the presence of hip pain ("did you have joint complaints of your right/left hip during the last month"), presence of morning stiffness and lower limb disability. Lower limb disability was assessed using a modified version of the Stanford Health Assessment Questionnaire (9). A lower limb disability index (LDI) was obtained by calculating the mean score of answers to the following six questions: 'Are you able to stand up from a straight chair without using your arms for support?', 'Are you able to get in and out of bed?', 'Are you able to walk outdoors on flat ground?', 'Are you able to climb up five steps?', 'Are you able to bend down to pick up clothing from the floor?' and 'Are you able to get in and out of a car?'. The answers were scored as follows: 0 = yes, without difficulty, 1 = yes, with some difficulty, 2 = yes, with much difficulty, 3 = no, unable to do (needs help). Moderate disability was defined as a score higher than 0.5 and severe disability as a score higher than 1.0 on the lower limb disability index. Moderate disability is present whenever there is at least some difficulty with three out of six daily activities of the LDI (9).

Statistical analysis

For the inter-rater reliability, Kappa and Intraclass Correlation Coefficient (ICC) was assessed for the different radiological individual features and the three definitions of hip ROA.

Because of the absence of a "gold standard" we expressed the validity in the construct validity and in the predictive validity. The construct validity is operationalized by the ability to identify patients with symptoms (presence of hip pain, morning stiffness or lower limb disability) of hip OA (20, 21). The predictive validity is expressed as the ability of the definition to predict important long-term outcomes of disease (21). For the construct validity, the association between baseline radiological osteoarthritis of the hip according to the three definitions and the separate baseline clinical symptoms (hip pain, morning stiffness and lower limb disability) was tested by means of Generalised Estimating Equations (GEE) (cross-sectional design). This is a procedure of repeated measurements. It is used here to take account of the correlation between the left and right hip. Additionally, the sensitivity and specificity was assessed using the main symptom of hip OA, hip pain, as "gold standard". We used different cut-off points for the three definitions of hip ROA and additionally stratified the results for gender and age. A two-sided p-value of 0.05 was considered significant. For the predictive validity, we assessed the proportion THR, after a clinically meaningful follow-up period of 6.6 years, in patients identified by each definition as having hip ROA at baseline

(longitudinal design). We also calculated the association between the different definitions of hip ROA and THR at follow-up by means of the GEE method (odds ratios). We used SPSS version 11.0 (SPSS Inc., Chicago, USA) and SAS software, version 8.0 (SAS Institute, Cary, NC) for all analyses.

Results

Study population

Table 1 shows the demographic characteristics and prevalence data on radiographic hip OA, stratified for gender of the study population of 3,585 participants. Women were older, had a higher BMI, and were shorter. The prevalence of lower limb disability and hip pain is twice as high in women than in men. Men demonstrated a higher prevalence when defined by Kellgren & Lawrence or Croft grade than women. Of the subjects with hip pain, 98.8% had longer than 1 month pain, from which 30.2% between 1 and 5 years and 51.1% longer than 5 years. Radiological hip OA defined by Croft's grade 3 showed a much higher prevalence compared with the other definitions of a moderate hip OA. The prevalence of moderate radiological hip OA defined by Kellgren & Lawrence and MJS is similar.

Table 1: Demographic characteristics and prevalences of radiographic hip osteoarthritis stratified for gender

	Men (N = 1,499)	Women (N = 2,086)	Total (N = 3,585)
Mean age, years ± SD	65.5 ± 6.5	66.3 ± 7.2	66.0 ± 6.9
Mean BMI, kg/m^2 ± SD	25.9 ± 2.8	26.6 ± 4.0	26.3 ± 3.6
Mean height, cm ± SD	175.5 ± 6.6	162.4 ± 6.5	167.9 ± 9.2
Lower limb disability (≥ 0.5), %	8.2	16.3	12.9
Lower limb disability (≥ 1.0), %	4.7	9.6	7.6
Morning stiffness, %	24.9	36.9	31.9
Hip pain (left and/or right), %	7.1	14.9	11.7
– right	5.6	11.0	8.7
– left	4.5	9.9	7.6
Total Hip Replacements at baseline,			
– unilateral, number (%)	20 (1.0)	69 (2.6)	89 (2.0)
– bilateral, number (%)	6 (0.4)	18 (0.9)	24 (0.7)
Kellgren, %			
– ≥ grade 2	7.8	6.4	7.0
– ≥ grade 3	1.2	1.5	1.4
MJS, %			
– ≤ 2.5 mm	6.8	8.1	7.5
– ≤ 2.0 mm	2.6	3.3	3.0
– ≤ 1.5 mm	1.2	1.5	1.4
Croft grade, %			
– ≥ grade 3	39.1	30.0	33.9
– ≥ grade 4	4.5	4.4	4.4

BMI = Body Mass Index
MJS = Minimal Joint Space

Reliability

Table 2 shows the inter-rater reliability for different individual radiological features and three definitions of hip ROA. The inter-rater reliability for the different individual radiological features was relatively low, with the exception of the MJS as assessed as a continuous variable. Kellgren & Lawrence ≥ grade 2 and MJS ≤ 2.5 mm had a comparable reliability, whereas for Croft's grade ≥ grade 3 the reliability was somewhat lower.

Table 2: Inter-rater reliability for individual radiological features and three definitions of radiographic hip osteoarthritis studied (N = 148)

	Subchondral Sclerosis		Osteophytes		Cysts	Minimal Joint Space (continuous)
	Acetabulum	Femoral head	Acetabulum	Femoral head		
Inter-rater reliability	**0.51** (0.35–0.67)	**0.66** (0.22–1.00)	**0.23** (0.06–0.40)	**0.34** (0.19–0.50)	**0.32** (-0.17–0.82)	**0.85*** (0.80–0.89)
	Kellgren ≥ grade 2	MJS ≤ 2.5 mm	Croft grade ≥ grade 3			
Inter-rater reliability	**0.68** (0.44–0.92)	**0.62** (0.43–0.81)	**0.51** (0.35–0.67)			

Inter-rater reliability is presented by Intraclass Correlation Coefficient (two-way mixed effect model, consistency definition) for Minimal Joint Space (as continuous variable) and by Kappa for other individual radiological features and three definitions of hip OA studied, with 95% confidence interval between parentheses.

Construct validity

Table 3 shows the association between the three definitions of hip ROA for different cut-off points and clinical symptoms of hip OA; hip pain, morning stiffness and lower limb disability (moderate and severe). The percentages of subjects defined by these definitions according to the different cut-off points are shown in Table 1. Table 3 shows that severe hip ROA has a stronger association with symptoms than moderate hip ROA. The Kellgren & Lawrence grade and MJS demonstrate comparable associations with clinical symptoms of hip OA, especially with hip pain and lower limb disability for both moderate and severe hip ROA. Croft's grade shows the weakest associations with clinical symptoms of hip OA.

Gender as effect modificator

We found that men had on average a larger joint space width than women (4.2 versus 3.9 mm, respectively). Furthermore we found that height was positively correlated with the joint space width. Additionally, we also found a positive correlation within gender between height and the joint space width (respectively a beta of 0.16 for men and 0.14 for women). Because in the present study women were shorter than men, we adjusted for height. After adjustment for height, the gender effect disappeared.

Table 3: Association between different definitions of radiographic hip osteoarthritis and clinical symptoms of hip osteoarthritis (N = 3,561)

	Hip pain		Morning stiffness	Disability (LLD ≥ 0.5)	Disability (LLD ≥ 1.0)
	OR	Sensitivity / Specificity (%)	OR	OR	OR
Kellgren					
− ≥ grade 2	**2.6** (1.8–3.6)	20.7 / 92.7	**1.2** (0.9–1.6)	**2.4** (1.7–3.4)	**3.0** (2.0–4.4)
− ≥ grade 3	**6.6** (3.6–12.1)	12.5 / 97.9	**2.2** (1.2–3.9)	**7.0** (3.8–12.8)	**7.6** (4.0–14.6)
MJS					
− ≤ 2.5 mm	**2.4** (1.7–3.4)	14.9 / 93.3	**1.6** (1.2–2.1)	**2.7** (2.0–3.7)	**3.0** (2.0–4.4)
− ≤ 2.0 mm	**4.5** (2.9–7.0)	9.3 / 97.8	**1.7** (1.2–2.6)	**3.7** (2.4–5.9)	**4.1** (2.5–7.0)
− ≤ 1.5 mm	**6.6** (3.6–12.2)	5.5 / 99.1	**2.0** (1.1–3.7)	**5.3** (2.9–9.8)	**6.1** (3.1–12.1)
Croft grade					
− ≥ grade 3	**1.3** (1.1–1.7)	39.9 / 66.9	**0.9** (0.7–1.0)	**1.0** (0.8–1.3)	**1.1** (0.8–1.5)
− ≥ grade 4	**3.6** (2.4–5.2)	11.7 / 96.4	**1.6** (1.1–2.2)	**3.3** (2.3–4.9)	**3.7** (2.3–5.8)

Associations are presented by odds ratios with 95% confidence interval between parentheses.
LLD: Lower limb disability index, a score of ≥ 0.5 was defined as moderate disabled, and a score of ≥ 1.0 as severe disabled.
Associations are adjusted for body mass index and radiographic osteoarthritis of the other hip.

Table 4: Association between three different definitions of radiographic hip osteoarthritis studied and clinical symptoms of hip OA, stratified for gender (N = 3,561)

	Hip pain		Morning stiffness		LLD ≥ 0.5		LLD ≥ 1.0	
	men	women	men	women	men	women	men	women
Kellgren								
− ≥ grade 2	**1.6*** (0.8–3.2)	**3.5** (2.3–5.2)	**1.1** (0.7–1.7)	**1.3** (0.9–1.9)	**1.5*** (0.8–3.0)	**3.3** (2.2–4.9)	**1.0*** (0.3–2.7)	**4.6** (2.9–7.3)
− ≥ grade 3	**8.7** (3.1–24.5)	**5.7** (2.7–12.0)	**2.5** (0.9–6.9)	**1.9** (0.9–4.1)	**4.4** (1.4–13.9)	**8.7** (4.1–18.6)	**3.6** (1.8–16.4)	**9.6** (4.4–20.7)
MJS								
− ≤ 2.5 mm	**2.7** (1.5–5.0)	**2.3** (1.5–3.4)	**1.6** (1.0–2.5)	**1.5** (1.1–2.2)	**2.5** (1.4–4.6)	**2.7** (1.9–4.0)	**1.2*** (0.4–3.2)	**3.8** (2.4–5.8)
− ≤ 2.0 mm	**4.4** (1.9–10.0)	**4.4** (2.6–7.5)	**1.6** (0.8–3.3)	**1.7** (1.0–2.9)	**2.2** (0.8–5.8)	**4.4** (2.6–7.4)	**1.6** (0.4–6.6)	**5.1** (2.8–9.2)
− ≤ 1.5 mm	**6.6** (2.2–19.4)	**6.6** (3.1–14.0)	**1.9** (0.7–5.1)	**2.1** (1.0–4.4)	**3.0** (0.9–10.9)	**6.5** (3.1–13.9)	**3.7** (0.8–16.8)	**7.1** (3.2–15.7)
Croft grade								
− ≥ grade 3	**1.3** (0.9–2.0)	**1.5** (1.1–2.0)	**0.8** (0.7–1.1)	**0.9** (0.8–1.1)	**0.8** (0.5–1.3)	**1.2** (0.9–1.6)	**0.8** (0.4–1.4)	**1.4** (1.0–2.0)
− ≥ grade 4	**2.9** (1.4–6.0)	**4.1** (2.5–6.5)	**1.4** (0.8–2.5)	**1.7** (1.1–2.6)	**2.7** (1.3–5.5)	**3.8** (2.4–6.1)	**1.3*** (0.4–4.4)	**5.1** (3.0–8.6)

Associations are presented by odds ratios with 95% confidence interval between parentheses.
LLD: Lower limb disability index, a score of ≥ 0.5 was defined as moderate disabled, and a score of ≥ 1.0 as severe disabled.
Associations are adjusted for body mass index and radiographic osteoarthritis of the other hip.
* significant difference between men and women

Table 5: Association between two definitions of radiographic hip osteoarthritis and hip pain, stratified for gender and age

	Men		Women		Total	
	Younger	Older	Younger	Older	Younger	Older
Median age, years	60.8	70.1	60.5	71.1	60.6	70.6
	n = 776	n = 723	n = 1,016	n = 1,070	n = 1,792	n = 1,793
Kellgren ≥ grade 2	**0.3*** (0.0–1.9)	**3.7** (1.7–7.9)	**4.0** (2.0–8.2)	**3.2** (1.9–5.3)	**1.7** (1.0–3.2)	**3.2** (2.1–4.9)
Proportion with hip pain, %	2.1	16.9	36.1	31.7	16.7	25.5
MJS ≤ 2.5 mm	**1.6** (0.5–4.7)	**4.1** (1.9–8.8)	**2.2** (1.1–4.2)	**2.3** (1.4–3.8)	**2.0** (1.2–3.5)	**2.7** (1.8–4.2)
Proportion with hip pain, %	10.8	17.9	23.6	25.5	18.5	22.7

Associations are presented as odds ratios with 95% confidence interval between parentheses.
Men and women are divided in two equal groups, a younger (1) and older (2) group, divided by median (65.2 years).
Associations are adjusted for body mass index and radiographic osteoarthritis of the other hip.
* 48 cases with radiographic hip osteoarthritis according to Kellgren & Lawrence (5), only 1 of these reported hip pain.

Table 4 shows the association between different definitions of hip ROA and clinical symptoms (of hip OA) stratified for gender. For both definitions of Croft the results showed no significant gender difference except for the association between Croft's grade ≥ grade 4 and severe lower limb disability. We found, however, that gender was a significant effect modifier for Kellgren & Lawrence grade (≥ grade 2). For women the associations between symptoms (hip pain and lower limb disability) and hip OA defined by Kellgren & Lawrence (≥ grade 2) were significantly stronger than for men. We also found that for women the association between symptoms and hip ROA according to Kellgren & Lawrence (≥ grade 2) was stronger than according to the MJS (≤ 2.5 mm) or Croft's grade (≥ grade3). Women had a higher body mass index (BMI) than men, but after we adjusted for BMI the assessed associations did not change.

Table 5 shows the association between different definitions of hip ROA (Kellgren ≥ grade 2 and MJS ≤ 2.5 mm) and hip pain stratified for gender and age (2 categories). We divided men and women in two equal groups, a younger and older group (median of 65.2 years). Older persons had a stronger association between hip ROA and hip pain than younger persons, especially when defined by Kellgren & Lawrence. The trend was that hip ROA in younger men, especially when defined by Kellgren & Lawrence, had a weaker relationship with hip pain than in women (both age categories) and older men. These results were similar for the association with lower limb disability.

Predictive validity

Table 6 shows the predictive validity of the three definitions for THR at follow-up, indicated by the association between the different definitions of hip ROA at baseline and THR at

Table 6: Predictive validity of the three definitions for total hip replacement (THR) at follow-up (N = 3,561)

	Total incident THR at follow-up	Number of THR predicted by each definition / number ROA cases defined by each definition	OR (95% CI)
Kellgren & Lawrence (≥ grade 2)			
right	**57**	**33/154** (21.4)	**30.6** (17.5–53.5)
left	**42**	**22/110** (20.0)	**34.3** (18.1–65.2)
MJS (≤ 2.5 mm)			
right	**57**	**26/151** (17.2)	**18.6** (10.7–32.3)
left	**42**	**19/123** (15.4)	**22.6** (11.8–43.0)
Croft grade (≥ grade 3)			
right	**57**	**46/673** (6.8)	**16.0** (8.0–31.8)
left	**42**	**28/717** (3.9)	**6.7** (3.4–12.9)

Number of incident THR at follow-up predicted by each definition / number ROA cases at baseline, as defined by each definition of hip ROA. ROA = radiological osteoarthritis

follow-up. The Kellgren & Lawrence grading system predicted the highest ratio of number of incident THR at follow-up divided by the number ROA cases at baseline, and showed the strongest association with THR at follow-up, compared to the other definitions.

Discussion

Based on the results of the present study that Kellgren & Lawrence showed to be the best predictor for a THR at follow-up and that MJS is height dependent, we concluded that radiological hip OA might be better defined for epidemiological studies by the Kellgren & Lawrence grading system than by MJS.

The inter-rater reliability of Kellgren & Lawrence assessed in this study is similar to that described in literature (4, 6, 22, 23). In contrast to more recent studies, the original study of Kellgren & Lawrence showed a relatively low inter-rater reliability (ICC of 0.40) (10). In the present study we found an inter-rater reliability of the MJS according to Croft, which is similar to previous studies (4, 7, 22, 24, 25). The inter-rater reliability for Croft's grade (≥ grade 3) in the present study showed a Kappa-value of 0.51 compared to Kappa statistics of 0.37–0.79 in earlier studies (4, 7, 25, 26). The wide range of inter-rater reliability between these studies is mainly explained by the different cut-off levels used. One study (14) used the same cut-off level as the present study, and reported a similar Kappa value of 0.41. However, in the original study of Croft the presented Kappa values were based on measurement of the size of the individual radiological features and not on atlas-based grades. The inter-rater reliability reported in the present study was similar for subchondral sclerosis and for osteophytes compared with the reliability reported in the study of Croft (14).

The validity of the different definitions of hip ROA has been poorly investigated in previous studies. In the present study we investigated the construct and predictive validity. Because

of the absence of a "gold standard" we expressed the predictive validity as the ability of each definition to predict a THR at follow-up. The requirement for a THR has been proposed as a potential outcome measure based on the assumption that THR is performed only in patients with a severe disease from both a symptomatic (painful and disabling disease) and a structural point of view (overall severity or advanced JSN) (27, 28). The lower limb disability assessed by the HAQ in the Rotterdam Study is not a disease specific outcome measure, but it measures arthritic conditions in general. On the other hand is lower limb disability an important symptom of hip OA, and OA is the most important cause of disability of elderly people (29). Hence we included lower limb disability, assessed by the HAQ, besides the presence of hip pain and morning stiffness as an important symptom of hip OA in the analysis. Overall, the Kellgren & Lawrence grading system showed the best predictive validity when compared with the other definitions of hip ROA and similar associations with symptoms of hip OA (construct validity), with MJS. MJS came out better concerning the construct and predictive validity than Croft's grade. The weak associations reported in the present study between Croft's grade (≥ grade 3) and symptoms of hip OA, can be explained by the high prevalence of moderate hip OA and presumably therefore by the low specificity value, using hip pain as "gold standard" for Croft's grade (≥ grade 3). Therefore it is difficult to compare the definition of Croft (moderate hip OA, ≥ grade 3) with the other definitions. An earlier study reported similar prevalence of hip OA defined by Croft grade (≥ grade 3) and MJS (≤ 2.5 mm) and also similar prevalence of hip pain in "disease positive" hips (14). When we excluded those subjects with an incident hip fracture during follow-up time, and repeated the analysis for predictive validity, the results did not change essentially.

The second objective of the present study was to investigate whether the relationship between the three definitions and symptoms was gender dependent. Surprisingly, only the strength of the association between the Kellgren & Lawrence grading system and symptoms of hip OA was gender dependent. These findings are not reported in previous studies. A possible explanation for this gender dependency could be the stronger relationship between (femoral) osteophytes and hip pain for women. In women we found a stronger relationship between osteophytes and hip pain (OR of 1.7 for women versus 1.2 for men); however, the prevalence of osteophytes in women was lower (34.3% for women versus 43.6% for men). In contrast to our findings, we had expected that the strength of the association between MJS and symptoms of hip OA would have been gender dependent.

In older persons a stronger association between hip ROA and hip pain was found compared to younger persons. Because of a power problem (due to the smaller sample size for the younger men category) the gender difference was not significant. A possible explanation for this difference might be that younger persons have better muscle strength of the lower limb than older persons. Reduced muscle strength is regarded as a risk factor for pain and disability in OA (30–32) and exercise therapy, with the aim to improve muscle strength, has a beneficial effect on pain in patients with OA of the hip or knee (33, 34).

52

The results of the present study may be flawed by the quality of the radiographs. Especially measurement of joint space width (MJS) could be flawed because of the quality of the radiographs. Important variations in the radiographic procedure are the position of the central ray of the X ray beam relative to the centre of the joint, and the distance between the centre of the joint and the X ray film (focus to film distance). Centring the X ray beam on the umbilicus instead of on the superior aspect of the symphisis pubis resulted in a average increase in joint space width of about 10% (16). The focus to film distance may also modify the measurement (35). On the other hand in the study of Croft (14) the X ray beam was also centred 10 cm higher than a standard anteroposterior view of the pelvis.

The source of potential bias in this study is a likely health-based selection. The subjects in the present study had to be mobile enough to visit the research centre at baseline and follow-up and survive the follow-up period (mean 6.6 years). Overall, participants were generally healthier than non-participants. In other words, patients with the most severe symptoms were most likely not included. It seems probable that, in this younger and healthier population with less frequent lower limb disability and hip pain, the prevalence of hip ROA as well as the magnitude of the association between the different definitions of hip ROA and symptoms of hip OA is underestimated. Knowing that for older persons a stronger relationship was found between hip ROA and hip pain, especially when defined by Kellgren & Lawrence, this underestimation may particularly hold for Kellgren & Lawrence.

When we compared the results of Kellgren & Lawrence and MJS we found the following differences. Kellgren & Lawrence was the best predictor for a THR at follow-up. As described earlier by Buckland-Wright (36) and Lanyon et al. (37), we also found that men had larger joint spaces than women. After adjustment for height these joint space differences between men and women disappeared. Considering these results, it is doubtful whether the given cut-off point of MJS is valid for people of short stature, for example Asians.

When we stratified the associations between each definition and symptoms of hip OA (hip pain and lower limb disability) for gender, surprisingly we found for Kellgren & Lawrence significantly stronger associations for women with hip pain and lower limb disability than for men.

Based on these results, we concluded that Kellgren & Lawrence is still a useful definition for hip ROA for epidemiological studies.


Acknowledgments

53

References

1. Felson DT, Zhang Y. An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. Arthritis Rheum 1998;41(8):1343–55.
2. Birrell F, Croft P, Cooper C, Hosie G, Macfarlane G, Silman A. Health impact of pain in the hip region with and without radiographic evidence of osteoarthritis: a study of new attenders to primary care. Ann Rheum Dis 2000;59:857–63.
3. Sun Y, Gunther K, Brenner H. Reliability of radiographic grading of osteoarthritis of the hip and knee. Scand J Rheumatol 1997;26(3):155–65.
4. Hirsch R, et al. Hip osteoarthritis prevalence estimates by three radiographic scoring systems. Arthritis Rheum 1998;41(2):361–8.
5. Gunther KP, Sun Y. Reliability of radiographic assessment in hip and knee osteoarthritis. Osteoarthritis Cartilage 1999;7(2):239–46.
6. Reijman M, Hazes JMW, Koes BW, Verhagen AP, Bierma-Zeinstra SMA. Validity, reliability and applicability of seven definitions of hip osteoarthritis used in epidemiological studies: a systematic appraisal. Ann Rheum Dis 2004;63(3):226–32.
7. Croft P, Cooper C, Coggon D. Case definition of Hip Osteoarthritis in epidemiologic studies. J Rheumatol 1994;21(4):591–2.
8. Hofman A, Grobbee D, Jong PTde, VandenOuweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. Eur J Epidemiol 1991;7(4):403–22.
9. Odding E, Valkenburg H, Algra D, Vandenouweland F, Grobbee D, Hofman A. Associations of radiological osteoarthritis of the hip and knee with locomotor disability in the Rotterdam Study. Ann Rheum Dis 1998;57(4):203–8.
10. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494–502.
11. Kellgren JH, Jeffrey MR, Ball J. The epidemiology of chronic rheumatism. Atlas of standard radiographs of arthritis. Oxford: Blackwell Scientific Publications 1963.
12. Altman RD, Hochberg MC, Murphy WA, Wolfe F, Lequesne M. Atlas of individual radiographic features in osteoarthritis. Osteoarthritis Cartilage 1995;3:3–70.
13. Burnett S, Hart DJ, Cooper C, Spector TD. A radiographic atlas of osteoarthritis. London: Springer-Verlag 1994.
14. Croft P, Cooper C, Wickham C, Coggon D. Defining osteoarthritis of the hip for epidemiologic studies. Am J Epidemiol 1990;132(3):514–22.
15. Auleley G, Rousselin B, Ayral X, Edouard-Noel R, Dougados M, Ravaud P. Osteoarthritis of the hip: agreement between joint space width measurements on standing and supine conventional radiographs. Ann Rheum Dis 1998;57:519–23.
16. Auleley GR, Duche A, Drape JL, Dougados M, Ravaud P. Measurement of joint space width in hip osteoarthritis: influence of joint positioning and radiographic procedure. Rheumatology 2001;40(4):414–9.
17. Ravaud P, et al. Assessment of joint space width in patients with osteoarthritis of the knee: a comparison of 4 measuring instruments. J Rheumatol 1996;23:1749–55.
18. Donner A, Eliasziw M. A goodness-of-fit approach to inference procedures for the Kappa statistic: confidence interval construction, significance testing and sample size estimation. Stat Med 1992;11:1511–9.
19. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Stat Med 1998;17:101–10.
20. Buchbinder R, Goel V, Bombardier C, Hogg-Johnson S. Classification systems of soft tissue disorders of the neck and upper limb: do they satisfy methodological guidelines? J Clin Epidemiol 1996;49(2):141–9.
21. Felson DT, Anderson JJ. Methodological and statistical approaches to criteria development in rheumatic diseases. Baillieres Clin Rheumatol 1995;9(2):253–66.
22. Ingvarsson T, Hagglund G, Lindberg H, Lohmander L. Assessment of primary hip osteoarthritis: Comparison of radiographic methods using colon radiographs. Ann Rheum Dis 2000;59(8):650–3.

