CHECKing radiographic joint damage in early osteoarthritis

Margot Kinds

The research described in this thesis was financially supported by the Dutch Arthritis Association (Reumafonds).
Printing of this thesis was funded by the Dutch Arthritis Association, the J.E. Jurriaanse Stichting, the Anna fund (Annafonds), Merck Sharp & Dohme B.V., and Pfizer bv.

CHECK is funded by the Dutch Arthritis Association on the lead of a steering committee comprising 16 members with expertise in different fields of osteoarthritis, chaired by Prof. JWJ Bijlsma, and coordinated by J Wesseling. Involved are: Academic Hospital Maastricht; Erasmus Medical Center Rotterdam; Reade (Jan van Breemen Institute)/ VU Medical Center Amsterdam; Kennemer Gasthuis Haarlem; Martini Hospital Groningen/ Allied Health Care Center for Rheumatology and Rehabilitation Groningen; Medical Spectrum Twente Enschede/ Twenteborg Hospital Almelo; St. Maartenskliniek Nijmegen; Leiden University Medical Center; University Medical Center Utrecht; and Wilhelmina Hospital Assen.

Cover designed by Marc Tjin

Printed by Drukkerij Modern, Bennekom

ISBN: 978-90-393-5694-4

M.B. Kinds
University Medical Center Utrecht
Rheumatology & Clinical Immunology
PO Box 85500 (F02.127)
3508 GA Utrecht
The Netherlands

CHECKing radiographic joint damage in early osteoarthritis

CHECKen van radiologische gewrichtsschade in vroege artrose
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan,
ingevolge het besluit van het college voor promoties in het openbaar
te verdedigen op dinsdag 24 januari 2012 des middags te 2.30 uur

door

Margot Bernadette Kinds

geboren op 12 juni 1981 te Ede

Promotoren:	Prof.dr. F.P.J.G. Lafeber
	Prof.dr.ir. M.A. Viergever
	Prof.dr. J.W.J. Bijlsma
Co-promotoren:	Dr.ir. K.L. Vincken
	Dr. A.C.A. Marijnissen

Contents

Chapter 1

Introduction

Osteoarthritis

Osteoarthritis (OA) is a slowly developing disease of the synovial joint(s), characterized by pain and functional disability, and limiting quality of life. Structural changes affect all tissues of the joint and consist of cartilage degeneration, bone involvement by formation of osteophytes and changes in subchondral density, synovial inflammation, and impairment of soft tissue like ligaments and muscles[1].

Prevalence

OA is the most common joint disorder and accounts for more disability among the elderly than any other disease[2]. Although the disease can involve all synovial joints, the knees, hips, hands, and spine are most frequently affected. The impact on individuals is most apparent in case of affection of the larger joints, with symptomatic knee OA affecting 6% and symptomatic hip OA affecting 3% of the adult population (age 30 years and over)[3]. Radiographic OA of the knee is present in 53% of women and 33% of men aged older than 80 years[1]. The disease occurs in all ethnic groups and in all geographic areas.

Etiology

The etiology of OA is considered to be multi-factorial and to differ between individuals, with a role for (combinations of) genetic, metabolic, and mechanical factors. The mechanism of onset and progression remains unclear, although predisposing factors have been identified. Higher age is a risk factor for the development of OA[4], since with increasing age the tissue quality reduces[5] and the hormonal status changes (specifically in females)[6]. Obesity is recognized as an important predisposing factor specifically in the knee joint[7-9], but also in the hip[10,11] and hand joints[8,12,13]. The latter implies that next to mechanical impact, obesity influences metabolic processes (adiponectins) which contributes to the imbalance in synthesis and degradation of cartilage and bone (turnover)[14], and adds to synovial inflammation. Further, inflammation might add to progressive disease, specifically in erosive hand OA[15,16], but this is considered secondary to the degenerative process[2]. In addition to age and obesity, numerous factors have been suggested to add to the development and/or progression of OA. Mechanical factors like joint malalignment[17,18], bone shape[19], joint trauma[20,21] and occupational activities[21], and also joint shape[22] and family history[21] are considered to predispose to the development and/or progression of OA.