23. Bierma-Zeinstra SMA, Bohnen AM, Ginai AZ, Prins A, Verhaar JAN. Validity of American College of Rheumatology criteria for diagnosing hip osteoarthritis in primary care research. J Rheumatol 1999;26(5):1129–33.

24. Antoniades L, MacGregor AJ, Matson M, Spector TD. A cotwin control study of the relationship between hip osteoarthritis and bone mineral density. Arthritis Rheum 2000;43(7):1450–5.

25. Birrell F, Croft P, Cooper C, Hosie G, Macfarlane G, Silman A. Radiographic change is common in new presenters in primary care with hip pain. Rheumatology 2000;39:772–5.

26. MacGregor A, Antoniades L, Matson M, Andrew T, Spector TD. The genetic contribution to radiographic hip osteoarthritis in women: Results of a classic twin study. Arthritis Rheum 2000;43(11):2410–6.

27. Fox KM, Hochberg MC, Resnik CS, Kenzora JE, Hebel JR, Zimmerman SI, Hudson J, Magaziner J. Severity of radiographic findings in hip osteoarthritis associated with total hip arthroplasty. J Rheumatol 1996;23:693–7.

28. Dougados M, Guegen A, Nguyen M, Berdah L, Lequesne M, Mazieres B, Vignon E. Requirement for Total Hip Arthroplasty: an outcome measure of hip osteoarthritis? J Rheumatol 1999;26:855–61.

29. Brazier JE, Harper R, Munro J, Walter SJ, Snaith ML. Generic and condition-specific outcome measures for people with osteoarthritis of the knee. Rheumatology 1999;38:870–7.

30. Baar ME van, Dekker J, Lemmens JA, Oostendorp RA, Bijlsma JW. Pain and disability in patients with osteoarthritis of hip or knee: the relationship with articular, kinesiological, and psychological characteristics. J Rheumatol 1998;25(1):125–33.

31. Dekker J, Boot B, Woude LH van der, Bijlsma JW. Pain and disability in osteoarthritis: a review of biobehavioral mechanisms. J Behav Med 1992;15(2):189–214.

32. Steultjens MP, Dekker J, Baar ME van, Oostendorp RA, Bijlsma JW. Muscle strength, pain and disability in patients with osteoarthritis. Clin Rehabil 2001;15(3):331–41.

33. Baar ME van, Assendelft WJ, Dekker J, Oostendorp RA, Bijlsma JW. Effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a systematic review of randomized clinical trials. Arthritis Rheum 1999;42(7):1361–9.

34. Baar ME van, Dekker J, Oostendorp RA, Bijl D, Voorn ThB, Bijlsma JW. Effectiveness of exercise in patients with osteoarthritis of hip or knee: nine months' follow up. Ann Rheum Dis 2001;60:1123–30.

35. Ravaud P, Dougados M. Radiographic assessment in osteoarthritis. J Rheumatol 1997;24:786–91.

36. Buckland-Wright J. Quantitative radiography of osteoarthritis. Ann Rheum Dis 1994;53:268–75.

37. Lanyon P, Muir K, Doherty S, Doherty M. Age and sex differences in hip joint spaces among asymtomatic subjects without structural change. Arthritis Rheumatism 2003;48(4):1041–6.

4

A new marker for osteoarthritis:
cross-sectional and longitudinal
approach
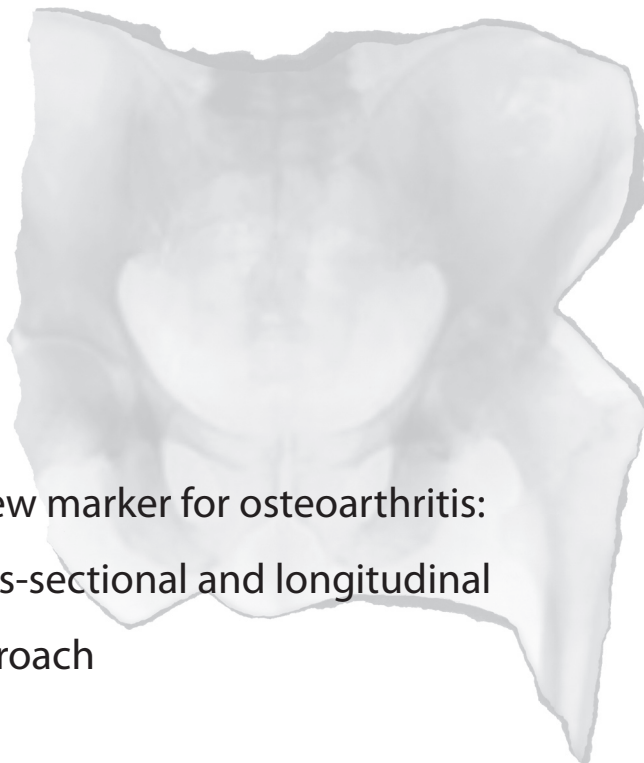
Abstract

Objective: To investigate the association between urinary concentrations of C-telopeptide fragments of collagen type II (CTX-II) and the prevalence and progression of ROA of the knee and hip.

Methods: The study population consisted of a sample of 514 men and 721 women aged 55 years and older, of the Rotterdam Study (population-based cohort study), with a mean follow-up time of 6.6 years. Prevalent ROA was defined by Kellgren ≥ grade 2 and progression of ROA as decrease of joint space width.

Results: Subjects with a CTX-II level in the highest quartile had a 4.2 times increased risk of having ROA of the knee (95% confidence interval (95% CI) 2.2–7.8) and at the hip (95% CI, 2.5–7.0) compared to subjects with a CTX-II level in the lowest quartile. We observed a stronger association for subjects with hip pain (17.1 (95% CI) 2.3–185.2) compared with those without hip pain (2.3 (95% CI) 1.5–6.0). Subjects with a CTX-II level in the highest quartile had a 6.2 times increased risk for progression of ROA at the knee (95% CI 1.2–31.6) and an 8.3 times increased risk for progression of ROA at the hip (95% CI 1.0–72.3). All of these associations were found to be independent of known risk factors for OA, such as age, gender and body mass index.

Conclusion: This study shows that CTX-II is associated with both prevalence and progression of ROA at the knee and hip. Importantly, this association is independent from known clinical risk factors for OA and seems stronger in subjects with joint pain.

Introduction

Osteoarthritis (OA) is a common age-related disabling locomotor disease characterized by degradation of articular cartilage. The most commonly used radiological method to assess cartilage damage is measurement of the joint space width. However, a limitation of using plain radiographs for detecting cartilage destruction is that significant cartilage degradation must have occurred in order to be visible on a radiograph (1). Therefore, cartilage degradation detectable on radiographs is considered as an already irreversible joint damage. Because of its relatively insensitive reflection of the disease process, it also takes at least one or two years to detect progression of damage that has been visualized on radiographs.

To overcome this, biochemical markers aiming to detect changes in OA with more reliability and sensitivity, preferably in an early stage of OA, have been developed (1–4). Biochemical markers are molecules derived from connective tissue matrices, which are released into biological fluid during the process of tissue turnover (1, 2). Such a biochemical marker might be useful for early identification of patients with OA, of patients at high risk for progression, for monitoring disease progression, and for assessing therapeutic response in OA all because of their improved responsiveness compared with radiographs (2, 4). One approach to identify such a marker could involve the analysis of cartilage metabolism. Proteoglycans and type II collagen are the major constituents of cartilage (4). Type II collagen is localized almost exclusively in cartilage, where it is a major structural component of the tissue. Hence, measurements of fragments derived from this protein may potentially represent a specific marker for cartilage degradation (1, 3). Recently, a specific marker of cartilage degradation, measured as the urinary concentration of C-telopeptide fragments of collagen type II (CTX-II), was developed (2, 3, 5). Mouritzen et al. described slightly increased concentration of CTX-II with increasing age, higher CTX-II concentration for women (both after 55 years of age) and higher CTX-II concentration in subjects with a higher body mass index (BMI) (5). Some evidence supporting the use of CTX-II as a marker has already been obtained. Urinary CTX-II levels are elevated in diseases with increased cartilage turnover, such as OA (1, 2) and rheumatoid arthritis (6). Garnero et al. reported weak associations of CTX-II with prevalent knee radiological OA (ROA) (1), and also modest associations with progression of knee ROA (2). However, these studies are small and it remains uncertain to what extent CTX-II is an independent marker for ROA and which factors could modify the relation between CTX-II and ROA.

We were therefore interested to explore to what extent the CTX-II marker can be considered to be independent from known risk factors for ROA such as age, gender and BMI. Because of the limited number of subjects included in the studies on CTX-II up to now there is a clear need to examine larger populations to obtain more accurate estimates. Furthermore, it is conceivable that in combination with factors that might reflect an ongoing OA process, such as the presence of joint pain, a dynamic change in cartilage metabolism can be detected.

As such, joint pain can be considered as a potential effect modifier of the relation between CTX-II and ROA.

Therefore the present study investigated the association between CTX-II and the prevalence and progression of ROA of the knee and hip in a large population of men and women aged 55 years and over. Additionally, we stratified the baseline associations between CTX-II and ROA of the knee or hip for the presence of pain at baseline (knee or hip).

Subjects and Methods

The study population consisted of participants of the Rotterdam Study, a prospective cohort of men and women aged 55 years and over. The objective of the Rotterdam Study is to investigate the incidence of, and risk factors for, chronic disabling diseases. The rationale and study design have been described previously (7). The focus is on neurogeriatric, cardiovascular, ophthalmologic and locomotor diseases. All 10,275 inhabitants of Ommoord, a district in Rotterdam, were invited to participate. The response rate was 78%, resulting in 7,983 subjects participating in the present study. Written informed consent was obtained from each participant. The Medical Ethics Committee of the Erasmus University Medical Centre has approved the Rotterdam Study.

For the present study a sample of 1,235 subjects of the Rotterdam study was used. The selection was based on the availability of the radiographs of the hip and knee both at baseline and follow-up, and the availability of urine samples at baseline. The fact that subjects had to be mobile enough to visit the research center at baseline and follow-up, and survive the follow-up period, led to the selection of a relatively younger and healthier population. Compared to the total Rotterdam study population, the present study population was indeed younger (66.6 years versus 70.6 years), had a lower prevalence of lower limb disability at baseline ($\geq$ index score of 0.5: 11.4% versus 35.5%) and a somewhat lower percentage of women (58.4% versus 61.1%). The baseline measurements were conducted between April 1990 and July 1993, and the follow-up measurements between 1996 and 1999, with a mean follow-up time of 6.6 years (range: 5.1–9.4 years).

Radiographic assessment

Weight bearing anteroposterior radiographs of the knee and hip were obtained at 70 KV, a focus of 1.8, and a focus to film distance of 120 cm, applying a Fuji High Resolution G 35 × 43 cm film. Radiographs of the pelvis were obtained with both feet in 10° internal rotation and the X-ray beam centred on the umbilicus, and of the knee with the patellae in central position. Two trained readers independently evaluated the radiographs of the knee and hip at baseline and follow-up, unaware of the clinical status of the patients. All radiographs

were grouped by patient and read by pairs chronologically ordered, the chronological order being known to the reader (chronologically ordered reading procedure) (8).

At baseline, ROA of the knee and hip was quantified by measurements following the Kellgren & Lawrence grading system (9–12) (atlas-based) in five grades (from zero to four). A person was considered to have ROA of the knee or hip, if the Kellgren & Lawrence score of one or both joints was equal to or larger than two.

At baseline and follow-up the minimal joint space width (JSW) of the knee and hip joints were measured using a 0.5 millimetres graduated magnifying glass directly laid over the radiograph (13). For the knee the medial and lateral compartment was measured and for the hip the lateral, superior and axial compartment, as described previously by Croft et al. (13). Joint space narrowing (JSN) was defined as the JSW at baseline minus the JSW at follow-up (Δ JSW). Because of the absence of consensus concerning the cut-off point for JSN, we used different cut-off points for JSN, namely 1.0, 1.5 and 2.0 mm decrease of the JSW between baseline and follow-up. JSN was evaluated per compartment and for the knee a JSN of minimally 1 (out of 2; medial and lateral (14)) compartment and for the hip a JSN of minimally 1 (out of 3; lateral, superior and axial (13)) was defined as a positive progression. Additionally, we also used a JSN of the medial compartment of the knee as a definition of progression. Radiographic progression of JSN can be regarded as the most reliable measurement of OA progression (15).

The radiographs of the knee were scored for OA by two independent observers who were blinded to all data for the participant, as described previously (14, 16). After each set of 150 radiographs, the scores of the two readers were evaluated. Whenever the Kellgren & Lawrence score differed, the two readers met to read the radiographs together, and a consensus score was determined. Two independent readers tested the inter-rater reliability of the hip in a random set of 148 radiographs. We determined the inter-rater reliability for Kellgren & Lawrence to be 0.68 (Kappa statistics), and for the minimal joint space width we obtained an Intraclass Correlation Coefficient (ICC) of 0.85 (12).

Biochemical measurement

Overnight fasting urine samples were obtained from all subjects at baseline and kept frozen at −20° C.

Monoclonal antibody mAbF46, specific for CTX-II C-telopeptide fragments, was used in a competitive enzyme-linked immunosorbent assay (ELISA) format developed for measurement of urine samples, as described previously (3). In order to ensure the reproducibility and performance of the assay, three genuine urine samples were added as controls on each microtitre plate to assure performance of

the assay, and the entire plate was re-run if any of the genuine controls were measured with a concentration more than 20% of the predetermined value. The concentration CTX-II (ng/l)

was standardised to the total urine creatinine (mmol/l), and the unit for corrected CTX-II concentration was ng/mmol.

Potential confounders and effect modifiers

At baseline, trained interviewers performed an extensive home interview on demographic characteristics, medical history, risk factors for chronic diseases and medication use. Lower limb disability was assessed using a modified version of the Stanford Health Assessment Questionnaire (HAQ) (16). A lower limb disability index (LDI) was obtained by calculating the mean score of answers to six questions, as described previously (12). We used the LDI as measurement of mobility of the participant. The presence of knee and hip pain ("did you have joint complaints of your right/left knee/hip during the last month") was asked during the home interview at baseline.

Height and weight were measured with participants wearing indoor clothing without shoes. BMI was calculated as weight in kilograms divided by height in squared meters (kg/m^2).

Statistical analysis

Differences in baseline characteristics were evaluated by analysis of variance (ANOVA) for continuous variables and by Chi-square for categorical variables. Distribution analysis by the Shapiro-Wilk test showed that biochemical markers were not normally distributed and, thus, were log transformed to obtain normal distribution before statistical analysis. Hereafter CTX-II concentrations refer to log-transformed CTX-II concentrations. Influences of age, gender, and BMI on baseline CTX-II concentration were tested by independent t-tests.

The cross-sectional associations between CTX-II concentration and ROA of the knee or hip were assessed using logistic regression analysis to calculate odds ratios (ORs), by means of Generalised Estimating Equations (GEE) (cross-sectional design). This is a procedure of repeated measurements, which is used here to take account of the correlation between the left and right hip, while using each joint (left or right) as the observation unit (17). The ORs were calculated per quartile (with the 1st quartile as reference) and per standard deviation CTX-II. For the baseline associations we calculated crude ORs and adjusted the crude ORs for age, gender, BMI and LDI. Additionally, we stratified these associations for the presence or absence of pain in knee or hip (during the last month).

The longitudinal associations between baseline CTX-II concentration and progression of ROA of the knee or hip were assessed using logistic regression analysis to calculate ORs to estimate the relative risk for progression, by means of GEE (longitudinal design). ORs were calculated per quartile and per standard deviation CTX-II. For the associations between baseline CTX-II and progression of ROA of the knee or hip we calculated crude ORs and adjusted for age, gender, BMI, LDI, baseline Kellgren score and follow-up time. The baseline Kellgren score is a known risk factor for radiologic progression (18, 19). Additionally, we assessed the longitudinal associations between CTX-II concentration and

63

incident osteophytes at follow-up of the knee or hip. A (two-sided) P-value of 0.05 was considered significant.

We estimated the magnitude of confounding by the degree of discrepancy between the unadjusted and adjusted estimate (the change-in-estimate-criterion) (20). We choose a cut-off point of 10% for what constitutes as an important change in the estimate.

We used SPSS version 11.0 (SPSS Inc., Chicago, USA) and SAS software, version 8.2 (SAS Institute, Cary, NC) for all analyses.

Results

Table 1 presents the baseline characteristics of the total study population stratified for the absence or presence of knee or hip ROA. In this study population, with a mean age of 66.6 years, 19.2% of the subjects had ROA of the knee and 10.0% had ROA of the hip (Kellgren & Lawrence ≥ grade 2). During the last month before the baseline interview 12.3% of all subjects had knee pain and 18.1% had hip pain. The median CTX-II concentration (not log transformed) of the study population, was 177.0 ng / mmol. Participants with knee ROA were 3.1 years older, more frequently female (70.5% versus 50.2%), 3.9 kg (2.1 kg/m^2) heavier and 2.2 cm shorter, compared to those without knee ROA. Subjects with hip ROA were 3.8 years older compared to those without hip ROA. Compared with those with ROA of the hip, persons with ROA of the knee were more often women (70.5% versus 58.5%) and 2.1 kg (1.3 kg/m^2) heavier.

Table 1: Baseline characteristics of the study population, stratified by the absence / presence of radiological osteoarthritis (ROA) of the knee or hip.

	Study population	Persons without Knee ROA	Persons with Knee ROA	Persons without Hip ROA	Persons with Hip ROA
Number	1,235	998	237	1,112	123
Gender, % women	58.4	50.2	70.5**	58.1	58.5
Age, years ± SD	66.6 ± 6.8	66.0 ± 6.6	69.1** ± 6.9	66.2 ± 6.7	70.0** ± 6.7
Weight, kg ± SD	73.8 ± 11.5	73.1 ± 11.2	77.0** ±12.0	73.7 ± 11.6	74.9 ± 10.8
Height, cm ± SD	167.5 ± 9.1	167.9 ± 9.2	165.7* ± 8.4	167.5 ± 9.1	167.5 ± 8.8
Body Mass Index, kg/m^2 ± SD	26.3 ± 3.6	25.9 ± 3.4	28.0** ± 3.9	26.3 ± 3.6	26.7 ± 3.4
Presence of knee pain, %	12.3	15.1	30.8**	17.6	23.0
Presence of hip pain, %	18.1	11.1	17.1	10.3	29.9**
Lower limb disability, %	11.4	9.0	21.7	8.5	34.6
Concentration CTX-II, ng/mmol (median)[†]	177.0	167.0	228.0**	172.0	231.5**

[†] Concentration CTX-II, not log transformed

Significant differences between persons with radiological osteoarthritis (ROA) and persons without ROA
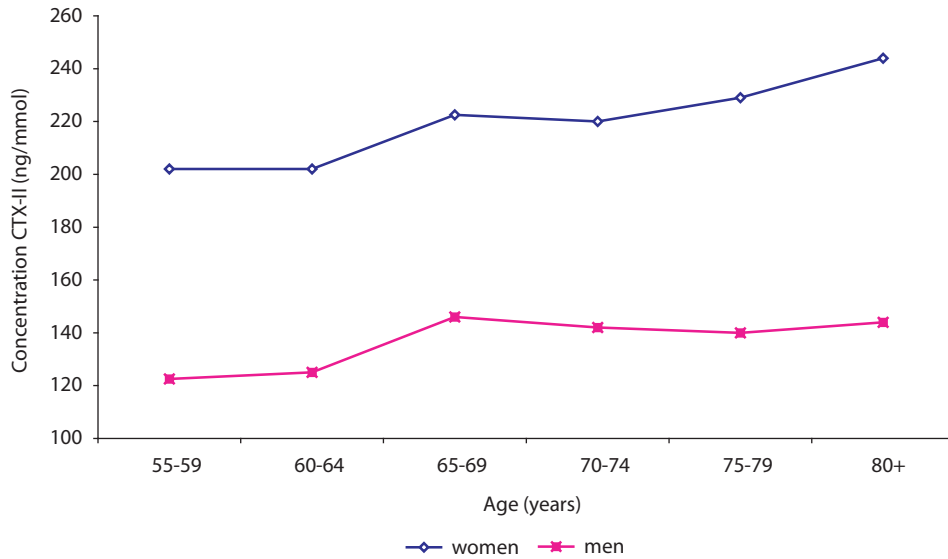
* P < 0.01, ** P < 0.001

Figure 1: Distribution of CTX-II (median value, ng/mmol) by age and gender.
Concentration of CTX-II not log transformed.

The CTX-II concentration was 72.3 ng/mmol higher in women than in men (P-value < 0.0001), increased 1.1 ng/mmol per year with age (P-value trend = 0.03) (Figure 1) and increased 3.3 ng/mmol per kg/m^2 with higher BMI (P-value trend < 0.0001). When we excluded participants with ROA of the knee or hip at baseline and those with incident ROA of the knee or hip at follow-up, only the gender difference in CTX-II concentration remained.

Table 2: Cross-sectional association between baseline CTX-II concentration and baseline radiological osteoarthritis (ROA) of the knee and/or hip (Kellgren & Lawrence ≥ grade 2).

Number of cases	Knee ROA		Hip ROA	
	237		**123**	
	crude OR	adj* OR	crude OR	adj* OR
CTX-II				
1st quartile (1.49–2.10)	**1**	**1**	**1**	**1**
2nd quartile (2.11–2.25)	**1.7** (1.0–2.9)	**1.7** (1.0–2.9)	**1.3** (0.7–2.5)	**1.5** (0.8–2.9)
3rd quartile (2.26–2.39)	**3.2** (2.0–5.2)	**2.8** (1.6–4.6)	**1.7** (0.9–3.1)	**2.1** (1.1–4.0)
4th quartile (2.40–3.11)	**5.2** (3.3–8.4)	**4.2** (2.5–7.0)	**3.6** (2.0–6.2)	**4.2** (2.2–7.8)
P-value for trend	< 0.0001	< 0.0001	< 0.0001	< 0.0001
CTX-II				
Per SD	**1.9** (1.6–2.2)	**1.7** (1.4–2.0)	**1.8** (1.5–2.1)	**1.8** (1.4–2.2)

Log-transformed CTX-II concentration is expressed in quartiles and standard deviation (SD)
Associations are presented by odds ratios (OR) with 95% confidence interval between parentheses, for risk of ROA by CTX-II levels.
*Associations are adjusted for age, gender, body mass index and lower limb disability index.
Knee and hip radiological osteoarthritis is defined as Kellgren & Lawrence ≥ grade 2 in minimally 1 joint.

Table 3: Associations between baseline CTX-II concentration and radiological progression of knee osteoarthritis.

Number of cases	JSN ≥ 1.0 mm		JSN ≥ 1.5 mm		JSN ≥ 2.0 mm	
	233		73		26	
	crude OR	adj* OR	crude OR	adj* OR	crude OR	adj* OR
CTX-II						
1st quartile (1.49–2.10)	**1**	**1**	**1**	**1**	**1**	**1**
2nd quartile (2.11–2.25)	**1.0** (0.7–1.5)	**0.9** (0.6–1.5)	**1.5** (0.7–3.3)	**1.3** (0.6–2.9)	**3.7** (0.8–17.7)	**4.1** (0.8–20.5)
3rd quartile (2.26–2.39)	**1.2** (0.8–1.8)	**1.1** (0.7–1.7)	**1.9** (0.9–4.1)	**1.5** (0.6–3.3)	**3.6** (0.7–17.5)	**4.5** (0.9–23.0)
4th quartile (2.40–3.11)	**1.2** (0.8–1.8)	**1.1** (0.7–1.7)	**2.5** (1.2–5.2)	**1.8** (0.8–4.1)	**5.2** (1.1–23.8)	**6.0** (1.2–30.8)
P-value for trend	0.219	0.730	0.009	0.120	0.033	0.064
CTX-II Per SD	**1.1** (1.0–1.3)	**1.1** (.9–1.3)	**1.5** (1.1–1.8)	**1.4** (1.0–1.8)	**1.5** (1.0–2.2)	**1.6** (1.0–2.5)

Log-transformed CTX-II concentration is expressed in quartiles and standard deviation (SD).

Associations are presented by odds ratios (OR) with 95% confidence interval between parentheses, for risk of radiological osteoarthritis by CTX-II levels.

*Associations are adjusted for age, gender, body mass index, lower limb disability index, baseline radiological osteoarthritis of the knee, baseline radiological osteoarthritis of the hip and follow-up time.

JSN (joint space narrowing) is defined as the joint space width at baseline minus the joint space width at follow-up (of the medial and lateral compartment), using different cut-off points.