Treatment

It is preferred to prevent the development of OA, since no cure for the disease is currently available[23]. But when an individual develops OA, treatment is aimed at relief of symptoms. The first treatment options are non-surgical, either by non-pharmacological or by pharmacological strategies. Non-pharmacological options consist of education, weight loss, exercise, braces and physical therapy. Pharmacological options include pain medication, anti-inflammatory medication, and potentially disease modifying agents[24,25]. In a later (end) phase of the disease surgical options with or without preservation of the joint are indicated[26].

Although joint replacement is quite cost-effective[27] and has good clinical benefit[25], this treatment option should be postponed as long as possible to prevent costly replacement surgery at a later phase with less good outcome than initial joint replacement.

Socioeconomic burden

As a result of ageing of the community and the increasing prevalence of obesity[26], OA will increase in prevalence and disease duration in the coming years (decades). Therefore, the social and economic burden to our society will grow due to limitations in quality of life and the increased need for treatment. Next to direct costs for e.g. medication and hospitalization, also indirect costs for e.g. work-related loss and home-care, add significantly to the economic burden[27].

Diagnosis

In clinical practice the diagnosis of OA is usually based on clinical complaints and then confirmed by radiographic evaluation of structural damage[26]. An inconsistent association between the radiographic characteristics (structural tissue damage) and the clinical symptoms (pain and functional disability) of OA[28-30] hampers the evaluation of the disease however. In fact, in clinical practice radiographs are primarily used to exclude other underlying pathologies responsible for the symptoms of pain and functional disability. Even in case of surgical intervention, clinical symptoms and not radiographic changes are leading in decision-making[31,32].

If structural tissue changes underlie the clinical symptoms, there are several possible explanations for this (apparent) discrepancy between radiographic and clinical characteristics of OA. Radiographs possibly do not show those structural changes that are related to pain (e.g. synovitis). The sensitivity of the outcome measures to evaluate onset and progression of disease might be limited. The way radiographs are presently read and graded, makes it hardly possible to detect subtle progression or treatment effects in a short time span and it is generally appreciated that a significant change in radiographic grade takes at least one or even two years[33,34]. This tempted many researchers to study the use of other imaging techniques such as magnetic resonance imaging (MRI). An important advantage of MRI is that it allows direct evaluation of different joint tissues[35], which may enable the detection of specific structural changes before they become evident on radiographs[36]. Importantly, using MRI a relation was found between pain and bone marrow lesions and bone attrition[37,38]. However, the sensitivity of joint space width (JSW) measurement was found to be similar for MRI and radiography[39]. Simple radiographic evaluation of joint space narrowing (JSN) performed even better than advanced MRI techniques to measure cartilage composition, i.e. delayed gadolinium-enhanced MRI and T2 mapping, when discriminating between knees with and without cartilage thinning[40]. Furthermore, although developments in using molecular markers to detect OA characteristics are ongoing, so far such markers did not perform better than radiography in detecting cartilage thinning[40,41].

In the meantime radiography has been improved by introducing more standardized acquisition protocols[42-45] for digital radiography (the standard nowadays). Although standardization might still need optimization, radiography is non-invasive, cheap, fast, and generally available and continues to be the gold standard for evaluation of structural damage in clinical practice[1,46,47]. Also the food and drug administration (FDA: guidance for industry at www.fda.gov/cder/guidance) still demands radiographic changes to prove disease modifying efficacy of treatment strategies.

Next to the imaging procedures, the heterogeneous manifestation of the disease might also hamper the detection of an association between radiographic and clinical OA. In clinical practice, it is recognized that some patients suffer from severe pain without evident radiographic damage, while other patients have evident radiographic damage and only mild or no symptoms[48]. Translated to structural damage, in some patients synovial inflammation might be most prominent, while in others changes in the bone mainly occur. Also, patients can have affection of an isolated joint or of multiple joints[49]. This implies the existence of different phenotypes of OA[49-51]. In each of these phenotypes the relation between structural changes and clinical symptoms might be different. Furthermore, clinical complaints can fluctuate over time since it is acknowledged that OA related pain has an intermittent character[52], specifically early in disease[53]. Even during the day complaints might vary, e.g. stiffness is known to be more severe in the morning or after long periods of immobilization, and pain may be more intense after demanding activities or less intense thanks to sufficient medication.