Baseline CTX-II concentration (in quartiles and per standard deviation) is higher in subjects with baseline ROA of the knee and hip compared to those without baseline ROA of the knee and hip (Table 2). The crude data showed a stronger (but not significant) association between the highest quartile CTX-II and ROA of the knee than for ROA of the hip. After adjustment for gender and age, the risk estimate increased for the hip and decreased for the knee, resulting in similar ORs for the hip and knee. Additional adjustment for BMI and lower LDI did not essentially change the risk estimates for the knee and hip. Overall, we observed a clear trend that the higher the CTX-II concentration, the stronger the association with prevalent ROA of the knee and hip.

Table 3 shows the associations between baseline CTX-II concentration (in quartiles and per standard deviation) and progression of knee ROA using different cut-off points for JSN. We found significant crude associations between a decrease in joint space of ≥ 1.5 mm or ≥ 2.0 mm, and the highest quartile of CTX-II. After adjustment for BMI, age, gender, LDI and baseline ROA of the knee the risk estimates changed importantly, and only the association between JSN ≥ 2.0 mm and the 4th quartile of CTX-II reached significance with an OR of 6.2. We observed a clear trend, especially for a JSN of ≥ 2.0 mm, but also for a JSN of ≥ 1.5 mm, the higher the CTX-II concentration, the stronger the association with progression of knee ROA. Additionally, we also assessed the association between CTX-II and progression of the medial compartment. These associations did not essentially differ with the abovementioned associations (JSN ≥ 1.5 mm and 4th quartile CTX-II: adjusted OR of 2.0 (95% CI) .8–5.1).

Table 4 shows the associations between baseline CTX-II concentration (in quartiles and per standard deviation) and progression of hip ROA (for different cut-off points as defined by

Table 4: Associations between baseline CTX-II concentration and radiological progression of hip osteoarthritis.

	JSN ≥ 1.0 mm		JSN ≥ 1.5 mm	
Number of cases	73		24	
	crude OR	adj* OR	crude OR	adj* OR
CTX-II				
1st quartile (1.49–2.10)	**1**	**1**	**1**	**1**
2nd quartile (2.11–2.25)	**1.1** (0.5–2.6)	**1.0** (0.4–2.4)	**4.2** (0.5–37.4)	**3.9** (0.4–36.9)
3rd quartile (2.26–2.39)	**2.3** (1.1–4.7)	**2.1** (0.9–4.6)	**8.5** (1.1–68.5)	**8.3** (1.0–72.2)
4th quartile (2.40–3.11)	**2.8** (1.4–5.8)	**1.7** (0.7–4.0)	**12.3** (1.6–95.5)	**8.4** (1.0–72.9)
P-value for trend	< 0.0001	0.05	< 0.0001	0.005
CTX-II Per SD	**1.6** (1.3–2.0)	**1.3** (1.0–1.8)	**2.2** (1.5–3.2)	**1.9** (1.2–3.0)

Log-transformed CTX-II concentration is expressed in quartiles and standard deviation (SD)
Associations are presented by odds ratios (OR) with 95% confidence interval between parentheses, for risk of ROA by CTX-II levels.
*Associations are adjusted for age, gender, body mass index, lower limb disability index (LDI), baseline radiological osteoarthritis of the hip, baseline radiological osteoarthritis of the knee and follow-up time.
JSN (joint space narrowing) is defined as the joint space width at baseline minus joint space width at follow-up (of the lateral, superior and axial compartment), using the different cut off points.
The associations with the cut-off point ≥ 2.0 mm were not given because of too low statistical power (N = 11).

JSN) are shown. The results for the JSN cut-off point ≥ 2.0 mm are not presented because the power is too low (11 cases). At the hip we found a trend similar to prevalent ROA of the knee and hip; i.e. the higher the CTX-II concentration, the stronger the association with progression of ROA. After adjustment for BMI, age, gender, LDI and baseline ROA of the hip only the associations between JSN ≥ 1.5 mm and the 4th quartile of CTX-II reached significance with an OR of 8.3. When we compared the association between the different aspects of ROA as measured by the Kellgren & Lawrence score, i.e. osteophytes and JSN, we observed no association with incident osteophytes of the knee and the hip. The ORs for the 4th quartile of CTX-II were 0.3 for both knee and hip (P-values of 0.288 and 0.232, respectively).

Figure 2 shows the baseline associations between high CTX-II concentrations (4th quartile) and ROA of the knee and hip, stratified for the absence or presence of knee or hip pain. For this analyses we compared subjects with a high CTX-II concentration (4th quartile) with those with a low concentration (1st quartile), resulting in lower number of subjects as reported before. We observed substantially stronger associations between CTX-II levels and ROA for subjects with hip pain (OR 20.4) compared to those without hip pain (OR 3.0). Adjustment for potential confounders changed the risk estimates importantly (from 17.1 to 20.4 and from 2.3 to 3.0) for subjects with and without hip pain, respectively. The difference between the ORs for subjects with versus without hip pain just failed to reach significance (P-value 0.105). In case of ROA of the knee the differences in ORs between participants with and without knee pain were similar but smaller than found for the hip. After adjustment for
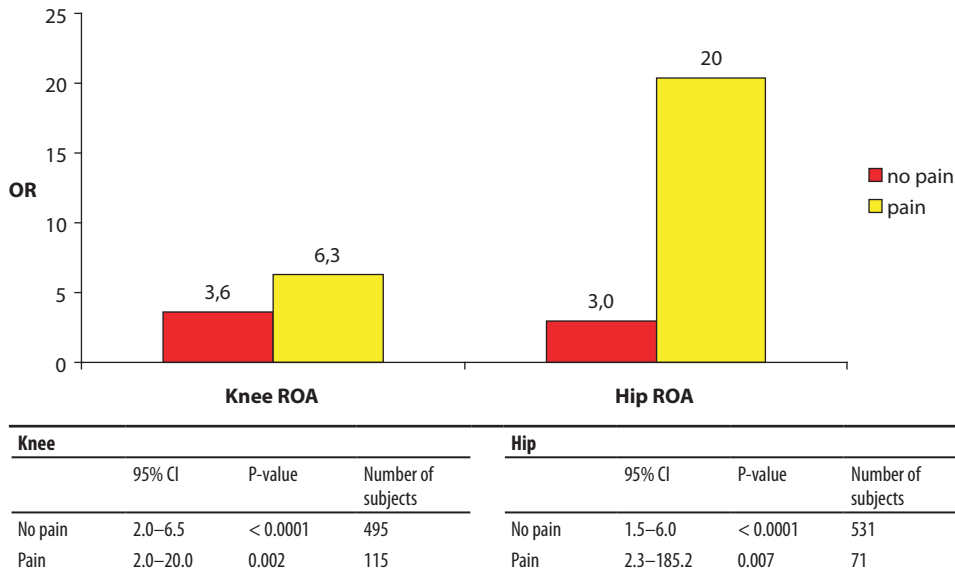
67

Figure 2: Associations between baseline CTX-II concentration (log transformed, highest quartile) and baseline radiological osteo-arthritis (ROA) of the knee or hip (Kellgren & Lawrence ≥ grade 2), stratified for the presence of (knee/hip) pain.
Associations are adjusted for age, gender, body mass index and lower limb disability index.

Knee				Hip			
	95% CI	P-value	Number of subjects		95% CI	P-value	Number of subjects
No pain	2.0–6.5	< 0.0001	495	No pain	1.5–6.0	< 0.0001	531
Pain	2.0–20.0	0.002	115	Pain	2.3–185.2	0.007	71

potential confounders the risk estimates changed importantly for participants with (OR from 7.1 to 6.3) and without (OR from 4.3 to 3.6) knee pain.

Discussion

We report the analysis of CTX-II levels in urine in a large population-based prospective cohort study that indicates a strong relation between CTX-II levels and risk of ROA. For persons with a CTX-II level in the highest quartile, we observed a more than four times increased risk of having prevalent ROA of the knee or hip, and a more than 6 to 8 times increased risk for progression of ROA at the knee and the hip, respectively. All these associations were found to be independent from known risk factors for ROA, including age, gender, BMI and baseline Kellgren score. Furthermore, CTX-II seems to be a specific marker for cartilage degradation, since CTX-II is associated with joint space narrowing but not with incident osteophytes.

The baseline associations seemed stronger for those participants with hip pain compared with those without hip pain. Because of the low numbers of subjects with hip pain the confidence interval of the association for those with hip pain are huge (95% CI 2.3–185.2) and overlaps with the CI for those without hip pain. We confirm that women had a higher CTX-II concentration than men, and found that this is not explained by prevalent or incident

ROA of the knee or hip. Thus the present study shows that a single degradation marker (CTX-II) can identify patients who are at high risk for rapid progression of joint destruction.

The distribution of CTX-II concentration by age, gender and BMI in the present study was similar to that described by Mouritzen et al (5). We found a slight rise in urinary CTX-II with increasing age (after 55 years of age), a significantly higher level for women compared to men (after 55 years of age), and a significantly higher CTX-II concentration in subjects with higher BMI. The increased concentration with age seems to reflect the increase in prevalence of ROA with increasing age. However, the higher concentration found in women remains after we excluded participants with prevalent (at baseline) and incident ROA (at follow-up) of the knee or hip. In line with this, Mouritzen et al.(5) reported a sudden and marked increase in CTX-II concentration after the menopause. This observation may be explained by a higher turnover rate for cartilage in women after menopause. Indeed, a recent study in cynomolgus monkeys showed that ovariectomy induced OA lesions of articular cartilage (21). Furthermore, in a cross-sectional observational study, Wluka et al. (22) reported that the use of estrogen replacement therapy (ERT) for more than five years is associated with greater knee cartilage volume. Similarly, a number of retrospective and observational studies indicated that ERT is associated with decreased prevalence of OA, but this finding is not universal (23). Finally, polymorphisms in the estrogen receptor α gene have been identified as genetic risk factors for knee OA (14). Altogether these data suggest that estrogen can prevent cartilage erosion, and thereby identify the estrogen endocrine system as a significant regulator of cartilage turnover and structural integrity (21). However, the exact mechanism whereby estrogen influences cartilage metabolism needs further investigation (21, 23–26).

Type II collagen markers are probably a specific tool for detecting changes in OA (27). Other proposed markers of OA, such as collagen crosslinks, proteoglycan, cartilage oligomeric matrix protein, matrix metalloproteinases and inflammatory markers (27), reflect general remodeling of the various tissues of the cartilage, bone and synovium. Up to now, increased serum or urine levels of the different markers have been obtained from small cross-sectional studies (27, 28). This is the first large follow-up study that investigated the use of CTX-II as biomarker for cartilage degradation and disease progression.

The strengths of the present study are its size, its population-based prospective design and the clinically meaningful follow-up period of 6.6 years. A potential limitation of the present study might be that the results are based on a single determination of CTX-II at baseline. Because of a possible diurnal variability of the CTX-II level, we obtained overnight fasting urine samples of all subjects. However, we found no an indication for the presence of a systematic bias due to the inherent variability of the measurements of CTX-II. Another limitation is a potential health-based selection bias. The subjects in the present study had to be mobile enough to visit the research center at baseline and follow-up, and survive the follow-up period (mean 6.6 years). Overall, participants were generally healthier than non-participants. In other words, patients with the most severe symptoms were most likely

not included. Therefore it seems probable that, in this younger and healthier population with less frequent lower limb disability and (knee and hip) pain, the prevalence of knee and hip ROA at baseline and the number of cases with progression of ROA at follow-up is underestimated. This could have resulted in an underestimation of the reported associations. Another limitation is the used radiographic procedure of the knee, the serial anterior-posterior radiograph. The reliability of radiographic JSW measurements in the knee increases when an anterior-posterior radiograph of the knee in 20–30 degrees flexion was used (28, 29), and therefore this procedure has been recommended for longitudinal studies (30, 31). The procedure used in the present study could have resulted in an under- or overestimation of the reported associations of the knee. The reliability of the measurements of the joint space width of the knee radiographs of the present study was not assessed. As reported by Günther and Sun (32, 33) the lateral joint space width measurement is less reliable compared to the medial joint space. Additionally, we repeated the analyses between CTX-II and progression of knee ROA with another definition of progression, namely JSN of only the medial compartment. The associations we found did not differ essentially with the associations reported in the present study.

Based on the results of the present study, we conclude that CTX-II is markedly associated with the prevalence and progression of ROA of the knee and hip, and that these associations are independent of known risk factors for ROA. The presence of joint pain seems to augment this relationship, which might reflect the effects of an ongoing OA process. The increase of CTX-II in women after menopause may reflect a protective effect of estrogen on cartilage loss. Further research is necessary to establish the clinical utility of this novel biomarker for OA.

## Acknowledgments

References

1. Garnero P, Piperno M, Gineyts E, Christgau S, Delmas PD, Vignon E. Cross sectional evaluation of biochemical markers of bone, cartilage, and synovial tissue metabolism in patients with knee osteoarthritis: relations with disease activity and joint damage. Ann Rheum Dis 2001;60:619–26.

2. Garnero P, et al. Uncoupling of Type II collagen synthesis and degradation predicts progression of joint damage in patients with knee osteoarthritis. Arthritis Rheum 2002;46(10):2613–24.

3. Christgau S, et al. Collagen Type II C-telopeptide fragments as an index of cartilage degradation. Bone 2001;29(3):209–15.

4. Vignon E, et al. Recommendations for the registration of drugs used in the treatment of osteoarthritis: an update on biochemical markers. Osteoarthritis Cartilage 2001;9:289–93.

5. Mouritzen U, Christgau S, Lehmann H-J, Tanko LB, Christiansen C. Cartilage turnover assessed with a newly developed assay measuring collagen type II degradation products: influence of age, sex, menopause, hormone replacement therapy, and body mass index. Ann Rheum Dis 2003;62:332–6.

6. Garnero P, Gineyts E, Christgau S, Finck B, Delmas PD. Association of baseline levels of urinary glucosyl-galactosyl-pyridinoline and type II collagen C-telopeptide with progression of joint destruction in patients with early rheumatoid arthritis. Arthritis Rheum 2002;46:21–30.

7. Hofman A, Grobbee DE, Jong PTde, VandenOuweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. Eur J Epidemiol 1991;7(4):403–22.

8. Auleley G, Girardeau B, Dougados M, Ravaud P. Radiographic assessment of hip osteoarthritis progression: impact of reading procedures for longitudinal studies. Ann Rheum Dis 2000;59:422–7.

9. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494–502.

10. Reijman M, Hazes JMW, Koes BW, Verhagen AP, Bierma-Zeinstra SMA. Validity, reliability and applicability of seven definitions of hip osteoarthritis used in epidemiological studies: a systematic appraisal. Ann Rheum Dis 2004;63(3):226–32.

11. Kellgren JH, Jeffrey MR, Ball J. The epidemiology of chronic rheumatism. Atlas of standard radiographs of arthritis. Oxford: Blackwell Scientific Publications 1963.

12. Reijman M, Hazes JMW, Pols HAP, Bernsen RMD, Koes BW, Bierma-Zeinstra SMA. Validity and reliability of three definitions of hip osteoarthritis. Ann Rheum Dis 2004; in press.

13. Croft P, Cooper C, Wickham C, Coggon D. Defining osteoarthritis of the hip for epidemiologic studies. Am J Epidemiol 1990;132(3):514–22.

14. Bergink AP, et al. Estrogen receptors alpha gene haplotype is associated with radiographic osteoarthritis of the knee in elderly men and women. Arthritis Rheum 2003;48(7):1913–22.

15. Conrozier T, et al. Quantitative measurement of joint space narrowing progression in hip osteoarthritis: a longitudinal retrospective study of patients treated by total hip arthroplasty. Br J Rheumatol 1998;37(9):961–8.

16. Odding E, Valkenburg H, Algra D, Vandenouweland F, Grobbee D, Hofman A. Associations of radiological osteoarthritis of the hip and knee with locomotor disability in the Rotterdam Study. Ann Rheum Dis 1998;57(4):203–8.

17. Zhang Y, Glynn RJ, Felson DT. Musculoskeletal disease research: Should we analyse the joint or the person? J Rheumatol 1996;23:1130–4.

18. Dougados M, et al. Radiological progression of hip osteoarthritis: definition, risk factors and correlations with clinical status. Ann Rheum Dis 1996;55:356–62.

19. Lievense AM, Bierma-Zeinstra SMA, Verhagen AP, Verhaar JAN, Koes BW. Prognostic factors of progress of osteoarthritis of the hip: A systematic review. Arthritis Rheum 2002;47(5):556–62.

20. Rothman K.J, Greenland S. Introduction to stratified analysis. In: Rothman KJ, Greenland S, editors. Modern epidemiology. Second edition. Boston/Toronto: Little, Brown and Company. 1998:311–26.

21. Ham KD, Loeser RF, Lindgren BR, Carlson CS. Effects of long-term estrogen replacement therapy on osteoarthritis severity in cynomolgus monkeys. Arthritis Rheum 2002;46:2081–90.

22.   Wluka AE, Davis SR, Bailey M, Stuckey SL, Cicuttini FM. Users of oestrogen replacement therapy have more knee cartilage than non-users. Ann Rheum Dis 2001;60:332–6.

23.   Felson DT, Nevitt MC. Estrogen and osteoarthritis: how do we explain conflicting study results? Prev Med 1999;28:445–8.

24.   Dennison EM, Arden NK, Kellingray S, Croft P, Coggon D, Cooper C. Hormone replacement therapy, other reproductive variables and symptomatic hip osteoarthritis in elderly white women: a case-control study. Br J Rheumatol 1998;37:1198–1202.

25.   Spector TD, Nandra D, Hart DJ, Doyle DV. Is hormone replacement therapy protective for hand and knee osteoarthritis in women? The Chingford Study. Ann Rheum Dis 1997;56:432–4.

26.   Nevitt MC, et al. Association of estrogen replacement therapy with the risk of osteoarthritis of the hip in elderly white women. Study of Osteoporotic fractures research group. Arch Intern Med 1996;156:2073–80.

27.   DeGroot J, Bank RA, Tchetverikov I, Verzijl N, TeKoppele JM. Molecular markers for osteoarthritis: the road ahead. Curr Opin Rheumatol 2002;14:585–9.

28.   Vignon E, et al. Measurement of radiographic joint space width in the tibiofemoral compartment of the osteoarthritic knee: comparison of standing anteroposterior and Lyon schuss views. Arthritis Rheum 2003;48(2):378–84.

29.   Mazzuca SA, Brandt KD, Lane KA, Katz BP. Knee pain reduces joint space width in conventional standing anteroposterior radiographs of osteoarthritic knees. Arthritis Rheum 2002;46(5):1223–7.

30.   Dieppe P. Recommended methodology for assessing the progression of osteoarthritis of the hip and knee joints. Osteoarthritis Cartilage 1995;3:73–7.

31.   Altman R, et al. Design and conduct of clinical trials in patients with osteoarthritis. Recommendations from a task force of the osteoarthritis Research Society. Osteoarthritis Cartilage 1996;4:217–43.

32.   Gunther KP, Sun Y. Reliability of radiographic assessment in hip and knee osteoarthritis. Osteoarthritis Cartilage 1999;7(2):239–46.

33.   Sun Y, Gunther K, Brenner H. Reliability of radiographic grading of osteoarthritis of the hip and knee. Scand J Rheumatol 1997;26(3):155–65.

5

X-ray findings strongly predict progression of osteoarthritis of the hip
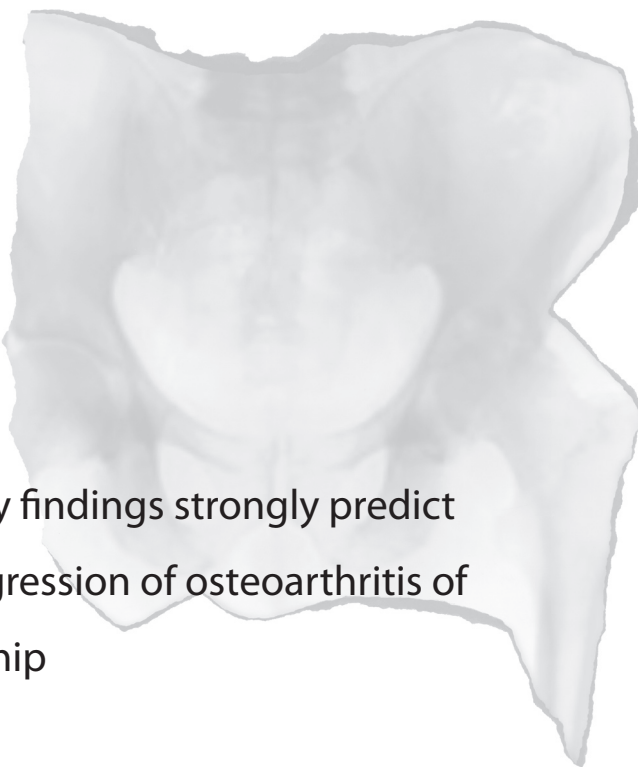
Abstract

Objectives: To investigate which variables identify persons at high risk for progression of hip osteoarthritis (OA).

Methods: In 1,920 men and women aged 55 years and older from the Rotterdam Study (a population-based cohort study) potential determinants of progression of hip OA were collected at baseline. X-rays of the hip at baseline and follow-up (mean follow-up time of 6.6 years) were evaluated. Radiologic progression of hip OA was defined as a decrease of joint space width ($\geq$ 1.0 mm) at follow-up, or an incident total hip replacement. Using multi-variate logistic regression models, the association between potential risk factors and progression of hip OA was assessed.

Findings: In 13.1% of the study population (mean age of 66.1 years) radiologic progression of hip OA was evident. Starting with a simple model of only directly obtainable variables, the Kellgren & Lawrence score at baseline, when added to the model, was a strong predictor with an odds ratio (OR) of 5.6, increasing in those subjects with hip pain at baseline to 31.7. A cartilage degradation marker (CTX-II) had an independent additional association with progression (OR of 2.2 and 3.9, respectively).

Conclusion: The Kellgren & Lawrence score at baseline was by far the strongest predictor for progression of hip OA, especially in those with existing hip pain at baseline. In patients with hip pain, an X-ray has strong additional value to identify those at high risk for progression of hip OA.

Submitted

Introduction

Osteoarthritis (OA) of the hip is one of the main causes of disability among the elderly and the prevalence of hip OA will increase with the aging of the Western society (1, 2).

The management of patients with OA focuses on symptom relief and preservation of function (3, 4) including, in case of severe symptomatic OA, the consideration of joint replacement (3, 4). The attempts to identify potential disease-modifying OA drugs (DMOADs) that may halt or retard joint destruction have produced differential results. Hence, identification of persons at high risk for rapid progression of OA is important for at least two reasons. Firstly, well-characterized 'high risk' groups may be useful in clinical trials and, secondly, assuming that DMOADs do become available in the future, to identify primary target groups in need of such therapy. Additionally, the identified non-progressors can be given some reassurance about their disease status.

There is no consensus on how to define progression of hip OA (5). International committees have suggested to evaluate both structural (joint space narrowing) and symptomatic variables of OA (pain, functional impairment, overall assessment by the patient) in clinical studies (6, 7). A potential composite outcome measure is the need for surgery (total hip replacement; THR), based on the assumption that THR is performed only in patients who have a severe symptomatic OA together with structural damage of the hip (8, 9).

Potential factors that may identify persons at risk for progression of hip OA include systemic factors (e.g. metabolic, hormonal, genetic, age and gender), local biomechanical factors, such as mechanical workload, body mass index (BMI) and acetabular dysplasia, and already existing osteoarthritic changes such as radiological signs, clinical symptoms and signs of cartilage degradation. In a recent review, Lievense et al. (10) reported that radiological features were the main mediators of progression of hip OA; however, all the included studies had a small study population, a short follow-up time, and were hospital based.

The present study investigated in a large open population with a long-term follow-up period, which easily measurable determinants (clinically relevant risk factors) will best identify those persons at high risk for progression of hip OA.

Subjects and Methods

The study population consisted of participants of the Rotterdam Study, a prospective cohort of men and women aged 55 years and over. The objective of the Rotterdam Study is to investigate the incidence of, and risk factors for, chronic disabling diseases; the rationale and study design have been described previously (11). Written informed consent was obtained from each participant. The Medical Ethics Committee of the Erasmus Medical Center has approved the Rotterdam Study.

For the present study a sample of 1,920 subjects from the Rotterdam study was used. The selection was based on the availability of radiographs of the hip both at baseline and follow-up and the presence of radiographic osteoarthritic signs at baseline defined by the Kellgren & Lawrence index ≥ grade 1. Of a subset of 754 subjects, CTX-II assessments were available. The baseline measurements were conducted between April 1990 and July 1993, and the follow-up measurements between 1996 and 1999 with a mean follow-up time of 6.6 years. Because our study group had to be mobile enough to visit the research center at baseline and at follow-up, and survived the follow-up period, implies a healthy cohort effect. Compared to the total population of the Rotterdam study, the present study group was younger, had a lower prevalence of lower limb disability at baseline, and a lower prevalence of hip pain at baseline as reported earlier (12).