As for evaluation of radiographic OA changes, the sensitivity of methods to evaluate clinical OA changes also needs further research. The heterogeneous and subjective character of symptoms makes precise assessment of clinical OA very difficult, which might hamper the detection of an association between radiographic and clinical OA characteristics. On the other hand, structural damage might be a predisposing factor for pain but might not be the origin of the pain[54,55]. Or pain might indeed be originated by structural damage, but evolve to a more chronic pain syndrome in a later stage and then lose the relation with structural damage.

Identifying a relation between clinical and radiographic OA is considered of major relevance to further understand this disabling and costly disease. If such a relation can actually be confirmed, it is sensible to evaluate individuals with early complaints that might lead to the development of OA. By following the course of complaints, by assessing individual characteristics (risk factors) and by evaluating structural changes, specific features might be identified that are of importance in the development of OA. Clinical study design should aim at identifying such specific features by thorough evaluation of individuals from an early phase of the disease when structural changes are not evident yet. This is expected to advance the identification of phenotypes of OA. With that, specific treatment strategies can be developed and moreover treatment can be started from an earlier phase of the disease when symptoms and structural damage are not that disabling yet. To enable such advances

in the understanding of OA, detailed evaluation of validated clinical and radiographic characteristics from an early phase of the disease is a prerequisite.

Clinical osteoarthritis

Symptoms of OA include pain, stiffness, and limitations in daily functioning. Clinical OA is considered to develop over time; pain becomes chronic[2,53,56] and functional limitations increase due to reduced range of motion and decreased muscle strength[57]. When the development of OA is evaluated on the individual level, variation in the assessment of severity commonly occurs[57]. Besides actual fluctuation in severity, the method to evaluate clinical OA might introduce variation. Commonly, the presence of clinical OA is evaluated by reporting whether individuals suffered pain, stiffness or discomfort during most days of at least one month during the past year[56]. As frequently described for other diseases, intensity of pain can be graded on an ordinal scale (qualitative) or measured on a continuous scale (quantitative). To enable evaluation of clinical manifestation more sensitively, specific OA measures are preferred. Therefore the Western Ontario and McMasters Universities Osteoarthritis Index (WOMAC) was designed, which is a disease specific questionnaire including three subscales with five items on pain, two items on stiffness, and seventeen items on function[58]. Although quantitative and OA specific methods are more precise, variation in the assessed level of symptoms occurs within and between individuals. Furthermore, in evaluation of clinical characteristics it needs to be taken into account that some patients suffer from (severe) complaints in one joint while others suffer from (mild) complaints in multiple joints. Specifically when an association with radiographic OA (of the individual or a specific joint) is evaluated, it should be considered whether symptoms are assessed as an overall measure of an individual or of a specific joint.

Radiographic osteoarthritis

OA affects all structures of the joint and comprises cartilage loss, bone remodeling, capsular stretching, muscle weakness, synovitis, laxity of the ligaments, and lesions in the bone marrow[59]. On radiographs structural changes are evaluated by assessment of joint space narrowing (JSN) as a measure for cartilage thickness, osteophyte formation as a measure for bone remodeling, cyst formation, and sclerosis (increased bone density) as a measure for changes in the subchondral bone. Whether bone density indeed increases is not evident however. While the apparent density of the subchondral bone seems to increase, probably by the increased number and decreased separation of trabeculae, the material density of the bone is significantly lower than in individuals without OA[60].