Radiographic assessment

Weight bearing anteroposterior radiographs of the hip and knee were obtained at 70 KV, a focus of 1.8, and a focus to film distance of 120 cm, applying a Fuji High Resolution G 35 × 43 cm film. Radiographs of the pelvis were obtained with both feet in 10° internal rotation and the X-ray beam centered on the umbilicus, and of the knee with the patellae in central position. For the hand, standard anteroposterior radiographs were obtained. One trained reader (M.R.) evaluated the radiographs of the hip obtained at baseline and at follow-up, unaware of the clinical status of the patients. Three trained readers independently evaluated the baseline radiographs of the knee (E.O. and A.P.B.) and the hand (S.D.), also unaware of the clinical status of the patients. All radiographs of the hip were grouped per patient and read by pairs in chronological order, the order being known to the reader (chronologically ordered reading procedure) (13).

Outcome measure

We defined progression of OA of the hip as a joint space narrowing (JSN) of ≥ 1.0 mm, or an incident total hip replacement (THR) at follow-up. At baseline and follow-up the minimal joint space width (JSW) of the hip joints was measured using a 0.5 millimeters graduated magnifying glass laid directly over the radiograph (14). The lateral, superior and axial compartments of the hip were measured, as described previously by Croft et al.(14). JSN was defined as the JSW of baseline minus the JSW of follow-up (Δ JSW), and a JSN ≥ 1.0 mm of minimal 1 (out of 3) compartment was defined as a progression.

77

Potential determinants of progression

Radiographic determinants

At baseline, radiographic osteoarthritis (ROA) of the hip, knee and hand was quantified by measurements of the Kellgren & Lawrence grading system (15, 16) (atlas-based) in five grades (from zero to four). A person was considered to have ROA of the hip or knee, if the Kellgren & Lawrence score of one or both joints was equal to or larger than two. Hand ROA was defined as the presence of a Kellgren & Lawrence score ≥ grade 2 in at least one joint of two out of the three groups of hand joints (distal interphalangeal, proximal interphalangeal and first carpometacarpal joint group) of one or both hands.

The presence of a superior, axial or medial migration of the femoral head was evaluated to be present or absent.

The inter-rater reliability of the hip was 0.68 for Kellgren & Lawrence (Kappa statistics), and 0.85 (Intraclass Correlation Coefficient) for the minimal JSW, as reported earlier (17). The radiographs of the knee were scored for OA by two independent observers, as described previously (18, 19). For the hand, an inter-rater reliability was reported for Kellgren & Lawrence of 0.68 and 0.77 (Kappa statistics) (20).

Determinants collected by questionnaire

At baseline, trained interviewers conducted an extensive home interview addressing demographic characteristics, medical history, risk factors for chronic diseases and medication use. Lower limb disability was assessed using a modified version of the Stanford Health Assessment Questionnaire (HAQ). A lower limb disability index (LDI) was obtained by calculating the mean score of the answers to six questions, as described previously (range: 0–3) (17, 19). The presence of hip pain ("Did you have joint pain of your right/left hip during the last month") and morning stiffness ("Did you experience morning stiffness of the hips") was asked. Data on age at and type of menopause (spontaneous or artificial) were collected. Menopause was defined as the cessation of menses for at least one year. For women reporting natural menopause, age at menopause was defined as the self-reported age of last menstruation (2, 21). The family history of OA in parents and in siblings was asked. The current or last occupation was asked including the duration in years of this occupation. The jobs were coded according to a job title scheme used at Statistics Netherlands (22). A subject was considered to be exposed to heavy mechanical workload if the subject performed heavy physically demanding work indoors or outdoors and the exposure time of this job was longer than 8 years (3rd and 4th quartile of exposure time).

Determinants collected by physical examination

At the research center, a clinical examination was performed. Height and weight were measured with participants wearing indoor clothing without shoes. Body mass index (BMI)

was calculated as weight in kilograms divided by height in squared meters (kg/m²), and BMI ≥ 30 kg/m² was defined as obesitas. Blood pressure was measured at the right brachial artery using a random-zero sphygmomanometer with the participant in sitting position; the mean of two consecutive measurements was used in the analysis. Hypertension was defined as a systolic blood pressure of 160 mm Hg or higher or a diastolic blood pressure of 95 mm Hg or higher, or current use of antihypertensive drugs for the indication of hypertension (2, 21). Diabetes mellitus was considered present when the subject reported the use of antidiabetic therapy (code A010 of the Anatomical Therapeutical Chemical Classification index, WHO 1992), or when the pre- or post-load serum glucose level was equal to or higher than 11.0 mmol/l (2, 21, 23).

In a subset of 525 subjects, the range of motion was tested. In supine position internal and external rotation, flexion and extension of the hips were tested.

Overnight fasting urine samples were obtained from all subjects at baseline and kept frozen at −20° C. In these samples a cartilage degradation marker, CTX-II (C-telopeptide fragments) was measured by monoclonal antibody mAbF46, and was used in a competitive enzyme-linked immunosorbent assay (ELISA) format, as described previously (24). The concentration of CTX-II (ng/l) was standardized to the total urine creatinine (mmol/l), and the unit for corrected CTX-II concentration was ng/mmol.

Statistical analysis

Of all potential determinants of progression we first performed univariate logistic regression analyses and those determinants with a P-value < 0.1 were used for the multivariate analyses.

For the multivariate analyses we chose a practical approach and in three different models assessed which determinants best identified persons with progression of hip OA.

Model 1

In the first model only those determinants were included which are easily and directly obtainable by the physician such as age, gender, family history of OA, age at menopause, hypertension, diabetes, BMI, mechanical work load, lower limb disability, the presence of hip pain, and morning stiffness.

Model 2

In the second model we added the information obtained from additional radiographic testing; i.e. using the Wald test (cut-off value of P = 0.05) we assessed whether radiographic variables offered additional value to model 1 (with only those variables that are easily and directly obtainable by the physician).

79

Model 3

In the third model we added a cartilage degradation marker (CTX-II) to the model.

To investigate which variables will identify the progressors of hip OA in a clinical situation, we repeated the same procedure for those subjects with existing hip pain at baseline. Pain was considered as a potential marker for symptomatic OA of the hip.

The associations between the potential determinants and progression of the hip were estimated by calculating ORs, by means of Generalized Estimating Equations (GEE). This procedure takes into account the correlation between the left and right hip, using each patient as the observation unit and the hips as repeated measurements (25). A (two-sided) P-value of 0.05 was considered significant. We also calculated Receiver Operating Curves (ROC) of the predicted probabilities of each model for progression of hip OA. Additionally, a clinically useful cut-off point for CTX-II was estimated by ROC analyses. The cut-off point with the best accuracy (i.e. the highest sum of sensitivity and specificity) was used for the development of a prediction rule.

All multivariate analyses were adjusted for follow-up time.

We used SPSS version 11.0 (SPSS Inc., Chicago, USA) and SAS software, version 8.2 (SAS Institute, Cary, NC) for all analyses.

Results

Table 1 presents baseline characteristics of the study population, and the univariate associations with progression of hip OA. In this study population (n = 1,920) with a mean age of 66.1 years, 13.1% had progression of ROA of the hip after a mean follow-up time of 6.6 years. Of these progressors, 38.3% had an incident THR during the follow-up period. During the last month before the baseline interview, 12.7% of all subjects had lower limb disability, 9.5% had hip pain, and 29.8% had morning stiffness. In the univariate analyses the following potential determinants of progression of hip OA had a P-value < 0.1: age, gender, BMI ≥ 30 kg/m², family history of OA, (low) age at menopause, presence of hip pain, lower limb disability, presence of morning stiffness, JSW at baseline, ROA of the hip, ROA of the hand and CTX-II; these determinants were therefore included in the multi-variate analyses.

Table 2 shows the associations between determinants and progression of hip ROA in the total study population for the three models used (see Methods section). In this population-based cohort of men and women aged 55 years and over, the first model (which included easily obtainable variables) showed that age (per year), gender (female), a lower limb disability index of ≥ 0.5 and the presence of hip pain were independent determinants of progression of hip ROA. When radiographic variables were added to this model we found that especially a Kellgren & Lawrence score at baseline of ≥ grade 2 had a strong independent additional (P-value of < 0.0001) association with progression of hip OA, with an OR of 5.6. In model 3,

Table 1: Baseline characteristics of the study population and univariate associations with progression of hip osteoarthritis.

	Study population	OR (95% CI)
Number	1,920	
Gender, % women	51.3	**2.0*** (1.6–2.6)
Age, years ± SD	66.1 ± 6.8	**1.1*** (1.1–1.1)
Body Mass Index, ≥ 30 kg/m², %	12.6	**1.6*** (1.1–2.2)
Diabetes (type II), %	8.9	.9 (.6–1.3)
Hypertension, %	30.0	0.9 (.7–1.1)
Family history of OA, %	9.9	**1.4*** (1.0–2.0)
Age at menopause, (reference group > 50 years)		
− ≤ 45 years	24.6	**1.3*** (0.9–1.9)
− 46–50 years	36.9	**1.6*** (1.1–2.2)
Heavy mechanical workload, %	13.0	.8 (.5–1.1)
Presence of hip pain, %	9.5	**3.4*** (2.5–4.6)
Lower limb disability, %	12.7	**3.2*** (2.5–4.3)
Presence of morning stiffness, %	29.8	**1.7*** (1.3–2.2)
Joint space width at baseline ≤ 2.5 mm, %	7.5	**7.2*** (5.2–9.9)
ROA of the hip, %	12.6	**8.9*** (6.8–11.6)
ROA of the knee, %	16.6	1.0 (.6–1.4)
ROA of the hand, %	23.3	**2.0*** (1.5–2.6)
CTX-II, 4th quartile, %	−	**4.0*** (2.4–6.7)

† Progression of the hip was defined as a joint space narrowing ≥ 1.0 mm or an incident total hip replacement at follow-up.
* P-value of < 0.1

Table 2: Association between determinants and progression of hip osteoarthritis of complete study population in three models (n = 1,920).

	Model 1 (clinical variables)	Model 2 (including radiological variables)	Model 3 (including CTX-II)
Age (years)	**1.07** (1.05–1.08)	**1.06** (1.04–1.08)	**1.03** (1.00–1.06)
Gender	**1.7** (1.3–2.2)	**1.8** (1.3–2.3)	**1.6** (1.0–2.5)
Disability (index score ≥ 0.5)	**1.7** (1.2–2.4)	-	-
Hip pain (presence of)	**2.7** (1.9–3.8)	**2.2** (1.5–3.2)	**2.0** (1.1–3.6)
Baseline JSW (≤ 2.5 mm)	*	**1.9** (1.2–2.9)	**2.1** (1.1–3.8)
Baseline K & L (≥ grade 2)	*	**5.6** (3.9–8.1)	**5.6** (3.3–9.5)
CTX-II (4th quartile)	*	*	**2.2** (1.2–4.1)
Explained variance (R² Nagelkerke)	.104	.220	†

Associations are presented by odds ratios (OR) with 95% confidence interval between parentheses.
Those determinants were included in a model with a P-value < 0.05.
All odds ratios were adjusted for follow-up time.
Progression of the hip was defined as a joint space narrowing ≥ 1.0 mm or an incident total hip replacement at follow-up.
* not tested in this model.
† CTX-II only available in a subset of this population (n = 754).

81
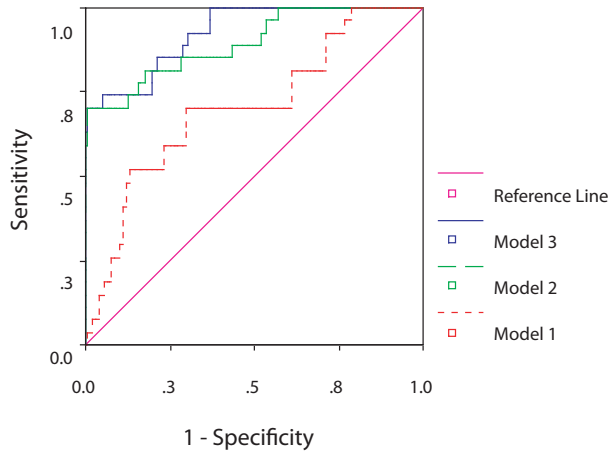
	Area Under the Curve
Model 1	0.703
Model 2	0.895
Model 3	0.926

Figure 1: Receiver operating curves of the predicted probabilities of the three models with progression of hip osteoarthritis for participants with hip pain (N = 195).
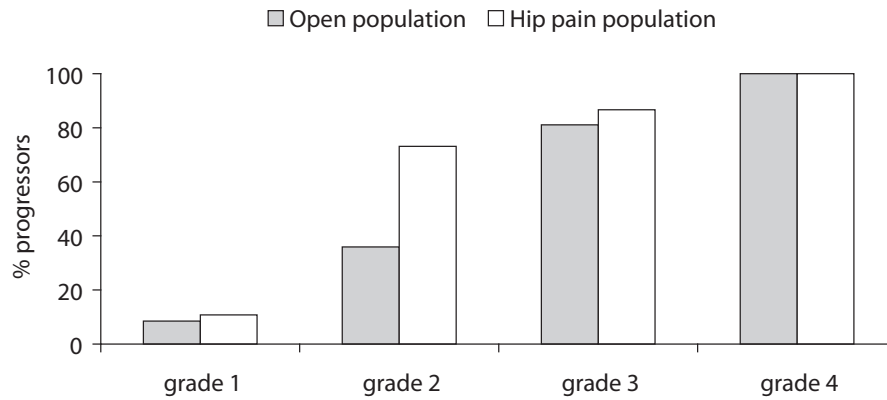

Figure 2: Percentage progressors of hip OA by Kellgren & Lawrence grade at baseline.
Progressors are those subjects with a joint space narrowing ≥ 1.0 mm or the presence of an incident total hip replacement at follow-up.

in a subset of the population (n = 754), we added the information of CTX-II level to model 2 and found that a CTX-II level in the upper quartile also had an independent additional (P-value of 0.001) association with progression of hip ROA, with an OR of 2.2. In the final model, we found that a Kellgren & Lawrence score at baseline of ≥ grade 2 and CTX-II both

Table 3: Association between determinants and progression of hip osteoarthritis for participants with hip pain (n = 411) in three models.

	Model 1 (clinical variables)	Model 2 (inclusive radiological variables)	Model 3 (inclusive CTX-II)
Age (years)	**1.08** (1.03–1.13)	–	–
Disability (index score ≥ 0.5)	**2.5** (1.4–4.4)	–	**4.8** (1.2–19.1)
Baseline K & L (≥ grade 2)	*	**31.7** (10.2–98.3)	**35.6** (11.1–114.2)
CTX-II (≥ 235.5 mmol/l)	*	*	**3.9** (1.3–11.8)
Explained variance (R^2 Nagelkerke)	.123	.504	†

Associations are presented by odds ratios (OR) with 95% confidence interval between parentheses.

Those determinants were included in a model with a P-value < 0.05.

All odds ratios were adjusted for follow-up time and duration of hip pain (longer than 1 year).

Progression of hip osteoarthritis was defined as a joint space narrowing ≥ 1.0 mm and the presence of hip pain at follow-up or an incident total hip replacement at follow-up.

* not tested in this model.

† CTX-II only available in a subset of this population (n = 195).

had an independent association with progression of hip ROA. Furthermore, the lower limb disability index disappeared in the final model.

We repeated the same procedure for those subjects with prevalent hip pain at baseline (n = 411). In this subset of the study population, using by ROC analyses we calculated the best accuracy cut-off point for CTX-II to be ≥ 235.5 mmol/l. We found an impressively stronger association of a baseline Kellgren & Lawrence score ≥ grade 2 with progression of hip ROA in those with initial hip pain with an OR of 35.6 (additional to model 1, P-value < 0.0001 and also a higher area under the curve, Figure 1: 0.895 versus 0.703) (Table 3). CTX-II also showed an independent additional (P-value of < 0.0001 and also a higher area under the curve, Figure 1: 0.926 versus 0.895) association with progression of hip ROA in subjects with initial hip pain. Surprisingly age disappeared in the final model. In addition we repeated all analyses in a subset (n = 525) for whom data on limited range of motion were available. In this subset, we found that a restricted flexion of the hip of more than 20% had an independent association in the final model (OR of 3.1; 95% CI 2.1–4.7) with progression of hip OA. However, the strong additional value of radiographic findings still holds.

Figure 2 shows the percentage progressors and incident THR of the total population and of those with probably symptomatic hip OA, stratified by Kellgren & Lawrence grade at baseline. All subjects with a Kellgren & Lawrence grade 4 at baseline had an incident THR at follow-up. Of the subjects with hip pain and a Kellgren & Lawrence grade 2 at baseline, 73% developed progression during follow-up, compared to 36% in the total study population.

Discussion

In this large population-based prospective cohort study with a long-term follow-up we found that the presence of a Kellgren & Lawrence score of ≥ grade 2 at baseline was the strongest identifier of those persons at high risk for progression of hip OA. This holds particularly for in patients with a prevalent hip pain at baseline. In addition, a lower limb disability index of ≥ 0.5 and a CTX-II concentration ≥ 235.5 mmol/l were also independent identifiers of these high-risk persons.

In the present study we defined progression of hip OA as the presence of a JSN of ≥ 1.0 mm or an incident THR. The choice of how to define progression is arbitrary because there is no consensus about the definition of progression of hip OA. Because JSN is more sensitive to change compared to the Kellgren & Lawrence index (7, 9), we used the above-mentioned definition. We also used an incident THR as a definition of progression of hip OA, based on the assumption that THR is performed in patients whit severe symptomatic OA together with structural damage of the hip. Although a hip fracture is also a reason for THR and may have flawed our results, when we excluded subjects with an incident fracture (n = 16) from the analyses we found similar associations between the risk factors and progression of hip OA.

In the total study population the independent identifiers of the high-risk group for progression of hip OA were age, female gender, the presence of hip pain, JSW at baseline ≤ 2.5 mm, Kellgren & Lawrence score of ≥ grade 2 at baseline, and a high CTX-II level. Of these factors only age and gender are relatively independent factors of the disease, whereas the other predictive factors are signs of the presence or severity of OA. These findings are in agreement with those reported by Lievense et al. in a systematic review (10). In subjects who consulted a general practitioner for hip pain, Birrell et al. showed that a simple scoring system based on both radiographic severity and clinical measures could clearly identify groups at high likelihood of being put on a waiting list for THR (26). Based on the results of the present study and of the two above-mentioned studies, it is clear that progression of hip OA has the strongest associations with signs of the presence or severity of OA, in other words with the disease status of the subject. The absence of an association between BMI and progression of hip OA in the present study was also reported by Lievense et al (10).

It is striking that all other potential determinants of progression that are independent of the disease, were excluded when signs of the presence or severity of hip OA were added to the model. We expected that especially local biomechanical factors (such as mechanical workload and sport activity) would have independent associations with progression of hip OA. A possible explanation for the lack of association in the present study may that we used information about historical workload and not of the workload during the follow-up period. Therefore, we may have missed information about important determinants of progression

of hip OA in our study, such as mechanical load during follow-up of an already existing osteoarthritic joint.

A possible limitation of the present study is the presence of health-based selection bias; overall, participants were generally healthier than non-participants. Therefore, it is likely that the generalizability of the reported findings holds particularly for those subjects who are mobile enough to visit a physician. Furthermore, the reported model should ideally be tested in another population.

Based on the results of the present study, we conclude that a Kellgren & Lawrence score of ≥ grade 2 at baseline is the strongest predictor of progression of hip OA, especially in those with prevalent hip pain at baseline. CTX-II seems to be a moderate predictor of progression of hip OA compared to the variables collected by history taking, physical examination and an X-ray. Overall we conclude that in a clinical situation and for clinical trials an X-ray has strong additional value to identify persons who are at high risk for progression of hip OA.

Acknowledgments

References

1.   Felson DT, Zhang Y. An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. Arthritis Rheum 1998;41(8):1343–55.
2.   Felson DT, et al. Osteoarthritis: New insights. Part 1: The disease and its risk factors. Ann Intern Med 2000;133:635–46.
3.   Felson DT, et al. Osteoarthritis: New insights. Part 2: Treatment approaches. Ann Intern Med 2000;133:726–37.
4.   Altman RD, Hochberg MC, Moskowitz RW, Schnitzer TJ. Recommendations for the medical management of osteoarthritis of the hip and knee. Arthritis Rheum 2000;43:1905–15.
5.   Dougados M, et al. Radiological progression of hip osteoarthritis: definition, risk factors and correlations with clinical status. Ann Rheum Dis 1996;55:356–62.
6.   Altman RD, Brandt KD, Hochberg MC, Moskowitz RM. Design and conduct of clinical trials in patients with osteoarthritis: recommendations from a task force of the Osteoarthritis Research Society. Osteoarthritis Cartilage 1996;4:217–43.

85

7.   GREES. Recommendations for the registration of drugs used in the treatment of osteoarthritis. Ann Rheum Dis 1996;55:552–7.

8.   Dougados M, et al. Requirement for Total Hip Arthroplasty: an outcome measure of hip osteoarthritis? J Rheumatol 1999;26:855–61.

9.   Dougados M. Monitoring osteoarthritis progression and therapy. Osteoarthritis Cartilage 2004;12(suppl.A):S55–60.

10. Lievense AM, Bierma-Zeinstra SMA, Verhagen AP, Verhaar JAN, Koes BW. Prognostic factors of progress of osteoarthritis of the hip: A systematic review. Arthritis Rheum 2002;47(5):556–62.

11. Hofman A, Grobbee DE, Jong PT de, VandenOuweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. Eur J Epidemiol 1991;7(4):403–22.

12. Reijman M, et al. A new marker for osteoarthritis: cross-sectional and longitudinal approach. Arthritis Rheum 2004;50(8):2471–8.

13. Auleley G, Girardeau B, Dougados M, Ravaud P. Radiographic assessment of hip osteoarthritis progression: impact of reading procedures for longitudinal studies. Ann Rheum Dis 2000;59:422–7.

14. Croft P, Cooper C, Wickham C, Coggon D. Defining osteoarthritis of the hip for epidemiologic studies. Am J Epidemiol 1990;132(3):514–22.

15. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494–502.

16. Kellgren JH, Jeffrey MR, Ball J. The epidemiology of chronic rheumatism. Atlas of standard radiographs of arthritis. Oxford: Blackwell Scientific Publications 1963.

17. Reijman M, Hazes JMW, Pols HAP, Bernsen RMD, Koes BW, Bierma-Zeinstra SMA. Validity and reliability of three definitions of hip osteoarthritis: cross-sectional and longitudinal approach. Ann Rheum Dis 2004; in press.

18. Bergink AP, et al. Estrogen receptors alpha gene haplotype is associated with radiographic osteoarthritis of the knee in elderly men and women. Arthritis Rheum 2003;48(7):1913–22.

19. Odding E, Valkenburg HA, Algra D, Vandenouweland FA, Grobbee DE, Hofman A. Associations of radiological osteoarthritis of the hip and knee with locomotor disability in the Rotterdam Study. Ann Rheum Dis 1998;57(4):203–8.

20. Dahaghin S, Bierma-Zeinstra SMA, Reijman M, Pols HAP, Hazes JMW, Koes BW. Prevalence and determinants of one-month hand pain and hand-related disability in the elderly. The Rotterdam Study. Ann Rheum Dis 2004 in press.

21. Felson DT. An update on the pathogenesis and epidemiology of osteoarthritis. Radiol Clin N Am 2004;42:1–9.

22. Standaard Beroepenclassificatie. Voorburg/Heerlen: Centraal Bureau voor de statistiek. 1985.

23. Cerhan JR, Wallace RB, El-Khoury GY, Moore TE. Risk factors for progression to new sites of radiographically defined osteoarthritis in women. J Rheumatol 1996;23:1565–78.

24. Christgau S, et al. Collagen Type II C-telopeptide Fragments as an index of Cartilage Degradation. Bone 2001;29(3):209–15.

25. Zhang Y, Glynn RJ, Felson DT. Musculoskeletal Disease Research: Should we analyse the joint or the person? J Rheumatol 1996;23:1130–4.

26. Birrell F, et al. Predictors of hip joint replacement in new attenders in primary care with hip pain. Br J Gen Pract 2002;53:26–30.

6

Acetabular dysplasia predicts
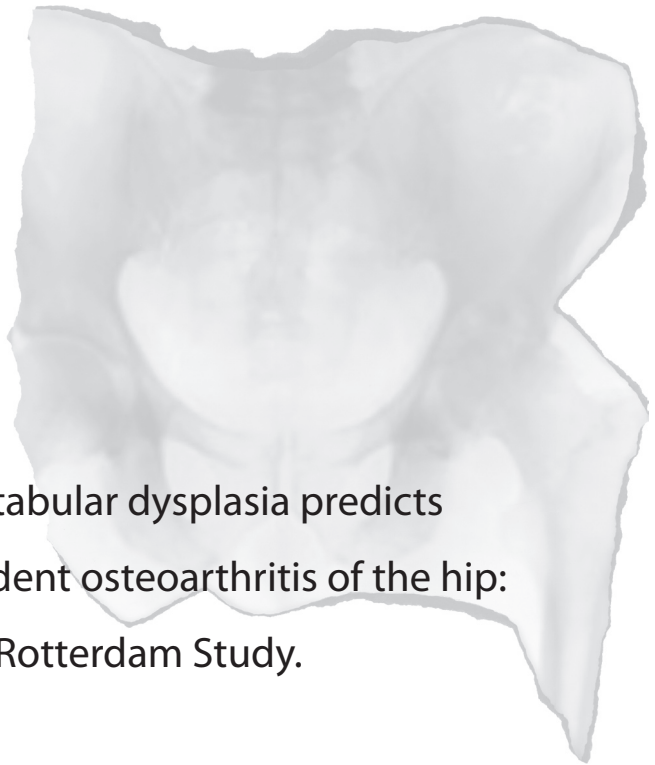
incident osteoarthritis of the hip:

the Rotterdam Study.