Radiographic OA (progression) is commonly evaluated by use of the Kellgren & Lawrence (K&L) grade[61] (figure 1). This method provides a summary grade (0-IV) for the whole joint and is commonly interpreted as a measure for absence (grade 0 or I) or presence (grade ≥II) of structural damage. Drawbacks of this method are the low sensitivity to change and the assumption of a fixed sequence in the development of radiographic OA characteristics.

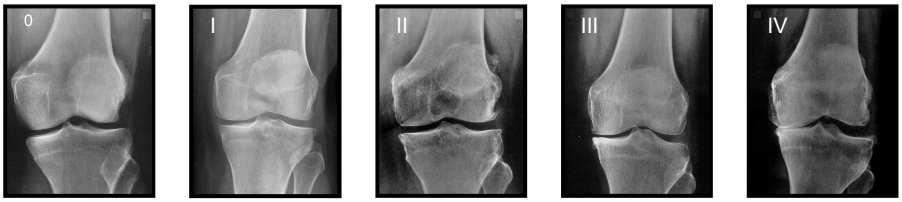

Figure 1 Kellgren & Lawrence grade 0-IV of the knee joint

Assessment of separate radiographic features enables evaluation of whether such a sequence actually exists in all patients. JSN is already frequently evaluated by grading on an ordinal scale[62], but also other features like osteophytes and bone density may be assessed separately[63]. E.g. the Altman atlas uses different radiographic views to evaluate separate features on a 0-3 (or 0-1) scale[64]. However, these ordinal grades are not sensitive to change and are not commonly applied in clinical practice.

The sensitivity to change might be improved when progression is evaluated more thoroughly by quantitative measurement of separate radiographic features. Digital image analysis techniques like Knee Images Digital Analysis (KIDA: figure 2) have been developed for (semi-) automatic measurement of JSW[65-67], but also other radiographic features like varus alignment[68], osteophytes[69], eminence height, and bone density can be measured quantitatively.

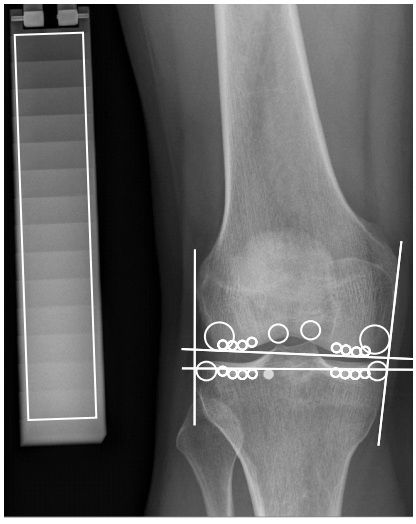

Figure 2 Screenshot of KIDA

For (longitudinal) application of quantitative measurement of radiographic features in clinical practice and for research purposes, it needs to be taken into account that variation in radiographic acquisition settings and joint positioning[70,71] within and between technologists and hospitals occurs. Standardization of radiographic procedures is important for reproducible joint positioning and hence sensitive measurement of separate radiographic OA features[43,72]. The detection of presence, progression, and sequence of separate radiographic features might enable the detection of OA phenotypes with different (combinations of) clinical and radiographic characteristics. And more importantly, the use of quantitative measurement might enable detection of OA features earlier in the process of disease.

Early osteoarthritis

An association between structural damage and clinical symptoms of OA is more commonly reported in patients with established disease[73]. This implies that OA is usually diagnosed in a late stage of disease, when structural damage is already irreversible and treatment options are largely palliative. For better understanding of the disease process and for the development of adequate (preventive) treatment options for all those suffering from OA, focus needs to shift to individuals at high risk of OA development or to those with early disease in which structural changes are still absent or reversible[74]. When phenotypes of OA can be identified this can improve clinical trial design to evaluate specific treatment strategies[1,50]. Individuals at high risk of OA and those in an early phase of OA, are increasingly studied in large cohorts e.g. the population-based Framingham study[75], the MOST study[76], and in recent years the Osteoarthritis Initiative (OAI) which was initiated throughout the United States[77].