Abstract

Objective: To investigate the association between acetabular dysplasia and the incidence of radiographic osteoarthritis (ROA) of the hip, in a population-based sample of elderly subjects.

Methods: In 835 men and women (aged 55 years and older) from the Rotterdam Study X-rays of the hip at baseline and at follow-up (mean follow-up time of 6.6 years) were evaluated. Included were subjects with a baseline Kellgren & Lawrence score of grade 0 or 1 in both hips. Incident hip ROA was defined as a decrease of joint space width of the hip (≥ 1.0 mm) at follow-up. Acetabular dysplasia was assessed using the center-edge angle as defined by Wiberg, and the acetabular depth as defined by Murray. The association between acetabular dysplasia and incident hip ROA was assessed by calculating odds ratios using multivariate regression analysis.

Results: This study population with a mean age of 65.7 ($\pm$ 6.6) years, 9.3% developed an incident hip ROA. Subjects with acetabular dysplasia had a 4.3 times increased risk for incident hip ROA (95% CI; 2.2–8.7) compared to subjects without acetabular dysplasia. These associations were independent of known determinants of hip OA such as age, gender, and BMI, but tended to be enhanced by female gender, heavy mechanical workload, and low body mass index.

Conclusion: In a study population aged 55 years and over, acetabular dysplasia is still a strong independent determinant of incident hip ROA.

Submitted

Introduction

Hip osteoarthritis (OA) is one of the main causes of disability among the elderly and the prevalence will increase with the aging of the Western society (1, 2).

OA is a multifactorial disease involving first of all systemic factors such as metabolism, hormones, genetics, age and gender and, secondly, local biomechanical factors such as mechanical workload, body mass index (BMI) and acetabular dysplasia. Dieppe introduced the model that showed that the joint becomes susceptible for OA by systemic factors and that local biomechanical factors play the final role in determining site and severity of OA (1, 3). Marked acetabular dysplasia is a well-known cause of premature hip OA (4, 5), whereas the influence of a mild acetabular dysplasia on the development of hip OA is less clear. It has been proposed that in some patients with primary hip OA, the disease occurs as a consequence of a mild acetabular dysplasia that persists into adult life. Support for this theory comes from radiological observations in patients with OA of the hip (6), and from follow-up studies of subjects with dysplastic hips (7). A recent review (8) of the available literature on the influence of hip dysplasia on the development of hip OA, revealed that only one study investigated the influence of this parameter in a prospective cohort design with a long follow-up period (9). The authors showed that acetabular dysplasia is associated with a modestly increased risk of incident hip OA in elderly white women. Whether the influence of acetabular dysplasia on the development of hip OA is modified by other known determinants (such as gender, BMI or mechanical load) is not yet known.

Therefore, the present study investigated the association between radiographic evidence of acetabular dysplasia in participants without radiological OA of the hip at baseline and an incident hip OA, in a large population of men and women aged 55 years and over with a long-term follow-up. We also investigated whether the association between acetabular dysplasia and incident hip OA could be modified by other determinants of hip OA.

Subjects and Methods

The study population consisted of participants of the Rotterdam Study, a prospective cohort of men and women aged 55 years and over. The Rotterdam Study investigates the incidence of, and risk factors for, chronic disabling diseases; the rationale and study design have been described previously (10). Written informed consent was obtained from each participant. The Medical Ethics Committee of the Erasmus Medical Center has approved the Rotterdam Study.

The present study used a selected sample of 875 subjects from the Rotterdam study, based on the availability of radiographs of the hip both at baseline and at follow-up. Only those participants with a Kellgren & Lawrence score at baseline of grade 0 or 1 in both hips were

included. Forty participants were excluded because of a hip fracture during the follow-up period, resulting in a final study population of 835 subjects.

Baseline measurements were conducted between April 1990 and July 1993 and the follow-up measurements between 1996 and 1999, with a mean follow-up time of 6.6 years.

Radiographic assessment

Weight bearing anteroposterior radiographs of the hip were obtained at 70 KV, a focus of 1.8, and a focus to film distance of 120 cm, applying a Fuji High Resolution G 35 × 43 cm film. Radiographs of the pelvis were obtained with both feet in 10° internal rotation and the X-ray beam centered on the umbilicus. One independent trained reader (MR) evaluated the radiographs of the hip made at baseline and at follow-up, unaware of the clinical status of the patients. All radiographs of the hip were grouped per patient and read by pairs in chronological order, the order being known to the reader (chronologically ordered reading procedure) (11). At baseline, radiographic OA (ROA) of the hip was quantified by measurements of the Kellgren & Lawrence grading system (12, 13) (atlas-based) in five grades (from zero to four). A person was considered to have ROA of the hip if the Kellgren & Lawrence score of one or both joints was equal to or larger than two. Those persons with ROA of the hip at baseline were excluded from this study.

Outcome measure

In the present study two definitions of an incident hip ROA were used: firstly, defined by a joint space narrowing JSN) ≥ 1.0 mm, and secondly defined as a Kellgren & Lawrence score of ≥ grade 2 at follow-up.

The minimal joint space width (JSW) of the hip at baseline and follow-up was measured using a 0.5 millimetres graduated magnifying glass laid directly over the radiograph (14). The lateral, superior and axial compartments of the hip were measured as described previously by Croft et al. (14). JSN was defined as the JSW at baseline minus the JSW at follow-up, and a JSN of minimally 1 (out of 3) compartment was defined as a radiological osteoarthritic change.

Acetabular dysplasia

Acetabular dysplasia was assessed using the center-edge (CE) angle as defined by Wiberg (15), and the acetabular depth as defined by Murray (16). A period of a few months elapsed between the evaluation of the radiographs of the hip and the assessment of the acetabular dysplasia. The trained reader assessed the CE angle and the acetabular depth unaware of the outcome status of the subjects. The CE angle was defined as the angle formed by a line from the center of the femoral head to the lateral margin of the acetabular roof, and a line perpendicular to that joining the centers of the two femoral heads (Figure 1). The centers
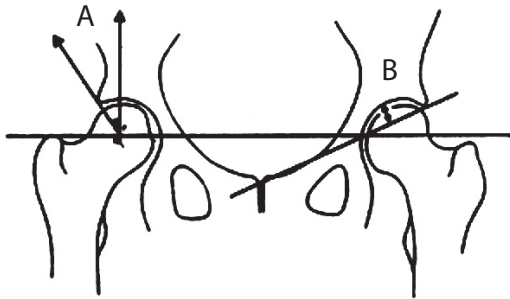
91

Figure 1: Diagram showing the center edge angle and acetabular depth measurements.
A = center edge angle; B = acetabular depth.

of the femoral heads were located with the aid of a transparent plastic sheet marked with concentric circles. The CE angles were measured using a transparent plastic protractor.

Acetabular depth was defined as the greatest perpendicular distance from the acetabular roof to a line joining the lateral margin of the acetabular roof and the upper corner of the symphysis pubis on the same side (Figure 1).

Two independent readers (MR and SMAB) tested the inter-rater reliability of the measurements of the CE angle and the acetabular depth in a random subset of 105 radiographs. The inter-rater reliability for Kellgren & Lawrence (Kappa statistics of 0.68) and for the JSW (Intraclass Correlation Coefficient, ICC of 0.85) has been tested and reported earlier (17). For the CE angle the ICC was 0.90 and for the acetabular depth the ICC was 0.89.

Clinical measures

At baseline, trained interviewers conducted an extensive home interview addressing demographic characteristics, medical history, risk factors for chronic diseases and medication use. The presence of hip pain ("Did you have joint pain of your right/left hip during the last month") was asked. The current or last occupation was asked including the duration in years of this occupation. The jobs were coded according to a job title scheme, used at Statistics Netherlands (18). A subject was considered to be exposed to heavy mechanical workload if that person performed heavy physically demanding work indoors or outdoors and the exposure time of this job was longer than 8 years (3rd and 4th quartile of exposure time).

At the research center, a clinical examination was performed. Amongst other measurements, height and weight were measured with participants wearing indoor clothing without shoes. BMI was calculated as weight in kilograms divided by height in squared meters (kg/m^2).

Statistical analysis

Differences in baseline characteristics were evaluated by analysis of variance (ANOVA) for continuous variables and by a Chi-square test for categorical variables.

Associations between baseline measurements of acetabular dysplasia and the two definitions of incident hip ROA were assessed using logistic regression analysis to calculate odds ratios (ORs), by means of Generalized Estimating Equations (GEE). This procedure takes into account the correlation between the left and right hip, while using each joint (left or right) as the observation unit (19).

To assess the association between acetabular dysplasia and incident hip ROA, we calculated crude ORs as well as ORs adjusted for age, gender, BMI and follow-up time. To investigate whether the associations between baseline measurements of acetabular dysplasia and an incident hip ROA were modified by gender or by mechanical load, we stratified these associations for gender, heavy workload and BMI (≥ 27 kg/m^2) at baseline. In addition, we stratified for the presence or absence of hip pain in order to investigate whether acetabular dysplasia is associated with incident hip ROA for subjects with hip pain. To establish the above-mentioned stratifications differed significantly, formal testing was applied using the formula: $Z = \beta1 - \beta2 / \sqrt{(SE \beta1)^2 + (SE \beta2)^2}$; where $\beta1$ and $\beta2$ stand for the Beta value of strata 1 and 2, and SE $\beta1$ and SE $\beta2$ for the standard error of Beta 1 and 2.

A (two-sided) P-value of 0.05 was considered significant in all analyses. We estimated the magnitude of confounding by the degree of discrepancy between the unadjusted and adjusted estimate (the change-in-estimate criterion) (20). A cut-off point of 10% was chosen to designate an important change in the estimate. SPSS version 11.0 (SPSS Inc., Chicago, USA) and SAS software, version 8.2 (SAS Institute, Cary, NC) were used for all analyses.

Results

Table 1 presents the baseline characteristics of the total study population stratified for the absence or presence of a JSN ≥ 1.0 mm. In this study population (with a mean age of 65.6 years), 9.3% developed JSN during the follow-up period. Subjects with JSN were 2.9 years older and more often female (68.9% versus 55.9%) compared to those without JSN. The mean CE angle in this population was 35.1° ($\pm$ SD 5.6°) and the mean acetabular depth was 12.2 ($\pm$ SD 2.8) mm.

Table 2 shows the association between baseline acetabular dysplasia and an incident hip ROA, as defined by a JSN ≥ 1.0 mm. The crude data showed a strong association between all acetabular dysplasia measurements and JSN of the hip. After adjustment for age, gender, BMI and follow-up time the risk estimates did not change substantially.

During the follow-up period, 16.9% developed an incident ROA of the hip as defined by Kellgren & Lawrence. For the association between baseline acetabular dysplasia and an incident ROA of the hip, as defined by Kellgren & Lawrence, we found lower but significant ORs (Table 3). After adjustment for age, gender, BMI and follow-up time these risk estimates

93

Table 1: Baseline characteristics of the study population, stratified by the absence or presence of an incident hip osteoarthritis (OA) (joint space narrowing ≥ 1.0 mm) at follow-up.

	Total n = 835	Absence of incident hip OA n = 757	Presence of incident hip OA n = 78
Age, years ± SD	65.6 ± 6.5	65.3 ± 6.5 ***	68.2 ± 6.6
Gender, % women	57.2	55.9 **	68.9
BMI, kg/m^2 ± SD	26.5 ±3.5	26.4 ± 3.4	26.6 ± 4.0
Heavy mechanical workload, %	12.2	12.5	9.9
CE angle			
<30°, %	19.2	17.9	31.0
<25°, %	4.8	3.8	12.6
Acetabular depth < 9 mm, %	12.0	10.7	23.0

Significant differences between persons with and without incident hip OA (joint space narrowing ≥ 1.0 mm) at follow-up.
*P < 0.05, ** P < 0.001, *** P < 0.0001

Table 2: Association between acetabular dysplasia and a joint space narrowing ≥ 1.0 mm of the hip.

	Crude OR	Adjusted OR
CE angle < 30°	**2.4** (1.6–3.5)	**2.8** (1.9–4.2)
CE angle < 25°	**4.1** (2.1–7.9)	**4.3** (2.2–8.7)
Acetabular depth < 9 mm	**2.8** (1.8–4.4)	**2.8** (1.8–4.5)

Associations are presented by odds ratios (OR) with 95% confidence interval between parentheses.
Odds ratios were adjusted for age, gender, BMI and follow-up time.
Joint space narrowing ≥ 1.0 mm was used as definition of an incident hip osteoarthritis.

Table 3: Association between acetabular dysplasia and incident hip (by Kellgren & Lawrence ≥ grade 2 at follow-up) osteoarthritis.

	Crude OR	Adjusted OR
CE angle < 30°	**1.6** (1.1–2.4)	**1.7** (1.2–2.5)
CE angle < 25°	**2.3** (1.2–4.5)	**2.4** (1.2–4.7)
Acetabular depth < 9 mm	**2.3** (1.5–3.5)	**2.3** (1.5–3.5)

Associations are presented by odds ratios (OR) with 95% confidence interval between parentheses.
Odds ratios were adjusted for age, gender, BMI and follow-up time.
Incident hip osteoarthritis is defined by baseline Kellgren & Lawrence ≤ grade 1, and follow-up Kellgren & Lawrence ≥ grade 2.

did not change substantially. This also applies for any definition of acetabular dysplasia as used in the present study.

The strong associations between acetabular dysplasia and incident hip ROA in the present study could be flawed by the pre-existence of a Kellgren & Lawrence grade 1 at baseline. Therefore we repeated the analyses separately for subjects with a baseline Kellgren & Lawrence score of 0 and for those with a baseline score of 1. Surprisingly, we found stronger associations for those with a Kellgren & Lawrence grade 0 at baseline compared to those with a Kellgren & Lawrence grade 1: 2.9 (95% CI; 1.4–6.0) and 1.5 (95% CI; .9–2.5), respectively. For an incident hip ROA defined by JSN, we found similar associations for both these Kellgren & Lawrence grades.

Table 4: Association between acetabular dysplasia and joint space narrowing stratified by gender, heavy mechanical workload, BMI, or the presence of hip pain.

| | CE angle | | Acetabular depth |
	< 25°	< 30°	< 9 mm
Overall	**4.5** (2.4–8.7)	**2.6** (1.8–3.9)	**2.6** (1.7–4.1)
Gender			
– Women	**5.9** (2.6–13.5)	**2.8** (1.8–4.4)	**3.2** (1.9–5.4)
– men	**2.6** (.8–8.6)	**2.2** (1.1–4.6)	**1.8** (.7–4.5)
Workload			
– low	**4.0** (1.9–8.4)	**2.6** (1.8–3.9)	**2.7** (1.7–4.3)
– heavy	**8.0** (1.5–42.4)	**2.4** (.6–9.9)	**2.0** (.4–11.2)
BMI			
– < 27 kg/m^2	**6.2** (2.8–13.9)	**3.4** (2.1–5.4)	**3.9** (2.2–6.9)*
– ≥ 27 kg/m^2	**2.5** (.8–8.2)	**1.8** (.9–3.4)	**1.6** (.7–3.3)
Hip pain			
– absent	**4.1** (2.1–7.8)	**2.3** (1.5–3.5)*	**2.2** (1.3–3.7)
– present	†	**7.0** (2.0–24.6)	**5.2** (1.6–16.9)

Associations are presented by odds ratios (OR) with 95% confidence interval between parentheses.
Odds ratios were adjusted for age, gender, BMI and follow-up time.
Joint space narrowing ≥ 1.0 mm was used as definition of an incident hip osteoarthritis (first definition of incident hip OA).
† The associations between a CE angle < 25° and the occurrence of radiological osteoarthritic changes of the hip stratified for the presence of hip pain, was not estimable because of empty cells (100% of the subjects with hip pain and a CE angle of < 25° developed radiological osteoarthritic changes of the hip during follow-up period).
* Significant difference with a P-value < 0.05.

Table 4 shows the association between baseline acetabular dysplasia and JSN stratified by gender, heavy workload, BMI and the presence of hip pain. Women had a stronger association between acetabular dysplasia (as defined by CE angle) and JSN compared to men. Those who had performed heavy physically demanding work had stronger associations between acetabular dysplasia and JSN, compared to those who had performed low physically demanding work. Surprisingly, we found that persons with a low BMI (< 27 kg/m^2) at baseline had stronger associations between acetabular dysplasia and JSN, compared to those with a high BMI (≥ 27 kg/m^2) at baseline. Persons with a prevalent hip pain at baseline had stronger associations between acetabular dysplasia and JSN, compared to those without a prevalent hip pain. The differences in associations did not reach significance, except for the stratification for overweight and non-overweight subjects with acetabular depth < 9 mm, and the stratification for the absence and presence of hip pain with a CE angle < 30°. Of all the subjects with hip pain at baseline, 8.7% developed an incident hip OA during follow-up period. If these subjects had also a CE angle < 30° or an acetabular depth < 9 mm this percentage increased to respectively 40.9% and 47.4%.

Discussion

In our study population of men and women aged 55 years and over, with no signs of ROA of the hip at baseline, we found that acetabular dysplasia, even in this elderly population, is a strong independent determinant of incident hip ROA. The association between acetabular dysplasia and incident hip ROA tends to be enhanced by female gender and mechanical workload.

In our study, the association in women was similar to that reported by Lane et al. (9) who investigated the association between acetabular dysplasia and incident hip ROA in women in a setting similar to ours. In the present study the prevalence of acetabular dysplasia was similar for men and women, but women more often developed JSN during follow-up compared to men (12.8% versus 6.8%, respectively). The reason why the association between acetabular dysplasia and incident hip ROA seems to be modified by gender is unclear. Different alignment of the lower extremity in women (21) and consequently another mechanical loading of the hip joint might be an explanation. To investigate whether high mechanical load of the joint is a modifier of the association between acetabular dysplasia and incident hip ROA we stratified for heavy workload, and found an indication for a stronger association for persons who had performed heavy physically demanding work compared to those who had performed low physically demanding work. A possible mechanism to explain the association between acetabular dysplasia and hip OA is that the presence of a subtle biomechanical abnormality, secondary to either joint incongruity (smaller acetabular depth) or decreased joint surface area (smaller CE angle), may increase joint stresses in the superolateral acetabular rim (22, 23). Hence, it seems plausible that a high mechanical load can modify this association. In an earlier study we found that heavy workload itself had no independent association with progression of hip OA as defined by JSN or an incident total hip replacement. However, the results of the present study suggest that high mechanical load might be associated with incident hip ROA, but only in subjects with acetabular dysplasia. We also investigated whether BMI is a modifier of the association between mild acetabular dysplasia and hip OA and surprisingly, found no stronger associations for persons with a low BMI (< 27 kg/m^2) at baseline compared to those with a high BMI ($\geq$ 27 kg/m^2) at baseline. A lower activity level in those with a high BMI might explain the lack of association between acetabular dysplasia and incident hip ROA in persons with a high BMI. We did indeed find that persons with a high BMI ($\geq$ 27 kg/m^2) were more often disabled at the lower limb versus those with a low BMI (15.6% versus 7.6%). In addition, BMI as such may not represent a high mechanical overload for the hip joints. Finally, we investigated whether acetabular dysplasia is associated with incident hip ROA for those with hip pain. The results suggest that for a person with hip pain at baseline but without the radiological evidence of hip OA, the presence of mild acetabular dysplasia is associated with an incident hip ROA during the follow-up period, even more than in subjects without hip pain.

Because a prevalent radiological hip ROA may alter hip geometry, such that the true prevalence of acetabular dysplasia cannot be assessed, it will always be difficult to investigate the association between mild acetabular dysplasia and prevalent hip ROA or the progression of ROA. In the present study we included only those subjects with a baseline Kellgren & Lawrence score of grade 0 or 1 in both hips. Nevertheless, associations between acetabular dysplasia and hip ROA have been suggested to be flawed by the prior existence of radiographic osteoarthritic signs at baseline, such as osteophytes at the lateral acetabular margin and medial migration of the femoral head (24). However, the results of the present study suggest that the reported strong association between acetabular dysplasia and incident hip ROA is even slightly stronger in those subjects without radiographic osteoarthritic signs of (grade 0) hip ROA at baseline.

Because there is no consensus on how to define incident hip ROA, in the present study we used two definitions, namely the presence of a JSN ≥ 1.0 mm and, secondly, the presence of a Kellgren & Lawrence ≥ grade 2 at follow-up. Strikingly, the overlap between the incident cases defined by both definitions was low. Of the incident cases defined by Kellgren & Lawrence, only 37.7% was defined by a JSN. The problems of using "incident cases" in a longitudinal approach were very clear described by Lohmander and Felson (25). "The distinction between incident cases of OA and progression of prevalent cases depends on where along the continuum patients are considered to have overt OA". In the present study we found that independent of which definition was used for incident hip ROA, the strong associations with acetabular dysplasia holds. The fact that these strong associations were independent of the definition used and even stronger for those with a Kellgren & Lawrence grade 0 at baseline makes the results of the present study even more convincing.

Furthermore, the associations between acetabular dysplasia and incident ROA of the hip may even be underestimated because of the relatively high mean age of the study population (65.7 ± 6.6 years). In other words, we assume that in a younger population the association between acetabular dysplasia and OA may be even higher.

Based on the results of the present study, we conclude that acetabular dysplasia is strongly associated with an incident hip ROA even in a population of men and women aged 55 years and over, and that these associations are independent of known risk factors for ROA of the hip. Furthermore, these associations might be enhanced by female gender, heavy mechanical workload and low BMI.

Acknowledgments

pharmacists, the many field workers at the research center in Ommoord and of course all participants.

References

1. Felson DT, Zhang Y. An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. Arthritis Rheum 1998;41(8):1343–55.
2. Felson DT, et al. Osteoarthritis: New insights. Part 1: The disease and its risk factors. Ann Intern Med 2000;133:635–46.
3. Dieppe P. The classification and diagnosis of osteoarthritis. In: Kuettner KE, Goldberg WM, editors. Osteoarthritic Disorders. Rosemont, IL: American Academy of Orthopaedic Surgeons. 1995:5–12.
4. Wedge JH, Wasylenko MJ. The natural history of congenital dislocation of the hip: a critical review. Clin Orthop 1978;137:154–62.
5. Weinstein SL. Natural history of congenital hip dislocation (CDH) and hip dysplasia. Clin Orthop 1987;225:62–76.
6. Murphy SB, Ganz R, Muller ME. The prognosis in untreated dysplasia of the hip. J Bone Joint Surg 1995;77(7):985–9.
7. Cooperman DR, Wallensten R, Stulberg SD. Acetabular dysplasia in the adult. Clin Orthop 1983;175:79–85.
8. Lievense AM, Bierma-Zeinstra SMA, Verhagen AP, Verhaar JAN, Koes BW. Influence of hip dysplasia on the development of osteoarthritis of the hip: A systematic review. Ann Rheum Dis 2004;63(6):621–6.
9. Lane NE, et al. Association of mild acetabular dysplasia with an increased risk of incident hip osteoarthritis in elderly white women. Arthritis Rheum 2000;43(2):400–4.
10. Hofman A, Grobbee DE, Jong PTde, Ouweland FA van den. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. Eur J Epidemiol 1991;7(4):403–22.
11. Auleley G, Girardeau B, Dougados M, Ravaud P. Radiographic assessment of hip osteoarthritis progression: impact of reading procedures for longitudinal studies. Ann Rheum Dis 2000;59:422–7.
12. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494–502.
13. Kellgren JH, Jeffrey MR, Ball J. The epidemiology of chronic rheumatism. Atlas of standard radiographs of arthritis. Oxford: Blackwell Scientific Publications 1963.
14. Croft P, Cooper C, Wickham C, Coggon D. Defining osteoarthritis of the hip for epidemiologic studies. Am J Epidemiol 1990;132(3):514–22.
15. Wiberg G. A measuring method for distinguishing between a normal and a maldeveloped acetabulum. Acta Chir Scand 1939;83 (suppl):28–38.
16. Murray RO. The aetiology of primary osteoarthritis of the hip. Br J Radiol 1965;38:810–24.
17. Reijman M, Hazes JMW, Pols HAP, Bernsen RMD, Koes BW, Bierma-Zeinstra SMA. Validity and reliability of three definitions of hip osteoarthritis: cross-sectional and longitudinal approach. Ann Rheum Dis 2004; in press.
18. Standaard Beroepenclassificatie. Voorburg/Heerlen: Centraal Bureau voor de statistiek. 1985.
19. Zhang Y, Glynn RJ, Felson DT. Musculoskeletal disease research: Should we analyse the joint or the person? J Rheumatol 1996;23:1130–4.
20. Rothman KJ, Greenland S. Introduction to stratified analysis. In: Rothman KJ, Greenland S, editors. Modern epidemiology. Second edition. Boston/Toronto: Little, Brown and Company 1998:311–26.
21. Laforgia R, Specchiulli F, Solarino G, Niti Luigi. Radiographic variables in normal and osteoarthritic hips. Bull Hospital Joint Dis 1996;54(4):215–21.