Cohort Hip & Cohort Knee (CHECK)

In the Netherlands a longitudinal study was initiated in 2001: the Cohort Hip & Cohort Knee (CHECK) study evaluates development of disease in individuals with early complaints related to hip and/or knee OA[78]. Most of these individuals are expected to develop OA in the following years, some progressively and some slowly. In this large multi-center study 1002 individuals are evaluated for multiple outcome measures. Next to yearly assessment of clinical symptoms using questionnaires and physical examination, blood samples, urine samples, and radiographs of the knees and hips are acquired at least four times during the ten-year follow-up period. Also in subpopulations of CHECK, MRI scanning, specific analyses of skin biopsies (Advanced Glycation Endproducts), and other additional examinations are performed.

Evaluation of these individuals is aimed at improving the understanding of the disease from early onset. Since multiple variables are collected of multiple joints (both knees and hips), it is expected that specific features can be identified that are important in the development of structural damage (amongst other OA characteristics). If phenotypes of OA exist with different patterns of radiographic damage, the relation between radiographic joint damage

and clinical symptoms can be studied more thoroughly by measuring specific radiographic features. Accordingly, the design of clinical trials aimed at developing targeted treatment for osteoarthritis can be facilitated in the future, by including only individuals with specific osteoarthritis features.

Outline of this thesis

The research described in this thesis was aimed at improving the understanding of the onset and development of radiographic joint damage due to OA, since radiography is still the gold standard to evaluate structural changes. Therefore, separate features of radiographic damage are measured quantitatively by digital image analysis in CHECK participants with very early complaints related to OA of knee and/or hip. The challenges in measurement of separate features are evaluated, and also the value of these measurements for application in clinical practice and in detecting a relation between damage and symptoms.

Section I: Methodology

There is ongoing debate on whether an association between radiographic and clinical characteristics of OA actually exists since only a limited number of studies reported an evident association. In this section of the thesis the hypothesis that the inconsistency in the association is caused by limited methodological quality, is studied. Therefore, in **chapter 2** the question was: *Which methodological criteria are important to detect an association between radiographic and clinical OA of hip and knee?* A systematic review was performed to evaluate whether the definition of OA for study inclusion, the radiographic protocol, and the outcome measures for radiographic and clinical characteristics are of importance to detect an association.

According to the results from this review, the CHECK study appeared suitable to evaluate the role of methodological quality in the evaluation of radiographic OA in a clinical study. In CHECK, participants were included with symptoms related to OA, and knee and hip radiographs were (are) acquired according to a standardized protocol. Moreover, in CHECK multiple outcome measures for radiographic and clinical OA were (are) collected. This enables the evaluation of the value of digital analysis in the quantitative measurement of radiographic OA features and the evaluation of an association between such features and clinical OA. In chapters 3 and 4 the ability to measure radiographic status and progression by digital image analysis was compared to conventional grading methods that are considered the gold standard.

In **chapter 3** the question was: *Does the quantitative measurement by knee images digital analysis (KIDA) result in higher sensitivity to detect progression of radiographic knee damage than qualitative grading according to the Altman atlas?* Quantitative measurement

by digital image analysis was compared to ordinal grading for the ability to detect changes in separate radiographic parameters of knee damage from baseline to two-year follow-up in CHECK.

In **chapter 4** it was described whether radiographic features represented the severity of OA at baseline, as defined according to conventional K&L grading. It was also studied whether radiographic features were related between different hip and knee joints within individuals and thus represented characteristics of an individual in addition to the severity of OA. The question was: *To what extent do radiographic features of knees and hips represent characteristics of an individual, in addition to osteoarthritis severity?* If a relation between joints of an individual is found, it might be useful to take into account the radiographic characteristics of different joints within individuals when studying a relation with clinical OA.

Section II: Clinical study

When evaluating joint damage in CHECK, the measurements of radiographic features are expected to cover a wide range of quantitative values. This range will partly represent variation between individuals, but will also represent the radiographic joint damage. This radiographic joint damage may affect the whole joint and even multiple joints, but may also affect one or more specific features more than others. In this second section it is evaluated how separate features can best be used in clinical practice to evaluate onset and progression of radiographic OA of the knee joint.