22.   Maxian TA, Brown TD, Weinstein SL. Chronic stress tolerance levels for human articular cartilage: two nonuniform contact models applied to long-term follow-up of CDH. J Biomech 1995;28(2):159–66.
23.   Hadley NA, Brown TD, Weinstein SL. The effects of contact pressure elevations and aseptic necrosis on the long-term outcome of congenital hip dislocation. J Orthop Res 1990;8(4):504–13.
24.   Croft P, Cooper C, Wickham C, Coggon D. Osteoarthritis of the hip and acetabular dysplasia. Ann Rheum Dis 1991;50:308–10.
25.   Lohmander LS, Felson DT. Can we identify a 'high risk' patient profile to determine who will experience rapid progression of osteoarthritis? Osteoarthritis Cartilage 2004;12(suppl A):S49–52.

7

Is there an association between
the use of different types
of NSAIDs and radiologic
progression of osteoarthritis?
The Rotterdam Study.

Abstract

Objective: To investigate the influence of the use of various types of NSAIDs on progression of OA of the hip and knee.

Methods: In 1,695 (for the hip) and 635 (for the knee) men and women aged 55 years and older from the Rotterdam Study (a population-based cohort study), X-rays of the hip and knee at baseline and follow-up (mean follow-up time of 6.6 years) were evaluated. Radiologic progression (ROA) was defined as a decrease of joint space width (hip $\geq$ 1.0 mm, knee $\geq$ 1.5 mm) or incident joint replacement at follow-up. NSAIDs were divided in those (indomethacin, naproxen and ibuprofen) which are supposed to have a deleterious effect on joint cartilage, and in a second group of NSAIDs (such as diclofenac and piroxicam) which are supposed to have a neutral effect on cartilage metabolism. The associations between the different types of NSAIDs and progression of ROA were assessed using multivariate logistic regression analysis.

Results: Those subjects who used NSAIDs supposed to have a neutral effect on cartilage metabolism (> 31 days) surprisingly had a 1.7-increased risk (95% CI, 1.0–2.9) for hip ROA, compared to the short-term user (1–30 days). This increased risk was mainly due to the long-term use (> 180 days) of diclofenac. No clear associations were found between the different types of NSAID and progression of knee ROA.

Conclusion: These data suggest that diclofenac may induce accelerated progression of hip OA. Whether this occurs due to a real deleterious effect on cartilage or due to excessive mechanical loading on an analgesic hip remains to be investigated.

Introduction

Osteoarthritis (OA) is a common age-related locomotor disease characterized by degradation of articular cartilage. OA of the hip and knee can be especially disabling because of the related pain and functional impairment, which results from the involvement of a large weight-bearing joint (1). Non-steroidal anti-inflammatory drugs (NSAIDs) have been widely used as a pharmacologic treatment to relieve pain in patients with OA (1–3). However, these agents are associated with major and minor side effects, particularly in the elderly population (1). Also, the rate of progression of OA might be negatively influenced by NSAIDs (4). Several in-vitro studies of human cartilage (5–7) and also animal studies (8–11) suggest that some types of NSAIDs inhibit the synthesis of articular cartilage metabolism. Based on the results of these studies, NSAIDs have been divided in those (indomethacin, naproxen and ibuprofen) which are supposed to have a deleterious effect on joint cartilage by inhibition of glucosaminoglycans (GAG) synthesis, and in a second group of NSAIDs (such as diclofenac and piroxicam) which are supposed to have a neutral effect on cartilage metabolism (12).

There are also clinical reports of an increased rate of progression of OA in patients receiving NSAIDs (13–15). One study concerned patients who used indomethacin and reached earlier the end point of OA (total hip replacement), but only in a selected group of patients with end stage OA awaiting surgery (14). The study of Huskisson et al. (13), a long-term prospective study, demonstrated in 812 patients with knee OA that those who used indomethacin had an increased rate of joint space narrowing compared to those who used tiaprofenic acid. However, a third patient-based study reported that naproxen had no toxic effect on osteoarthritic cartilage; this study, however, included only a small number of patients (15). Therefore, the two above-mentioned groups of NSAIDs have not been evaluated for a possible deleterious effect in a large population.

We investigated in the present study the associations between the two above-mentioned groups of NSAIDs and progression of OA of the hip and knee in a large open population of men and women aged 55 years and over. Additionally, we investigated the associations between the individual types of NSAIDs and progression of OA.

Subjects and Methods

The study population consisted of participants of the Rotterdam Study, a prospective cohort of men and women aged 55 years and over. The objective of the Rotterdam Study is to investigate the incidence of, and risk factors for, chronic disabling diseases. The rationale and study design have been described previously (16). The focus is on neurogeriatric, cardiovascular, ophthalmologic and locomotor diseases. All 10,275 inhabitants of Ommoord

(a district in Rotterdam, the Netherlands) were invited to participate. The response rate was 78%, resulting in 7,983 subjects participating in the study. Written informed consent was obtained from each participant. The Medical Ethics Committee of the Erasmus Medical Center has approved the Rotterdam Study.

For the present study a sample of 3,585 subjects of the Rotterdam study was used. The selection was based on the availability of the radiographs of the hip and knee at baseline and follow-up. Subjects with bilateral total hip replacement at baseline (n = 24) were excluded from analysis. At baseline there were no subjects with bilateral total knee replacement. The hypothesis that the rate of progression of OA is influenced by NSAIDs holds for osteoarthritic cartilage and not for normal cartilage (4). For the analyses of the hip we included those subjects with at least minimal osteoarthritic signs on the radiograph, defined by a Kellgren & Lawrence score of the hip of grade 1 or higher (in at least one joint), resulting in a study population of 1,695 subjects. For the analyses of the knee we included 635 subjects, since fewer numbers of radiographs of the knee at baseline and at follow-up were available.

The baseline measurements were conducted between April 1990 and July 1993, and the follow-up measurements between 1996 and 1999, with a mean follow-up time of 6.6 years. The fact that subjects had to be mobile enough to visit the research center at baseline and follow-up, and survived the follow-up time, caused a healthy cohort effect in our study population. Compared to the total Rotterdam study population, the present study population was younger, had a lower prevalence of lower limb disability at baseline and a lower prevalence of hip pain at baseline as reported earlier (17).

Radiographic assessment

Weight bearing anteroposterior radiographs of the hip and knee were obtained at 70 KV, a focus of 1.8, and a focus to film distance of 120 cm, applying a Fuji High Resolution G 35 × 43 cm film. Radiographs of the pelvis were obtained with both feet in 10° internal rotation and the X-ray beam centred on the umbilicus, and of the knee with the patellae in central position. Two trained readers independently evaluated the radiographs of the hip and knee at baseline and follow-up, unaware of the clinical status of the participants. All radiographs were grouped by participant and read by pairs in chronological order, the order being known to the reader (chronologically ordered reading procedure) (18). At baseline, radiological OA (ROA) of the hip and knee was quantified by measurements following the Kellgren & Lawrence grading system (19, 20) (atlas-based) in five grades (from zero to four).

We defined progression of the hip as a joint space narrowing (JSN) of ≥ 1.0 mm of minimal 1 (out of 3) compartment or an incident total hip replacement at follow-up (21). Progression of the knee was defined as a JSN of ≥ 1.5 mm of minimal 1 (out of 2) compartment or an incident total knee replacement at follow-up (21).

At baseline and follow-up the minimal joint space width (JSW) of the hip and knee joints were measured using a 0.5 millimetres graduated magnifying glass directly laid over the radiograph (22). For the hip the lateral, superior and axial compartment was measured and for the knee the medial and lateral compartment, as described previously (22, 23).

The inter-rater reliability for the Kellgren & Lawrence score of the hip was 0.68 (Kappa statistics), and for the minimal JSW 0.85 (Intraclass Correlation Coefficient), as reported earlier (17). The radiographs of the knee were scored for ROA by two independent observers, as described previously (23, 24). After each set of 150 radiographs, the scores of the two readers were evaluated. Whenever the Kellgren & Lawrence score differed, the two readers met to read the radiographs together, and a consensus score was determined.

Use of NSAIDs

Data on medication prescription were derived from the pharmacies in Ommoord. These pharmacies were fully automated and registered all prescriptions on drug use from January 1, 1991 through December 31, 1998. Prescriptions included the product name of the drug, the generic name, the Anatomical Therapeutic Chemical (ATC) code, the number of tablets, capsules or other vehicles in the filled prescription, the date of delivery of the product, the prescribed daily number of tablets to be taken, the daily drug dosage, and the duration of the prescription. Thus, for all NSAIDs prescriptions we had data on date of delivery of NSAIDs, duration and dosage of NSAID as well as type of NSAID.

Potential confounders

At baseline, trained interviewers performed an extensive home interview on demographic characteristics, medical history, risk factors for chronic diseases and medication use. Height and weight were measured with participants wearing indoor clothing without shoes. Body mass index (BMI) was calculated as weight in kilograms divided by height in squared meters (kg/m^2).

Statistical analysis

Differences in baseline characteristics were evaluated by analysis of variance (ANOVA) for continuous variables and Chi-square for categorical variables. The baseline characteristics were stratified for the presence or absence of progression of hip or knee ROA.

All prescribed NSAIDs were divided in two groups, those (indomethacin, naproxen and ibuprofen) that are supposed to have a deleterious effect on joint cartilage, and those NSAIDs (such as diclofenac and piroxicam) that are supposed to have a neutral effect on cartilage metabolism.

The associations between NSAIDs and progression of hip and knee OA were investigated in a stepwise procedure. Firstly, we assessed the association between the two groups of NSAIDs, those that are supposed to have a deleterious effect on joint cartilage and those

that are supposed to have a neutral effect on cartilage metabolism, and progression of OA. Those subjects who used NSAIDs from both groups were excluded from analysis. The duration of NSAID use was categorized in two groups; the short-term user (1–30 days), and the long-term user (longer than 31 days). For these analyses we used the 1–30 days users as reference group. Additionally, we investigated the associations between the individual types of NSAIDs (namely ibuprofen, naproxen, diclofenac and piroxicam) and progression of the hip. The numbers of users of indomethacin were too small to allow a subgroup analysis.

The associations between the use of NSAIDs and progression of ROA of the hip or knee were assessed using logistic regression analysis to calculate odds ratios (OR) as estimation for relative progression risk. All associations were adjusted for age, gender, BMI, baseline ROA (Kellgren & Lawrence dichotomized ≥ grade 2), follow-up time, the defined daily dosage (actual dosage/recommended dosage for an adult (WHO)), and the duration of use (continuous variable). A (two-sided) P-value of 0.05 was considered significant. We estimated the magnitude of confounding by the degree of discrepancy between the unadjusted and adjusted estimate (the change-in-estimate-criterion) (25). We chose a cut-off point of 10% to designate an important change in the estimate. SPSS version 11.0 (SPSS Inc., Chicago, USA) was used for all analyses.

Results

The baseline characteristics stratified for baseline Kellgren & Lawrence ≥ grade 1 of the hip and knee and for the presence or absence of progression of hip and knee ROA are shown in Table 1. Of the subjects with a Kellgren & Lawrence index ≥ grade 1 at baseline, 11.9% showed progression of ROA of the hip and 8.7% progression of ROA of the knee at follow-up. Of the hip progressors, 4.5% (n = 77) had an incident joint replacement during the follow-up period and of the knee progressors 1.4 % (n = 9). Those with progression of hip ROA were older and were more often women compared to those without progression. Furthermore, persons with progression of hip ROA more often had a Kellgren & Lawrence index grade ≥ 2 at baseline (53.7% versus 10.2%) compared to those without progression. The persons with progression of the knee had a higher BMI and slightly more often a Kellgren & Lawrence index grade ≥ 2 at baseline compared to those without progression (53.7% versus 48.8%).

Of the NSAIDs supposed to have a deleterious effect on joint cartilage by inhibition of GAG synthesis, naproxen was more often prescribed for short-term use (1–30 days) and ibuprofen more often for long-term use (> 180 days) (Table 2). The most frequently prescribed NSAID of the group supposed to have a neutral effect on cartilage metabolism was diclofenac. Similar to naproxen, diclofenac was also prescribed more often for short-term use and less

Table 1: Baseline characteristics of the study population, stratified by the presence or absence of hip or knee progression.

	Study population	Hip (K & L ≥ grade 1)		Knee (K & L ≥ grade 1)	
		Non progressors	Progressors	Non progressors	Progressors
Number	3,585	1,494	201	580	55
Age, years ± SD	66.0 ± 6.9	65.6 ± 6.7**	69.0 ± 6.8	67.9 ± 6.8	68.2 ± 6.9
Gender, % women	58.2	50.7**	67.7	64.8	72.7
BMI, kg/m^2 ± SD	26.3 ± 3.6	26.2 ± 3.3*	26.8 ± 4.5	27.1 ± 3.7*	28.5 ± 4.2
Kellgren & Lawrence ≥ grade 2, %		10.2	53.7	48.8	53.7
Presence of pain at baseline, %		10.4	28.0	22.5	42.6

K & L = Kellgren & Lawrence index.
BMI = body mass index.
Progression of the hip was defined as the presence of a joint space narrowing ≥ 1.0 mm or an incident total hip replacement at follow-up.
Progression of the knee was defined as the presence of a joint space narrowing ≥ 1.5 mm or an incident total knee replacement at follow-up.
Significant differences between persons without progression of osteoarthritis (OA) and persons with progression of OA.
* P < 0.05, ** P < 0.001

Table 2: Numbers of patients using the different types of (only data of frequently used) NSAIDs of the total study population (n = 3,585), stratified by duration of use.

	1–30 days	31–90 days	91–180 days	> 180 days
Group 1	786 (56.1)	354 (25.3)	118 (8.4)	142 (10.2)
Ibuprofen	320 (37.7)	301 (35.5)	92 (10.8)	136 (16.0)
Naproxen	507 (65.6)	173 (22.3)	46 (6.0)	47 (6.1)
Group 2	833 (54.6)	386 (25.3)	129 (8.4)	179 (11.7)
Diclofenac	833 (61.0)	321 (23.5)	94 (6.9)	118 (8.6)
Piroxicam	117 (49.2)	75 (31.5)	14 (5.9)	32 (13.4)

Group 1: those NSAIDs (ibuprofen and naproxen) that are supposed to have a deleterious effect on joint cartilage.
Group 2: those NSAIDs (such as diclofenac and piroxicam) that are supposed to have a neutral effect on cartilage metabolism.
The relative (%) use of the different types of NSAIDs is presented between parentheses.

often for long-term use. The duration of NSAIDs use of the two groups was dichotomized, and short-term use (1–30 days) was used as reference group in the following analyses.

As reported in Table 3, for the hip we unexpectedly found a significantly increased risk for progression in those subjects who used NSAIDs that were supposed to have a neutral effect on cartilage metabolism (group 2) for longer than 31 days with an adjusted OR of 1.7 (95% CI, 1.0–2.9), compared to the short-term user (1–30 days). The risk estimate for those who used NSAIDs supposed to have a deleterious effect on joint cartilage (group 1) for longer than 31 days was only 1.3 (OR) and failed to reach significance. The highest percentage of persons with progression of the hip was found in the group that used NSAIDs with a neutral effect for longer than 180 days (of group 2) (Figure 1). In both groups we found a

Table 3: Associations between use of NSAIDs and progression of hip and knee osteoarthritis.

| | Hip (n = 1,695) | | Knee (n = 635) | |
	crude	adjusted*	crude	adjusted*
Group 1	1.6 (1.0–2.4)	1.3 (0.8–2.1)	1.5 (0.7–3.0)	0.9 (0.4–2.0)
Group 2	2.4 (1.6–3.7)	1.7 (1.0–2.9)	1.0 (0.5–1.9)	0.9 (0.4–2.0)

Associations are presented by odds ratios (OR) with 95% confidence interval between parentheses.
Reference group is the short time user, between 1–30 days.
* Adjusted for age, gender, BMI, baseline radiological osteoarthritis, follow-up time and Defined Daily Dosage.
Group 1: those NSAIDs (indomethacin, naproxen and ibuprofen) that are supposed to have a deleterious effect on joint cartilage.
Group 2: those NSAIDs (such as diclofenac and piroxicam) that are supposed to have a neutral effect on cartilage metabolism.

Figure 1: Percentage persons with progression of hip osteoarthritis by duration of use of group 1 and group 2 NSAIDs.
Group 1: those NSAIDs (indomethacin, naproxen and ibuprofen) that are supposed to have a deleterious effect on joint cartilage.
Group 2: those NSAIDs (such as diclofenac and piroxicam) that are supposed to have a neutral effect on cartilage metabolism.
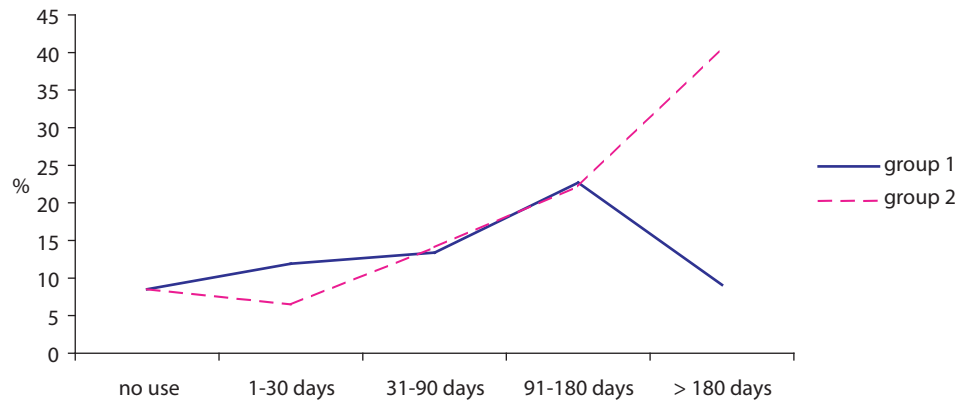Numbers of subjects per category:
Group 1:   555 / 193 / 67 / 22 / 22
Group 2:   555 / 232 / 85 / 27 / 32

clear trend that the longer the duration of use, the higher the percentage of persons with progression of hip ROA (Figure 1). For the knee we found no significant associations when we compared the long-term users with the short-term users of NSAIDs and progression of ROA.

To investigate whether individual types of NSAIDs could explain the significant increased risk for progression of the hip, we also assessed the associations between the use of the individual types of NSAIDs and progression of ROA of the hip (Table 4). For these analyses we used the four most frequently prescribed types of NSAIDs within the study population, namely ibuprofen, naproxen, diclofenac and piroxicam. The duration of the different types of NSAIDs use was categorized in four groups, i.e. 1–30 days, 31–90 days, 91–180 days and longer than 180 days. The short-term use (1–30 days) was used as reference group in the following analyses (Table 4). We found a significant crude association between long-term

**Table 4:** Associations between use of different types of NSAIDs and progression of hip osteoarthritis.

| (Numbers) | Hip | |
	crude	*adjusted
Ibuprofen		
– short (172)	**0.9** (0.5–1.6)	**0.9** (0.4–1.7)
– medium (55)	**1.1** (0.5–2.5)	**0.6** (0.2–1.7)
– long (62)	**1.8** (0.9–3.6)	**1.0** (0.4–2.3)
Naproxen		
– short (106)	**1.0** (0.5–1.9)	**1.0** (0.5–2.0)
– medium (33)	**2.2** (1.0–5.1)	**1.4** (0.5–3.9)
– long (26)	**0.2** (0.0–1.9)	**0.1** (0.0–1.2)
Diclofenac		
– short (167)	**1.3** (0.8–2.3)	**1.1** (0.6–2.0)
– medium (57)	**1.4** (0.6–3.2)	**1.4** (0.6–3.6)
– long (63)	**5.8** (3.2–10.5)	**3.6** (1.7–7.5)
Piroxicam		
– short (45)	**1.3** (0.5–3.5)	**1.5** (0.5–4.5)
– medium (6)	**1.0** (0.1–9.6)	**0.9** (0.1–10.1)
– long (20)	**0.9** (0.2–3.6)	**0.8** (0.2–3.9)

Duration of use; reference group is = 1–30 days, short = 31–90 days,
medium = 91–180 days, long = > 180 days.
Associations are presented by odds ratios (OR) with 95% confidence interval between parentheses.
* Adjusted for age, gender, BMI, baseline radiological osteoarthritis, follow-up time, and Defined Daily Dosage.

use (> 180 days) of diclofenac and progression of hip OA, with an OR of 5.8 (95% CI, 3.2–10.5). After adjustment for age, gender, BMI, baseline ROA of the hip, follow-up time, and defined daily dosage the risk estimate decreased importantly to an OR of 3.6 (95% CI; 1.7–7.5), but was still significantly increased. The association could also be confounded by
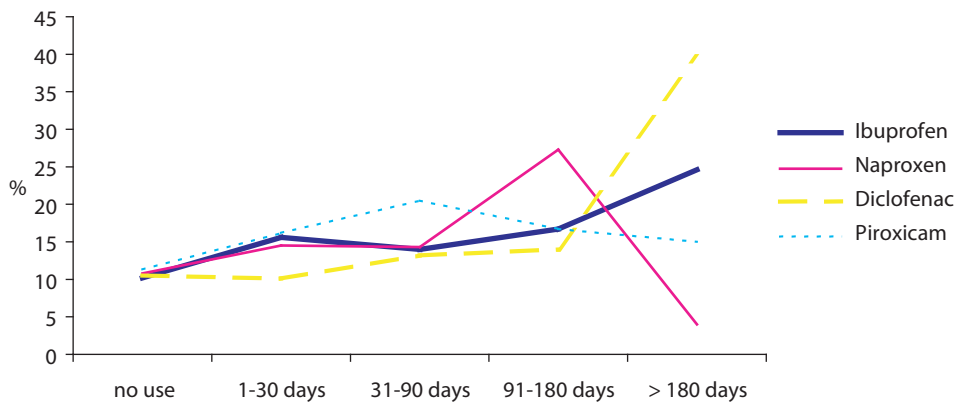
Figure 2: Percentage of persons with progression of hip osteoarthritis by duration of use of different types of NSAIDs.
Numbers of subjects per category:
Ibuprofen: 1461 / 170 / 172 / 55 / 62
Naproxen: 1479 / 276 / 106 / 33 / 26
Diclofenac: 1176 / 457 / 167 / 57 / 63
Piroxicam: 1775 / 74 / 45 / 6 / 20

110

the pain severity. Since we had information only on the absence or presence of hip pain, we repeated the analysis for those subjects with a prevalent hip pain at baseline. Of the subjects with hip pain at baseline (n = 243), 98.8% had pain for longer than 1 month, of which 30.2% had pain for 1 to 5 years and 51.1% had pain for longer than 5 years. In those subjects with hip pain the risk estimate for the long-term use of diclofenac was similar as reported (OR = 3.5), but failed to reach significance (95% CI; .9–13.9) due to the small numbers. Finally, the association could also be confounded by the "activity" of the OA process. Therefore, we additionally adjusted for the baseline erythrocyte sedimentation rate (as inflammation marker) and for the baseline CTX-II (fragments derived from type II collagen as marker for cartilage degradation). Again, the risk estimates (OR = 3.3) did not change importantly after adjustment.

In addition, for ibuprofen and diclofenac a trend was found that the longer the duration of use, the higher the percentage of persons with progression of hip ROA (Figure 2). This was not the case for naproxen and piroxicam.

No significant associations were found between the use of the individual types of NSAIDs and progression of ROA of the knee (data not reported in Table 4).

Discussion

In a large population-based prospective cohort study the division of NSAIDs into groups based on their supposed negative influence on cartilage metabolism, appears to be questionable. If any, the negative effect of NSAIDs on progression of hip ROA was found for the long-term use of diclofenac, surprisingly a NSAID considered to be a neutral type for cartilage metabolism. No clear associations were found between NSAIDs and progression of knee ROA.

The strengths of the present study are its population-based prospective design, its size, and the long follow-up period of 6.6 years, which enabled us to study all these different types of NSAIDs together. Because the present study comprised a healthy selection of the total population of the Rotterdam study, under-representation of symptoms may have resulted in an underestimation of the reported associations.

An important source of bias in the interpretation of the results concerning the association between NSAIDs and progression of OA in observational studies, is confounding by indication. The question is if the reported association between long-term use of NSAIDs and progression of ROA of the hip has been confounded by the severity of hip ROA or the presence of symptoms and/or side effects. After adjustment for radiologic severity at baseline and also for potential risk factors of severity (e.g. age, gender and BMI), the risk estimate decreased but was still significantly increased. Even after adjustment for inflammation and cartilage degradation markers, as sign of "activity" of the OA process, the risk estimate did

111

not change importantly. Also, in the subgroup of subjects with initial hip pain, the risk estimate was similar but failed to reach significance, probably because of the small size of this subgroup. At least the associations we found did not disappear after adjustment for these variables. It appears that the long-term use of NSAIDs is not harmless, with regard to their influence on progression of OA. In the present study, estimation of the use of NSAIDs is based on the prescriptions of NSAID and not on the actual NSAID intake in subjects. Therefore, it may well be that the actual duration of NSAID use has been overestimated and probably resulted in an underestimation of the reported associations.