In **chapter 5** the use of quantitatively measured radiographic features was analyzed by a cross-sectional approach. By use of baseline, two-year and five-year follow-up measurement of radiographic features the development of the separate features, the correlation between these features, and the relation with clinical characteristics were evaluated. The question was: *Can specific features measured by knee images digital analyses be used to evaluate radiographic OA development over time from an early phase of the disease, and can these features be related to each other and to clinical characteristics of OA?*

To further evaluate the application in a clinical study, it was studied whether specific features at baseline were predictors for the onset of radiographic OA and for the persistence and progression of clinical OA in CHECK. The predictive value of the quantitatively measured radiographic features was compared to that of using clinical and demographic characteristics only, and to that of using K&L grading. **Chapter 6** describes *whether and which separate features, measured on knee radiographs of individuals with recent onset of knee pain, are associated with incidence of radiographic OA and persistence and/or progression of clinical OA during five-year follow-up.* When features are detected that predict unfavorable outcome, individuals that are susceptible for OA development can be identified in an early phase of the disease.

In **chapter 7** the question was: *Can phenotypes of progression of radiographic knee OA be identified by quantitative measurement of separate radiographic features?* If phenotypes of radiographic OA can be identified that differ in the level of severity, the phase of disease, and in clinical characteristics, this gives opportunities for the design of clinical trials to develop more tailored treatment strategies.

Section III: Radiographic acquisition

From the above mentioned studies it was learned that the variation in the measurement of separate features (and hence the application in clinical studies) is for a large part dependent on the quality of acquisition of radiographs. In a (multicenter) clinical study slight changes are likely to occur during the acquisition of radiographs, despite the use of a standardized image acquisition protocol. The influence of these changes is investigated and discussed in the third section of this thesis (chapter 8 and 9).

The acquisition of posteroanterior knee radiographs according to the standardized protocol as used in CHECK is subject of study in **chapter 8**. The question was: *What is the influence of changes in knee position during acquisition of radiographs on the measurement of radiographic characteristics?* To describe the implications for use of these measurements in clinical practice, the influence of changes in position was compared to radiographic differences that occurred due to OA during two-year follow-up in CHECK.

Chapter 9 describes whether plain digital radiography, which has almost completely replaced conventional film-screen radiography in the past years, can be used to evaluate bone density changes. With the transition to digital radiography additional post-processing was introduced, which optimized contrast and reduced noise for clinical reading but influenced the gray scale of the projected bone. It was evaluated whether bone density values were reliably measured by normalization of gray values of the bone to gray values of an aluminum reference in radiographs acquired according to clinical practice. The question was: *Is measurement of bone density feasible using plain digital radiographs and what are the effects of acquisition and post-processing settings on this measurement?* If bone density evaluation is possible on radiographs that will be acquired in individuals suspected of OA, this removes the need for additional imaging techniques and this will limit costs and X-ray exposure.

Finally, in **chapter 10** the results of the previous chapters are summarized and integrated. Chapter 10 also discusses the use of measurement of radiographic features of joint damage is reflected on in individuals presenting themselves to a physician with early complaints suspected for OA.