Diclofenac has a differential effect on progression of hip and knee OA, and this raises the question whether diclofenac also has a differential effect on the cartilage metabolism of the hip and knee. Until now, only two studies have investigated the influence of diclofenac on the cartilage metabolism of the knee, and both reported that diclofenac did not induce any degenerative processes (animal and in vitro study) (9, 26). However, the study of Vignon et al. (27) reported a slightly decreased proteoglycanase activity in human osteoarthritic cartilage of the hip. One problem is that these studies did not investigate the influence of diclofenac on cartilage metabolism of the hip or the knee joint simultaneously and therefore the question whether diclofenac has a differential effect remains unanswered.

It has been suggested that effective pain relief due to analgesic drugs causes a patients to become more active. Because of this (suggested) increased activity, the mechanical loading of a less painful joint will increase. This increased mechanical loading may modify the supposed deleterious effect of some NSAIDs on cartilage. This would only hold for the relation between diclofenac and the hip joint. Consequently, the efficacy of diclofenac in pain relief should be considerably better and also result in fewer side-effects compared to ibuprofen, naproxen and piroxicam. This was, however, not confirmed by Towheed et al. who summarized the literature on the efficacy and side effects of NSAIDs in hip OA (1). Moreover, a supposed better efficacy and fewer side effects of diclofenac suggest that diclofenac could be used longer than ibuprofen, naproxen and piroxicam. However, our findings (see Table 2) do not support the longer use of diclofenac.

Our data suggest that diclofenac may not be harmless and may induce accelerated progression of hip OA. Whether this occurs due to a real deleterious effect on cartilage or due to excessive mechanical loading on an analgesic hip remains to be investigated. In view of the effect of diclofenac, it would be interesting to know the effect of cyclooxygenase (COX)-II selective inhibitors on cartilage metabolism. We conclude that there is a clear need to further investigate the influence of individual types of NSAIDs on cartilage metabolism in a clinical situation.

Acknowledgments

References

1. Towheed T, Shea B, Wells G, Hochberg M. Analgesia and non-aspirin, non-steroidal anti-inflammatory drugs for osteoarthritis of the hip. Cochrane Library 2003; Issue 1.
2. Bierma-Zeinstra SMA, et al. How do general practitioners manage hip problems in adults? Scan J Prim Health Care 2000;18:159–64.
3. Hungin APS, Kean DWF. Nonsteroidal anti-inflammatory drugs: Overused or underused in osteoarthritis? Am J Med 2001;110(1A):8S-11S.
4. Ding C. Do NSAIDs affect the progression of osteoarthritis? Inflammation 2002;26(3):139–42.
5. Rainsford KD, Ying C, Smith FC. Effects of meloxicam, compared with other NSAIDs, on cartilage proteoglycan metabolism, synovial prostaglandin E2, and production of interleukins 1, 6, and 8, in human and porcine explants in organ culture. J Pharm Pharmacol 1997;49(10):991–8.
6. Rainsford KD, Skerry TM, Chindemi P, Delaney K. Effects of the NSAIDs meloxicam and indomethacin on cartilage proteoglycan synthesis and joint responses to calcium pyrophosphates crystals in dogs. Vet Res Commun 1999;23(2):101–13.
7. Muir H, Carney SL, Hall LG. Effects of tiaprofenic acid and other NSAIDs on proteoglycan metabolism in articular cartilage explants. Drugs 1988;35 Suppl.:15–23.
8. Gencosmanoglu BE, Eryavuz M, Dervisoglu S. Effects of some nonsteroidal anti-inflammatory drugs on cartilage of rats in an experimental model of osteoarthritis. Res Exp Med 2001;200(3):215–26.
9. Kalbhen DA. The influence of NSAIDs on morphology of articular cartilage. Scan J Rheumatol Suppl. 1988;77:13–22.
10. Palmoski MJ, Brandt KD. Effects of some nonsteroidal antiinflammatory drugs on proteoglycan metabolism and organization in canine articular cartilage. Arthritis Rheum 1980;23(9):1010–20.
11. Ratcliffe A, Azzo W, Saed-Nejad F, Lane NE, Rosenwasser MP, Mow C van. In vivo effects of naproxen on composition, proteoglycan metabolism, and matrix metalloproteinase activities in canine articular cartilage. J Orthop Res 1993;11(2):163–71.
12. Pelletier JP. The influence of tissue cross-talking on OA progression: role of nonsteroidal anti-inflammatory drugs. Osteoarthritis Cartilage 1999;7:374–6.
13. Huskisson EC, Berry H, Gishen P, Jubb RW, Whitehead J. Effects of anti-inflammatory drugs on the progression of osteoarthritis of the knee. J Rheumatol 1995;22:1941–6.
14. Rashad S, Hemingway A, Rainsford KD, Revell P, Low F, Walker F. Effect of non-steroidal anti-inflammatory drugs on the course of osteoarthritis. Lancet 1989;2:519–21.
15. Vignon E, Mathieu P, Charlet C, Balblanc JC, Richard M. Effects of naproxen (naprosyne) on the metabolism of arthrotic cartilage in man in vivo. Rev Rheum Mal Osteoartic 1991;58(3):11S-15S.
16. Hofman A, Grobbee DE, Jong PT de, VandenOuweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. Eur J Epidemiol 1991;7(4):403–22.
17. Reijman M, Hazes JMW, Pols HAP, Bernsen RMD, Koes BW, Bierma-Zeinstra SMA. Validity and reliability of three definitions of hip osteoarthritis: cross-sectional and longitudinal approach. Ann Rheum Dis 2004; in press.

18. Auleley G, Girardeau B, Dougados M, Ravaud P. Radiographic assessment of hip osteoarthritis progression: impact of reading procedures for longitudinal studies. Ann Rheum Dis 2000;59:422–7.

19. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494–502.

20. Kellgren JH, Jeffrey MR, Ball J. The epidemiology of chronic rheumatism. Atlas of standard radiographs of arthritis. Oxford: Blackwell Scientific Publications 1963.

21. Reijman M, et al. A new marker for osteoarthritis: cross-sectional and longitudinal approach. Arthritis Rheum 2004;50(8):2471–8.

22. Croft P, Cooper C, Wickham C, Coggon D. Defining osteoarthritis of the hip for epidemiologic studies. Am J Epidemiol 1990;132(3):514–22.

23. Bergink AP, et al. Estrogen receptors alpha gene haplotype is associated with radiographic osteoarthritis of the knee in elderly men and women. Arthritis Rheum 2003;48(7):1913–22.

24. Odding E, Valkenburg HA, Algra D, Vandenouweland FA, Grobbee DE, Hofman A. Associations of radiological osteoarthritis of the hip and knee with locomotor disability in the Rotterdam Study. Ann Rheum Dis 1998;57(4):203–8.

25. Rothman KJ, Greenland S. Introduction to stratified analysis. In: Rothman KJ, Greenland S, editors. Modern epidemiology. Second edition. Boston/Toronto: Little, Brown and Company 1998:311–26.

26. Blot L, Marcelis A, Devogelaer JP, Manicourt DH. Effects of diclofenac, aceclofenac and meloxicam on the metabolism of proteoglycans and hyaluronan in osteoarthritic human cartilage. Br J Pharmacol 2000;131(7):1413–21.

27. Vignon E, Mathieu P, Louisot P, Richard M. In vitro effect of nonsteroidal anti-inflammatory drugs on proteoglycanase and collagenase activity in human osteoarthritic cartilage. Arthritis Rheum 1991;34(10):1332–5.

8

General discussion

In the previous chapters, the main findings and limitations of each study were discussed. In this chapter the main topics addressed in this thesis will be discussed in a broader perspective. Finally, suggestions are made for future research.

Case definition of prevalent and incident hip OA

The first question in this thesis was "How can we define the presence of hip OA in epidemiological studies?" As already mentioned in the Introduction, in most epidemiological studies OA was assessed by means of radiologic evaluation. The literature reports a wide range of prevalences of hip OA, mainly because of the variation in the definitions used. This would not be a problem if the agreement between the different definitions were high; however, we found that the overlap between three commonly used definitions of hip OA was strikingly low. Only half of the cases defined by the Kellgren & Lawrence index ≥ grade 2 were also defined by a minimal joint space of ≤ 2.5 mm, and for the index of Croft this was even lower. Because of this problem, the reported estimates of risk factors for hip OA may have been underestimated up to now. We argue that international consensus on a universal definition (or definitions) is needed for epidemiological and clinical studies. Therefore the validity (especially the expert or predictive validity) of the commonly used definitions should be studied more thoroughly.

Chapter 3 presents the results of our study investigating the reliability and validity of the Kellgren & Lawrence grade, Minimal Joint Space (MJS), and the Croft grade. Kellgren & Lawrence proved to be the best predictor for a total hip replacement at follow-up. We also found that the baseline Kellgren & Lawrence grade had an impressively stronger independent association with progression of hip OA compared with MJS at baseline. Furthermore, we found that MJS is dependent on height. Recently, Lanyon et al. (1) reported that women have a significantly smaller joint space width than men and that this difference remained significant after adjusting for height. Goker and colleagues (2, 3) also found gender differences in JSW, but these differences were no longer significant after adjustment for height. Therefore, the somewhat higher prevalence of hip OA in women defined by MJS according to Croft, as reported in Chapter 3, might be explained by a smaller JSW in women compared with men.

Based on the above-mentioned findings and arguments, we conclude that for epidemiological studies radiological hip OA should preferably be defined by the Kellgren & Lawrence grading system.

Incident versus progression

To investigate the prognostic factors for progression of hip OA, we first had to deal with another issue; namely, how to define progression of hip OA?

OA is a slow progressive disease, in which the normal joint of the patient degenerates over time to end stage OA. During this ongoing process there is a moment at which the patient crosses the borderline between what is defined as absence of OA and presence of OA. The distinction between incident OA and progression of already existing OA depends on at which point along this continuum the cut-off point of present OA is defined (4). If OA is diagnosed earlier in the future, because of more sensitive diagnostic tools, cases formerly considered to be 'incident cases' will then be considered as 'progressive cases'.

In case of the Kellgren & Lawrence index, incident hip OA is usually defined as a grade 0 or 1 at baseline, and greater than or equal to grade 2 at follow-up. However, one can question whether the cut-off point of ≥ grade 2 is valid and whether it is correct to classify people with grade 1 as a 'normal' group? Recently Hart and Spector investigated whether the Kellgren & Lawrence grade 1 of the knee was a reliable indicator of knee OA in a longitudinal population-based study (5). After 10 years of follow-up, more than 60% of the subjects with grade 1 at baseline had developed grade 2 or higher, whereas, 20% of those with a Kellgren & Lawrence grade 0 at baseline had developed a grade 2 or higher. In our hip study, however, after 6 years of follow-up we found that 7.6% of those subjects with a Kellgren & Lawrence hip grade 1 versus 1.4% with grade 0, developed a grade 2 or higher. These results suggest that the cut-off point of ≥ grade 2 for the hip seems to be valid, whereas for the knee a cut-off point of ≥ grade 1 seems more appropriate.

In contrast, several international scientific committees have suggested that assessment of joint space narrowing as the most sensitive technique to assess radiographic changes over time. The Kellgren & Lawrence index appears to be a valid tool for case definition; however, the index may not be sensitive enough to evaluate osteoarthritic changes over time. In the future, the distinction between incident OA and progression of already existing OA may well be influenced by tools that can identify OA in an earlier stage compared to the traditional radiograph, such as the MRI or a biomarker. It is important to realize that the chosen diagnostic tool and cut-off point determine whether a person is defined as a prevalent OA case. For the time being, in case of an incident hip OA the Kellgren & Lawrence index, with a cut-off point of ≥ grade 2, appears to be the most useful tool. However, in case of progression of hip OA, joint space narrowing appears to be the most useful tool.

Overall we conclude that, from a methodological view, the distinction between incident OA and progression of OA is arbitrary and until now this problem remains unanswered.

Determinants of progression

Why is the distinction between incident hip OA and progression of hip OA so important? This would not be the case if the risk factors for developing an incident hip OA or progression of an already existing hip OA were similar. However, not all the persons with a present hip OA will progress (see Figure 1), indicating that the risk factors for incident hip OA and progression of hip OA are different. Within the participants of the Rotterdam Study with a radiographic OA of the hip (Kellgren & Lawrence ≥ grade 2), 45.3% showed progression during the follow-up period. However, if these subjects also had hip pain at baseline, we found that 71.4% showed progression during follow-up.

In Chapter 5 we reported that age, gender and signs of the presence or severity of OA were prognostic factors for progression of hip OA, whereas systemic factors like family history of OA, diabetes (type II), hypertension and age at menopause were not. There is growing evidence that the development of hip OA is under strong genetic influence (6, 7), whereas, the genetic influence on progression of hip OA is less clear. Because this suggests that the role of systemic factors in progression is less important, we constructed a prognostic model of progression of hip OA, i.e. a modified version of the model of Felson (8). The

Radiographic hip OA at baseline (n = 251)

Hip pain at baseline (24.1%)

Progression of hip OA (45.3%)

Participants with hip pain and radiographic hip OA at baseline who developed progression during follow-up (71.4%).
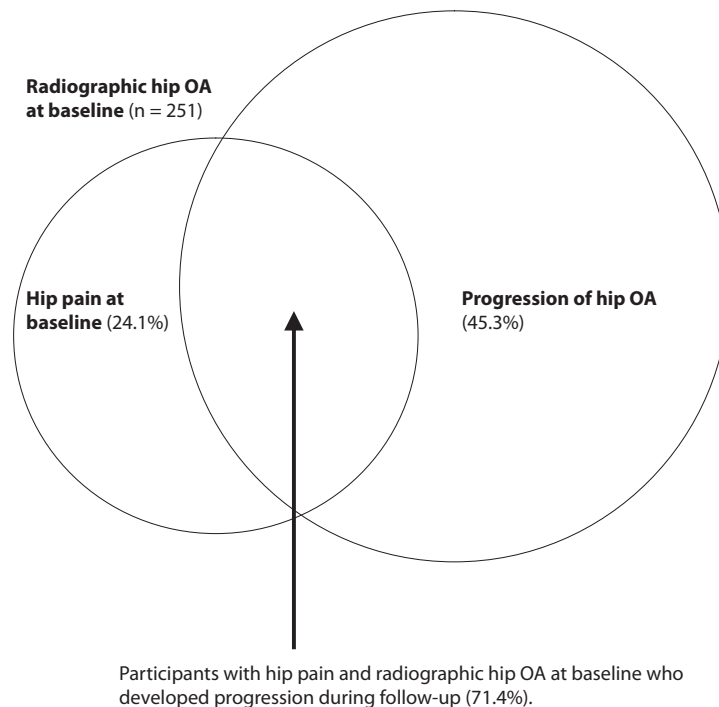
Figure 1: Relationship between radiographic hip osteoarthritis (OA), the presence of hip pain at baseline and progression of hip OA in participants of the Rotterdam Study.
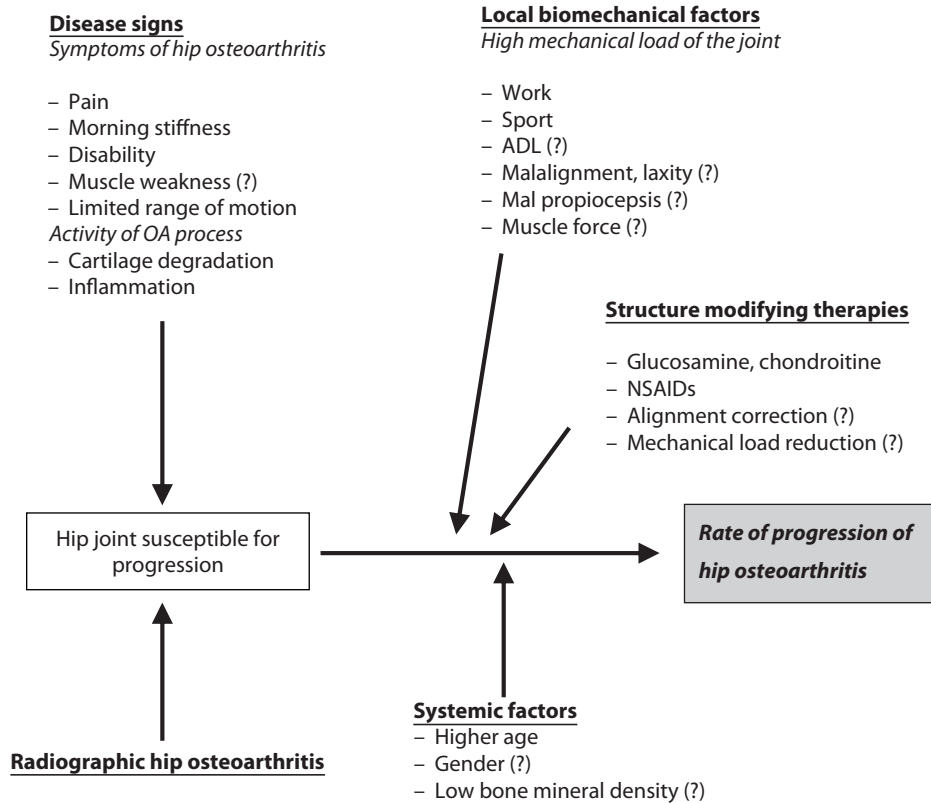
Disease signs
Symptoms of hip osteoarthritis

– Pain
– Morning stiffness
– Disability
– Muscle weakness (?)
– Limited range of motion
Activity of OA process
– Cartilage degradation
– Inflammation

Local biomechanical factors
High mechanical load of the joint

– Work
– Sport
– ADL (?)
– Malalignment, laxity (?)
– Mal propiocepsis (?)
– Muscle force (?)

Structure modifying therapies

– Glucosamine, chondroitine
– NSAIDs
– Alignment correction (?)
– Mechanical load reduction (?)

Hip joint susceptible for progression

Rate of progression of hip osteoarthritis

Systemic factors
– Higher age
– Gender (?)
– Low bone mineral density (?)

Radiographic hip osteoarthritis

Figure 2: Prognostic factors of hip OA and their interaction on the progression process; modified from (6).

model indicates that a person with radiographic OA of the hip will progress depending on the interaction with potential prognostic factors (see Figure 2). In Chapter 5 we reported that the proportion of progressors was higher in case of the presence of symptoms of hip OA and signs of an active OA process. Furthermore, we argue that besides the minor role of systemic factors (such as age and gender), local biomechanical factors play an important role in progression of hip OA. It is conceivable that local biomechanical factors such as workload, sports activity and mechanical load during daily activity interact with already existing osteoarthritic signs of the hip joint. Until now, however, all the studies that investigated the influence of mechanical load on hip OA have looked at the historical load of the hip joint and not at the load after the occurrence of hip OA. In addition to their direct effect, local factors may mediate in the pathway between other factors, such as structure modifying therapies and systemic factors, and OA (9). Moreover, we reported in Chapter 6 that the association between acetabular dysplasia and incident hip radiographic OA might be enhanced by female gender and mechanical workload. Whether the same holds for progression, however, is not yet clear. In case of the knee there is some evidence that malalignment is associated with progression of knee OA (9). Besides these determinants

of progression, the rate of progression might also be influenced by potential structure-modifying therapies, such as glucosamine, chondroitine, NSAIDs, alignment correction, or mechanical load reduction. The efficacy of these modifying therapies is currently under discussion and several studies are currently exploring this issue.

Overall, we conclude that the distinction between incident OA and progression of OA is important from a clinical view, because different determinants play a role in the onset of OA than in the progression of an already existing OA. In particular, the role of mechanical load and malalignment on progression should be thoroughly investigated.

Biomarkers

As described in Chapter 4, biomarkers aim to detect changes in OA with more reliability and sensitivity, preferably in an early stage. The first potential use of a biomarker is that a shorter follow-up time may be sufficient to observe changes in the joint. In our study we found that CTX-II had a strong association with progression of hip and knee OA (with an OR of 8.3 and 6.2 respectively). These results suggest that CTX-II may be able to identify subjects at high risk for progression of OA. However, whether CTX-II is capable to observe OA changes during a follow-up period, remains unanswered. As mentioned before, we found that CTX-II is associated with progression of OA; however, whether CTX-II has an additional diagnostic value is unclear. For such a clinical application it will be more informative to assess the prior probability, the posterior probability and the likelihood ratio.

Another potential use of biochemical markers is that the onset of osteoarthritic signs can be detected at an earlier stage. This could be useful to identify target groups in need of potential disease-modifying OA drugs (DMOADs), assuming that these drugs become available in the future. The hypothesis is that these DMOADs modify the synthesis of cartilage. Until now, however, the validity of CTX-II as an early detector of OA signs has not been investigated. Because of the absence of a gold standard that can detect the onset of OA in an earlier stage, another option is to assess the predictive validity of CTX-II. In this case the predictive validity is expressed as the ability of CTX-II to predict an incident OA. Because this has not yet been investigated, the ability of CTX-II to detect the onset of OA in earlier stage remains unanswered.

In addition, CTX-II seems to be a specific marker for cartilage degradation, since CTX-II is associated with joint space narrowing but not with incident osteophytes. However, the main limitation of CTX-II always will be that it is not completely specific for OA, since it is also associated with rheumatoid arthritis (10). Moreover, the CTX-II level express the total turnover rate of cartilage and can be modified by cartilage degradation of another joint next to the joint of interest. No studies have investigated the association between, for example, CTX-II and OA of the spine. It is conceivable that OA of the spine will influence the CTX-

II level as measured in urine. We conclude that besides these remarks concerning the specificity of CTX-II, the ability of CTX-II to detect changes in a shorter period of time and the ability of CTX-II to detect the onset of OA in earlier stage need to be investigated.

NSAIDs and progression

In Chapter 7 we discussed that different types of NSAIDs might influence the rate of progression. If any, the negative effect of NSAID was found for the long-term use of diclofenac and progression of hip OA.

Several sources of bias are important when interpreting the results of the association between NSAIDs and progression of radiographic OA, as reported in Chapter 7. Bias is generally divided into selection bias, information bias and confounding. Selection bias may occur when selection of subjects for the drug exposure group and the reference group of a cohort study differs between diseased and nondiseased persons. Selection bias may also occur if cases and controls are drawn from different source populations. In our study, however, it is unlikely that selection bias occurred because the study was a prospective one and population based. However, because our study included a healthy selection of the total population of the Rotterdam study, under-representation of symptoms may have resulted in an underestimation of the reported associations.

Information bias may occur if classification of disease status depends on exposure status, or vice versa. In the study reported in Chapter 7, the pharmacy records were used to avoid potential misclassification. One difficulty is that we used the amounts on the prescriptions of NSAID, and not the actual use of NSAIDs. Hence, it seems logical that the actual duration of use of NSAIDs has been overestimated and probably resulted in an underestimation of the reported associations. Overall, we assume that it is unlikely that information bias has occurred in this study since the pharmacy data were collected before the follow-up measurements.

Confounding by independent risk factors of progression of hip OA, which may also be associated with NSAID use, can usually be dealt with in the analyses of observational studies. A more difficult type of confounding is confounding by indication: this term is used when a variable is a risk factor for a disease among nonexposed and exposed persons and is associated with the exposure of interest in the population from which the cases derive, without being an intermediate step in the causal pathway between the exposure and the disease (11). This problem is thoroughly discussed in Chapter 7.

We conclude that at least the associations we reported in Chapter 7 did not disappear after adjustment for all the mentioned variables. Hence, there is a clear need to further investigate the influence of individual types of NSAID on cartilage metabolism in a clinical situation.

Advantages and disadvantages of a large cohort study

The main advantages of a large cohort study with a prospective design and a long follow-up period, such as the Rotterdam Study, are the large database and the possibility to investigate interactions between different risk factors.

Concerning the disadvantages, because of the long follow-up period, the outcome measures dependent on the baseline measurements may not have been updated according to the latest recommendations of international scientific committees. Nowadays, the standardization of radiographic procedures has become more precise to current opinion. Hence, for the radiographic procedure of the hip, especially the focus-to-film distance and the location of where the X-ray beam should be centered, are more precisely specified. Another disadvantage of using data from an already existing cohort is that the objectives of the present study are necessarily defined after the collection of the data. Consequently, for our study objectives we missed the optimal data collection of potential risk factors such as the exact location of the pain, the severity of the pain, sports activity, specificity and frequency of the components of a job, physical examination, and muscle strength.

In general, however, we were able to adequately answer our research question using the data of the Rotterdam Study.

Suggestions for future research

The results of the studies presented in this thesis give rise to new hypotheses for future research. Several research questions have already been presented in the Discussion section of the various chapters.

Our hypothesis is, as mentioned before, that the risk factors for incident hip OA are different from the prognostic factors for progression of hip OA. The first interesting topic is to identify persons at high risk for developing incident hip OA or progression of hip OA. More data are needed to define the risk or prognostic factors allowing to identify such a high-risk person. Secondly, tools that can detect the onset of OA at an earlier stage, such as biomarkers, need to be investigated more thoroughly. Furthermore, in case of progression of hip OA we assume that local factors play a more important role, than the systemic factors. Especially the influence of mechanical load on an already existing osteoarthritic joint and the influence of a malalignment on progression should be investigated. Finally, the relationship between body mass index and OA is not yet fully elucidated. We suggest that the influence of body mass index on the development and progression of OA of the hip, knee and hand should be investigated in the same large open population.