References

1. Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: an update with relevance for clinical practice. Lancet. 2011 Jun 18;377(9783):2115-26.
2. Felson DT. Developments in the clinical understanding of osteoarthritis. Arthritis Res Ther. 2009 Jan 30;11(1):203.
3. Felson DT, Zhang Y. An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. Arthritis Rheum. 1998 Aug;41(8):1343-55.
4. Oliveria SA, Felson DT, Reed JI, et al. Incidence of symptomatic hand, hip, and knee osteoarthritis among patients in a health maintenance organization. Arthritis Rheum. 1995 Aug;38(8):1134-41.
5. Carrington JL. Aging bone and cartilage: cross-cutting issues. Biochem Biophys Res Commun. 2005 Mar 18;328(3):700-8.
6. Zhang Y, Jordan JM. Epidemiology of osteoarthritis. Rheum Dis Clin North Am. 2008 Aug;34(3):515-29.
7. Reijman M, Pols HA, Bergink AP, et al. Body mass index associated with onset and progression of osteoarthritis of the knee but not of the hip: the Rotterdam Study. Ann Rheum Dis. 2007 Feb;66(2):158-62.
8. Grotle M, Hagen KB, Natvig B, et al. Obesity and osteoarthritis in knee, hip and/or hand: an epidemiological study in the general population with 10 years follow-up. BMC Musculoskelet Disord. 2008;9:132.
9. Jiang L, Tian W, Wang Y, et al. Body mass index and susceptibility to knee osteoarthritis: A systematic review and meta-analysis. Joint Bone Spine. 2011 Jul 29.
10. Karlson EW, Mandl LA, Aweh GN, et al. Total hip replacement due to osteoarthritis: the importance of age, obesity, and other modifiable risk factors. Am J Med. 2003 Feb 1;114(2):93-8.
11. Heliovaara M, Makela M, Impivaara O, et al. Association of overweight, trauma and workload with coxarthrosis. A health survey of 7,217 persons. Acta Orthop Scand. 1993 Oct;64(5):513-8.
12. Cicuttini FM, Baker JR, Spector TD. The association of obesity with osteoarthritis of the hand and knee in women: a twin study. J Rheumatol. 1996 Jul;23(7):1221-6.
13. Oliveria SA, Felson DT, Cirillo PA, et al. Body weight, body mass index, and incident symptomatic osteoarthritis of the hand, hip, and knee. Epidemiology. 1999 Mar;10(2):161-6.
14. Sowers M. Epidemiology of risk factors for osteoarthritis: systemic factors. Curr Opin Rheumatol. 2001 Sep;13(5):447-51.
15. Banks SE. Erosive osteoarthritis: a current review of a clinical challenge. Clin Rheumatol. 2011 Jul;29(7):697-706.
16. Vlychou M, Koutroumpas A, Malizos K, et al. Ultrasonographic evidence of inflammation is frequent in hands of patients with erosive osteoarthritis. Osteoarthritis Cartilage. 2009 Oct;17(10):1283-7.
17. Chapple CM, Nicholson H, Baxter GD, et al. Patient characteristics that predict progression of knee osteoarthritis: A systematic review of prognostic studies. Arthritis Care Res (Hoboken). 2011 Aug;63(8):1115-25.
18. Yusuf E, Bijsterbosch J, Slagboom PE, et al. Body mass index and alignment and their interaction as risk factors for progression of knees with radiographic signs of osteoarthritis. Osteoarthritis Cartilage. 2011 Sep;19(9):1117-22.
19. Sharma L, Song J, Felson DT, et al. The role of knee alignment in disease progression and functional decline in knee osteoarthritis. Jama. 2001 Jul 11;286(2):188-95.
20. Lohmander LS, Ostenberg A, Englund M, et al. High prevalence of knee osteoarthritis, pain, and functional limitations in female soccer players twelve years after anterior cruciate ligament injury. Arthritis Rheum. 2004 Oct;50(10):3145-52.
21. Zhang W, McWilliams DF, Ingham SL, et al. Nottingham knee osteoarthritis risk prediction models. Ann Rheum Dis. 2011 May 25.
22. Haverkamp DJ, Schiphof D, Bierma-Zeinstra SM, et al. Variation in joint shape of osteoarthritic knees. Arthritis Rheum. 2011 Aug 2.
23. Michael JW, Schluter-Brust KU, Eysel P. The epidemiology, etiology, diagnosis, and treatment of osteoarthritis of the knee. Dtsch Arztebl Int. 2010 Mar;107(9):152-62.
24. Zhang W, Moskowitz RW, Nuki G, et al. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. Osteoarthritis Cartilage. 2008 Feb;16(2):137-62.
25. Zhang W, Doherty M, Arden N, et al. EULAR evidence based recommendations for the management of hip osteoarthritis: report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). Ann Rheum Dis. 2005 May;64(5):669-81.
26. Hunter DJ, Felson DT. Osteoarthritis. Bmj. 2006 Mar 18;332(7542):639-42.