Consequently, the scientific evidence for therapeutic interventions of patients with OA should be expanded. Especially the efficacy of structural modifying interventions (such as

disease-modifying OA drugs, alignment correction or reduction of the mechanical load of the hip joint) needs to be investigated in clinical trials. Finally, the influence of different types of NSAIDs and the effect of the new COX-II selective inhibiters on progression of hip OA are topics for future research.

Implications for daily practice

Besides suggestions for future research, the reported results also have implications for the daily practice of a general practitioner (GP).

The main implication is the value of an X ray in case of the presence of hip pain. If a patient of 55 years or older visits a GP because of the presence of hip pain, an additional X ray can be rewarding for two reasons. Firstly, in combination with radiographic acetabular dysplasia there is an increased risk of developing an incident hip OA, and secondly, in combination with radiographic evidence of hip OA there is an increased risk of progression.

Based on the results that nearly all patients (of 55 years or older) with a Kellgren & Lawrence index ≥ grade 2 and hip pain will develop progression of hip OA, it seems reasonable that structure modifying therapies will be more effective in the early stage of OA. Hence, the primary target group of such therapies will be those patients with hip pain combined with minimal radiographic signs of hip OA. However, as mentioned in the previous paragraph, the efficacy of structural modifying interventions needs to be investigated more thoroughly.

The ACR and Eular guidelines for the medical management of hip OA recommended that overweight patients with hip OA should lose weight. However, results of our study and also of other epidemiological studies could not confirm that overweight is a risk factor for progression of hip OA. So far there has been no study that investigated the influence of mechanical load on an already existing osteoarthritic joint. Hence, the question of what should be recommended to a patient with hip OA, concerning mechanical loading of the joint, remains unanswered.

Besides the marginal surplus of NSAIDs in pain relief compared to acetaminophen and the known major and minor side effects of NSAIDs, our data suggest that the long-term use of diclofenac may not be harmless and may induce accelerated progression of hip OA. Based on these results we suggest a GP to be critical in the prescription of long-term use of NSAIDs in case of hip OA.

References

1.   Lanyon P, Muir K, Doherty S, Doherty M. Age and sex differences in hip joint spaces among asymtomatic subjects without structural change. Arthritis Rheumatism 2003;48(4):1041–6.
2.   Goker B, Sancak A, Arac M, Shott S, Block JA. The radiographic joint space width in clinically normal hips: effects of age, gender and physical parameters. Osteoarthritis Cartilage 2003;11:328–34.
3.   Goker B, Block JA. Correlations of hip joint space among asymptomatic subjects: comment on the article by Lanyon et al. Arthritis Rheum 2004;50(2):678.
4.   Lohmander LS, Felson DT. Can we identify a 'high risk' patient profile to determine who will experience rapid progression of osteoarthritis? Osteoarthritis Cartilage 2004;12(suppl A):S49–52.
5.   Hart DJ, Spector TD. Kellgren & Lawrence grade 1 osteophytes in the knee – doubtful or definite? Osteoarthritis Cartilage 2003;11:149–50.
6.   Lanyon P, Muir H, Doherty S, Doherty M. Assessment of a genetic contribution to osteoarthritis of the hip: sibling study. BMJ 2000;321:1179–83.
7.   MacGregor AJ, Antoniades L, Matson M, Andrew T, Spector TD. The genetic contribution to radiographic hip osteoarthritis in women: Results of a classic twin study. Arthritis Rheum 2000;43(11):2410–6.
8.   Felson DT. An update on the pathogenesis and epidemiology of osteoarthritis. Radiol Clin N Am 2004;42:1–9.
9.   Sharma L. Local factors in osteoarthritis. Curr Opin Rheumatol 2001;13:441–6.
10.   Garnero P, Gineyts E, Christgau S, Finck B, Delmas PD. Association of baseline levels of urinary glucosyl-galactosyl-pyridinoline and type II collagen C-telopeptide with progression of joint destruction in patients with early rheumatoid arthritis. Arthritis Rheum 2002;46:21–30.
11.   Salas M, Hofman A, Stricker BH. Confounding by indication: an example of variation in the use of epidemiologic terminology. Am J Epidemiol 1999;149(11):981–3.

9

Summary

Summary

Osteoarthritis (OA) is the most frequent disorder of the locomotor system and the prevalence of OA will increase with the aging of the Western society. Especially when the hip or knee is involved, OA causes considerable difficulty in walking, stair climbing and other lower extremity tasks. OA of the hip can be especially disabling because of the pain and functional impairment. The identification of patients at high risk for progression of hip OA is important for at least two reasons. Firstly, well-characterized 'high risk' groups may be useful in clinical trials and, secondly, assuming that disease-modifying OA drugs may become available in the future, to identify primary target groups in need of such therapy. Additionally, in a clinical situation the identified non-progressors can be reassured. Until now the prognostic factors of progression of hip OA have been investigated in small studies, with a short follow-up time, and only in a hospital setting. The overall aim of this thesis was to determine the prognostic factors of osteoarthritis of the hip in a large open population with a long-term follow-up.

Nearly all studies presented in this thesis were based on data from the Rotterdam Study, a large prospective population-based cohort study in the Netherlands. Participants of this study were men and women aged 55 years and over living in Ommoord, a suburb of Rotterdam.

Chapter 2 presents a study in which we systematically summarized the literature addressing the validity, reliability and applicability of seven commonly used definitions of hip OA for epidemiological studies, in order to use them primarily as classification criteria. Considering how frequently the definitions of hip OA are used for epidemiological studies, it is surprising that the validity of these definitions has been so poorly investigated. Summarizing the literature showed that the reliability of the minimal joint space (MJS) according to Croft, the Kellgren & Lawrence grade and the index according to Lane was similar, but the MJS had the highest relationship with hip pain in a male population. We recommend that the validity, especially the expert or predictive validity, of the commonly used definitions be studied more thoroughly. Moreover, the different definitions should be investigated in the same clinical setting.

Therefore, **Chapter 3** presents an evaluation of the reliability and validity of three frequently used radiological definitions of hip OA: namely, the Kellgren & Lawrence grade, the Minimal Joint Space, and Croft's grade that were used in the Rotterdam Study. The inter-rater reliability was similar for the Kellgren & Lawrence grade and MJS, but slightly lower for Croft's grade. The Kellgren & Lawrence grade and MJS showed the strongest associations with clinical symptoms of hip OA. Gender appears to be a significant effect modifier for the Kellgren & Lawrence grade in that women had a significantly stronger association with symptoms than men. The gender differences in joint space width, however, were attributed to differences

in height between women and men. The Kellgren & Lawrence grade showed the highest predictive value for total hip replacement at follow-up compared to the other definitions.

Chapter 4 focuses on the association between a cartilage degradation marker, urinary concentrations of C-telopeptide fragments of collagen type II (CTX-II), and the prevalence and progression of radiological OA (ROA) of the hip and knee. Subjects with a CTX-II level in the highest quartile had a 4.2 times increased risk of having ROA of the knee and at the hip compared to subjects with a CTX-II level in the lowest quartile. We observed a stronger association for subjects with hip pain at baseline compared with those without hip pain. Subjects with a CTX-II level in the highest quartile had a 6.2 times increased risk for progression of ROA at the knee and an 8.3 times increased risk for progression of ROA at the hip compared to subjects with a CTX-II level in the lowest quartile. All of these associations were found to be independent of known risk factors for OA, such as age, gender and body mass index (BMI).

In **Chapter 5** we investigated which determinants will best identify those persons at high risk for progression of hip OA. Of the study population, 13.1% of the subjects had progression of ROA of the hip during follow-up. Starting with a simple model of only directly obtainable variables collected by history taking, the Kellgren & Lawrence score at baseline, when added to the model, was a strong predictor of progression, especially in those subjects with hip pain at baseline. In addition, a lower limb disability index of ≥ 0.5 and a CTX-II concentration ≥ 235.5 mmol/l were also independent identifiers of these high-risk persons. CTX-II is only a moderate predictor of progression of hip OA compared to the variables collected by history taking, physical examination and an X-ray. Overall, we conclude that in a clinical situation and for clinical trials, an X-ray offers valuable additional information to identify persons at high risk for progression of hip OA.

Chapter 6 explores the association between radiographic evidence of acetabular dysplasia in participants without ROA of the hip at baseline, and an incident hip ROA. In this study population 9.3% developed an incident ROA of the hip during follow-up. Subjects with acetabular dysplasia had a 4.3 times increased risk for incident hip OA compared to subjects without acetabular dysplasia. These associations were independent of known determinants of OA, such as age, gender, and BMI. Furthermore, these associations seemed to be enhanced by female gender, heavy mechanical workload and low BMI.

Chapter 7 reports on the associations between two groups of NSAIDs, i.e. those (indomethacin, naproxen and ibuprofen) that are supposed to have a deleterious effect on joint cartilage, and those NSAIDs (such as diclofenac and piroxicam) assumed to have a neutral effect on cartilage metabolism, and progression of OA of the hip and knee. Those subjects who used NSAIDs supposed to have a neutral effect on cartilage metabolism (> 31 days) surprisingly had a 1.7 increased risk of hip ROA compared to the short-term user (1–30 days). This increased risk could be explained by the long-term use (> 180 days) of diclofenac. No clear associations were found between NSAID and progression of knee ROA.

Whether the increased risk is due to a real deleterious effect on cartilage or through the effect of mechanical loading an analgesic hip remains to be investigated.

In **Chapter 8** the main topics addressed in this thesis are discussed in a broader perspective. The two main problems of how to define the presence of hip OA, and how to define progression of hip OA are addressed. The discussion also focuses on the use of a biomarker and sources of potential bias in the association between NSAIDs and progression. Finally, recommendations are made for future research.

Samenvatting

Artrose is een van de meest voorkomende aandoeningen van het bewegingsapparaat en de prevalentie van artrose zal gezien het vergrijzen van de westerse maatschappij alleen maar toenemen. Voornamelijk als de heup of de knie is aangedaan, levert artrose veel problemen op bij lopen, traplopen en andere functies van de onderste extremiteiten. Door de pijn en het verminderd functioneren van de onderste extremiteit kunnen personen met heup artrose gehandicapt zijn. Het identificeren van personen met een verhoogd risico op progressie van artrose is voor tenminste twee redenen belangrijk. Ten eerste, kunnen specifieke 'hoog risico' groepen nuttig zijn voor klinische trials en ten tweede, om specifieke doelgroepen voor artrose-modificerende medicijnen te identificeren, ervan uitgaande dat deze therapie in de toekomst beschikbaar zijn. Bovendien kunnen in een klinische situatie de personen die geïdentificeerd worden waarbij de artrose niet zal verergeren, gerustgesteld worden. De prognostische factoren voor progressie van heup artrose zijn tot nu toe alleen maar onderzocht in kleine studies, met een korte follow-up tijd en in een ziekenhuis setting. De overall doelstelling van dit proefschrift was het bepalen van de prognostische factoren van heup artrose in een grote open populatie met een lange follow-up tijd.

Bijna alle studies die in dit proefschrift gepresenteerd worden, zijn gebaseerd op het Rotterdamse ERGO-onderzoek (Erasmus Rotterdam Gezondheid en Ouderen), internationaal bekend als "the Rotterdam Study". Dit is een groot prospectief bevolkingsonderzoek onder mannen en vrouwen van 55 jaar en ouder uit de Rotterdamse deelgemeente Ommoord.

In **hoofdstuk 2** presenteren we een systematisch overzicht van de literatuur betreffende de validiteit, de reproduceerbaarheid en de toepasbaarheid van zeven vaak gebruikte definities (classificatie criteria) van heup artrose binnen epidemiologische studies. Ervan uitgaande hoe frequent deze definities van heup artrose worden gebruikt binnen epidemiologische studies, is het verrassend dat de validiteit van de gebruikte definities zo weinig is onderzocht. De literatuur samenvattend, blijkt dat de reproduceerbaarheid van de "minimal joint space (MJS)" volgens Croft, de Kellgren & Lawrence index en de index volgens Lane vergelijkbaar is. De MJS vertoonde de sterkste relatie met heuppijn in een mannelijke populatie, echter dit is gebaseerd op slechts 1 studie. Onze aanbeveling is dat de validiteit, met name de expert of predictieve validiteit, van de gebruikte definities beter moet worden onderzocht. Daarnaast adviseren wij dat zo'n onderzoek gebeurt binnen een zelfde populatie.

In **hoofdstuk 3** evalueren wij de reproduceerbaarheid en validiteit van drie frequent gebruikte radiologische definities van heup artrose, namelijk de Kellgren & Lawrence index, de MJS en de index volgens Croft, getoetst binnen één grote open populatie, namelijk de Rotterdam studie. De inter-beoordelaar-reproduceerbaarheid was vergelijkbaar voor de Kellgren & Lawrence index en de MJS, maar wat lager voor de index volgens Croft. De Kellgren & Lawrence index en de MJS lieten beide sterke associaties met klinische symptomen van heup artrose zien. Geslacht bleek een significante effect modificator te zijn

voor de Kellgren & Lawrence index; vrouwen hadden significant sterkere associaties met klinische symptomen dan mannen. Het verschil in breedte van de gewrichtspleet tussen mannen en vrouwen was toe te schrijven aan het verschil in lengte tussen mannen en vrouwen. De Kellgren & Lawrence index liet de hoogste predictieve waarde zien voor het voorspellen van een "gewricht vervangende operatie" gedurende follow-up vergeleken met de andere twee definities.

Hoofdstuk 4 richt zich op de associatie tussen een kraakbeen degradatie marker, namelijk "C-telopeptide fragmenten van collageen type-II (CTX-II)" en de prevalentie en progressie van radiologische artrose van de heup en de knie. Personen met een hoge CTX-II concentratie (bovenste kwartiel), gemeten in de urine, hadden een 4.2 keer verhoogd kans op het hebben van prevalente radiologische artrose van de heup en de knie vergeleken met personen met een lage CTX-II concentratie (laagste kwartiel). Voor personen met heup pijn vonden we dat deze associatie sterker was vergeleken met personen zonder heup pijn. Personen met een hoge CTX-II concentratie (bovenste kwartiel; baseline meting) hadden een 6.2 keer verhoogd risico op progressie van knie artrose en een 8.3 keer verhoogd risico op progressie van heup artrose vergeleken met personen met een lage CTX-II concentratie. De gevonden associaties waren onafhankelijk van bekende risico factoren voor heup artrose, zoals leeftijd, geslacht en body mass index.

In **hoofdstuk 5** presenteren we welke determinanten het best die personen kunnen identificeren die een groot risico hebben op progressie van heup artrose. Binnen de studie populatie ontwikkelde 13.1% progressie van heup artrose gedurende de follow-up periode. De Kellgren & Lawrence index (baseline meting) bleek een sterke voorspeller te zijn voor progressie, onafhankelijk van variabelen verkregen door een anamnese. Dit gold voornamelijk voor die personen met heup pijn tijdens de baseline meting. In het uiteindelijke model waren een disability index van de onderste extremiteit ≥ 0.5 en een CTX-II concentratie ≥ 235.5 mmol/l ook onafhankelijke voorspellers van progressie. CTX-II bleek, vergeleken met variabelen verkregen door anamnese, lichamelijk onderzoek en röntgenonderzoek, een matige voorspeller te zijn voor progressie van heup artrose. Onze overall conclusie is dat in een klinische situatie en voor klinische trials, een röntgenfoto sterk toegevoegde waarde heeft om die personen te identificeren die een verhoogd risico hebben op progressie van heup artrose.

In **hoofdstuk 6** evalueren we de associatie tussen radiologische acetabulaire dysplasie en incidente radiologische heup artrose bij de follow-up meting, bij personen die bij de baseline meting geen radiologische heup artrose hebben. In deze populatie ontwikkelde 9.3% een incidente radiologische heup artrose gedurende de follow-up periode. Personen met acetabulaire dysplasie hadden een 4.3 verhoogd risico op het ontwikkelen van een incidente heup artrose vergeleken met de personen zonder acetabulaire dysplasie. Deze associatie was onafhankelijk van bekende determinanten van artrose zoals leeftijd, geslacht en body mass index. Verder bleek dat de associatie tussen acetabulaire dysplasie en incidente

heup artrose mogelijk wordt gemodificeerd door vrouwelijk geslacht, zwaar lichamelijk werk en body mass index.

De associaties tussen twee groepen NSAIDs; die NSAIDs (indomethacin, naproxen en ibuprofen) waarvan wordt verondersteld dat ze een negatief effect hebben op het kraakbeen metabolisme en die groep NSAIDs waarvan wordt verondersteld dat ze een neutraal effect hebben op het kraakbeen metabolisme; en progressie van heup en knie artrose worden gepresenteerd in **hoofdstuk 7**. De personen die NSAIDs hebben gebruikt waarvan wordt verondersteld dat deze een neutraal effect hebben op het kraakbeen metabolisme (> 31 dagen) hadden verrassend een 1.7 keer verhoogd risico op progressie van heup artrose vergeleken met de korte termijn NSAIDs (1–30 dagen) gebruikers uit dezelfde groep. Dit verhoogd risico zou mogelijk verklaard kunnen worden door het langdurig gebruik van diclofenac (langer dan 180 dagen). Voor de knie vonden we geen duidelijke associaties tussen NSAID gebruik en progressie van artrose. Of dit gevonden risico verklaard kan worden door een echt negatief effect op kraakbeen metabolisme of door het effect van mechanische belasting op een analgetische heup moet nog verder worden onderzocht.

In de algemene discussie in **hoofdstuk 8** worden de belangrijkste bevindingen van dit proefschrift besproken in een breder perspectief. Allereerst wordt er besproken hoe de aanwezigheid van heup artrose en hoe progressie van heup artrose te definiëren. De discussie richt zich daarna op het nut van een biomarker en tevens welke bronnen van potentiële bias de associatie tussen NSAIDs gebruik en progressie beïnvloeden. Tenslotte worden aanbevelingen gegeven voor toekomstig onderzoek en wat de gevonden resultaten betekenen voor de dagelijkse praktijk van een arts.

Dankwoord

Het meest gelezen hoofdstuk van een proefschrift is, helaas voor de desbetreffende promovendus, het dankwoord. De promovendus bedankt alle personen die in meer of mindere mate een bijdrage hebben geleverd aan het uiteindelijke product, het proefschrift. Gedurende de afgelopen drie jaar, heb ik met veel mensen mogen samenwerken die mij geholpen hebben bij dit enerverende proces van promoveren. Van mijn eerste analyses met SPSS tot aan het voorbereiden van weer een presentatie voor een internationaal congres (EULAR), afgelopen juni 2004 in Berlijn.

Allereerst mijn promotor Prof. Dr. J.M.W. Hazes. Beste Mieke, vanaf ons eerste overleg, was ik onder de indruk van jouw kwaliteit om anderen enthousiast te maken voor wetenschappelijk onderzoek. Je methodologische kennis en klinische blik op analyses hebben mij erg gestimuleerd en natuurlijk veel geleerd. Daarnaast wil ik je bedanken voor alle tijd (en dat was niet weinig) die je in mij geïnvesteerd hebt; hoe een wetenschappelijk artikel te schrijven.

Vervolgens mijn andere promotor Prof. Dr. B.W. Koes. Beste Bart, door onze gedeelde interesses in het bewegingsapparaat en voetbalclub klikte het vanaf het begin. Bedankt voor je methodologisch begeleiding en vooral voor je positieve feedback.

Sita, mijn copromotor, wil ik bedanken voor haar altijd openstaande deur. Jouw enthousiasme voor nieuwe resultaten, volgende versies van manuscripten, acceptaties van abstracts of artikelen is de mooiste feedback die je als promovendus kan krijgen. Ik denk met veel plezier terug aan onze uitgebreide discussies over van alles en nog wat en het maken van plannen voor toekomstig onderzoek. Daarnaast wil ik je bedanken voor de gezellige Eular congressen in Lissabon en Berlijn.

Mijn "derde promotor" Prof. dr. H.A.P. Pols, beste Huib, met name bedankt voor je betrokkenheid bij mijn onderzoek. Ik heb veel geleerd van je kritische klinische blik op statistische analyses en van je adviezen hoe de resultaten te verkopen op een congres of in een artikel. Een "sexy" auteur zal ik waarschijnlijk wel nooit helemaal worden.

Trudie bedankt voor de fraaie voorkant van dit proefschrift, voor de koffie op dinsdag ochtend en voor de gezelligheid.

Roos, mijn interesse voor de statistiek is door het samenwerken met jouw alleen maar groter geworden (waarschijnlijk statistisch significant) en daarnaast natuurlijk bedankt voor de gezellige samenwerking.

Mijn kamergenoten; Pepijn, Ton, Leonie, Esther, Arthur en Umit wil ik bedanken voor de vele discussies en vooral voor de lol die we samen gehad hebben. Het was goed om leuke maar ook frustrerende verhalen te delen.

Pim, ondanks je soms niet te volgen opmerkingen of redenaties hebben we ook vele serieuze en humorvolle gesprekken gehad. Bedankt voor je collegialiteit.

Verder natuurlijk alle Westzeedijkers: Marlies, Petra, Marie-Louise, Yvonne, Frieke, Rebecca, Hans, Hanneke, Corinne, Robbert, Michiel, Arjan, Diana, Anneke, Liesbeth, Heleen, Herman, Sander, Boris en Kees, bedankt voor jullie gezelligheid.

René bedankt voor je hulp ten aanzien van allerlei ondersteunende vragen en natuurlijk voor je vriendschap.

Saeede, Janneke, Rianne en Annet, de artrose onderzoekers bij Huisartsgeneeskunde, bedankt voor de gezellige samenwerking. Je hield je wel niet bezig met artrose, maar toch was het leuk om met je samen te werken, Bionca.

Mijn interesse voor wetenschappelijk onderzoek werd geboren tijdens mijn opleiding in Maastricht. Eén van de stimulerende mensen bij de vakgroep Epidemiologie ruilde Maastricht in voor Rotterdam. Arianne het was leuk om je in Rotterdam weer terug te zien en nog bedankt voor het stimuleren van deze wetenschappelijke interesse.

Eens in de drie maanden stond er een artrose meeting gepland, in samenwerking met Reumatologie, Interne Geneeskunde, Orthopedie en natuurlijk Huisartsgeneeskunde. Het was zeer leerzaam om nieuwe onderzoeksresultaten hier te presenteren en hierover te discussiëren. Vandaar Harry, Gerjo, Elian, Andre, Joyce, Behrooz, Marijke, Erwin, Bart, Mieke, Sita, Huib, Annet, Saeede en Rianne bedankt.

Ook wil ik de deelnemers aan de wekelijkse ERGO-werkbespreking; Hans, Andre, Mariette, Stephanie, Fernando, Joyce en Carola, bedanken voor hun actieve rol binnen dit overleg en de feedback op gepresenteerde resultaten.

Het raakvlak tussen toegepast en fundamenteel onderzoek is een uitermate interessant gebied. Met de vragen op het gebied van kraakbeen en bot synthese kon ik altijd aankloppen bij Gerjo en Harrie, hiervoor mijn dank.

Voor de data verzameling heb ik heel wat röntgenfoto's moeten beoordelen op de aan- of afwezigheid van artrotische kenmerken. Bij het leren beoordelen van de röntgenfoto's heb ik veel geleerd van dr. Ginai, waarvoor mijn dank.

Voor het bepalen van de CTX-II concentratie wil ik Stephan Christgau bedanken.

In het eerste jaar van mijn project heb ik op de afdeling Epidemiologie & Biostatistiek de röntgenfoto's beoordeeld. Mijn kamergenoten Stephanie en Simone wil ik bedanken voor de gezelligheid en steun gedurende dit monniken werk.

Als laatste wil ik Janine en onze kinderen Tim, Amber en Mats bedanken. Vooral ten tijde van mijn studie in Maastricht die veel tijd en energie vroeg stonden jullie altijd voor mij klaar. Ik had niet zo kunnen functioneren zonder jullie liefde en humor.

About the author

Max Reijman was born on October 12, 1963 in Den Haag, the Netherlands. In 1981 he passed secondary school at the "Lodewijk Makeblijde College" in Rijswijk (HAVO). He graduated in 1985 from the "Haagsche Academie voor Lichamelijke Opvoeding, academy for physical therapist training" in Den Haag. After he finished his military service in November 1988, he worked as physical therapist in "Diaconessenhuis" in Leiden, till September 2001. During this period he graduated in 1992 from "Stichting Opleiding Manuele Therapie (SOMT), academy for manual therapist training". In 1996 he started the study Health Sciences at the University of Maastricht where he specialized in Human Movement Science. He graduated in March 2001. In January 2002 he achieved his registration as Master of Science in Epidemiology. In September 2001 he started the work described in this thesis at the Department of General Practice (Prof. Dr. B.W. Koes), the Department of Epidemiology and Biostatistics and the Department of Rheumatology (Prof. Dr. J.M.W. Hazes) at the Erasmus Medical Center in Rotterdam.

Max is married to Janine Reijman-Booi and they have three children, Tim, Amber and Mats.