27. Bitton R. The economic burden of osteoarthritis. Am J Manag Care. 2009 Sep;15(8 Suppl):S230-5.
28. Bedson J, Croft PR. The discordance between clinical and radiographic knee osteoarthritis: a systematic search and summary of the literature. BMC Musculoskelet Disord. 2008;9:116.
29. Belo JN, Berger MY, Reijman M, et al. Prognostic factors of progression of osteoarthritis of the knee: a systematic review of observational studies. Arthritis Rheum. 2007 Feb 15;57(1):13-26.
30. Salaffi F, Carotti M, Grassi W. Health-related quality of life in patients with hip or knee osteoarthritis: comparison of generic and disease-specific instruments. Clin Rheumatol. 2005 Feb;24(1):29-37.
31. Birrell F, Afzal C, Nahit E, et al. Predictors of hip joint replacement in new attenders in primary care with hip pain. Br J Gen Pract. 2003 Jan;53(486):26-30.
32. Gossec L, Paternotte S, Bingham CO, 3rd, et al. OARSI/OMERACT Initiative to Define States of Severity and Indication for Joint Replacement in Hip and Knee Osteoarthritis. An OMERACT 10 Special Interest Group. J Rheumatol. 2011 Aug;38(8):1765-9.
33. Boegard TL, Rudling O, Petersson IF, et al. Joint space width of the tibiofemoral and of the patellofemoral joint in chronic knee pain with or without radiographic osteoarthritis: a 2-year follow-up. Osteoarthritis Cartilage. 2003 May;11(5):370-6.
34. Vignon E, Conrozier T, Hellio Le Graverand MP. Advances in radiographic imaging of progression of hip and knee osteoarthritis. J Rheumatol. 2005 Jun;32(6):1143-5.
35. Hunter DJ, Zhang W, Conaghan PG, et al. Systematic review of the concurrent and predictive validity of MRI biomarkers in OA. Osteoarthritis Cartilage. 2011 May;19(5):557-88.
36. Javaid MK, Lynch JA, Tolstykh I, et al. Pre-radiographic MRI findings are associated with onset of knee symptoms: the most study. Osteoarthritis Cartilage. 2010 Nov 11.
37. Torres L, Dunlop DD, Peterfy C, et al. The relationship between specific tissue lesions and pain severity in persons with knee osteoarthritis. Osteoarthritis Cartilage. 2006 Oct;14(10):1033-40.
38. Felson DT, Niu J, Guermazi A, et al. Correlation of the development of knee pain with enlarging bone marrow lesions on magnetic resonance imaging. Arthritis Rheum. 2007 Sep;56(9):2986-92.
39. Duryea J, Neumann G, Niu J, et al. Comparison of radiographic joint space width with magnetic resonance imaging cartilage morphometry: analysis of longitudinal data from the Osteoarthritis Initiative. Arthritis Care Res (Hoboken). 2010 Jul;62(7):932-7.
40. Eckstein F, Le Graverand MP, Charles HC, et al. Clinical, radiographic, molecular and MRI-based predictors of cartilage loss in knee osteoarthritis. Ann Rheum Dis. 2011 Jul;70(7):1223-30.
41. van Spil WE, DeGroot J, Lems WF, et al. Serum and urinary biochemical markers for knee and hip-osteoarthritis: a systematic review applying the consensus BIPED criteria. Osteoarthritis Cartilage. 2010 May;18(5):605-12.
42. Buckland-Wright C. Protocols for precise radio-anatomical positioning of the tibiofemoral and patellofemoral compartments of the knee. Osteoarthritis Cartilage. 1995 Sep;3 Suppl A:71-80.
43. Mazzuca SA, Brandt KD, Buckwalter KA, et al. Field test of the reproducibility of the semiflexed metatarsophalangeal view in repeated radiographic examinations of subjects with osteoarthritis of the knee. Arthritis Rheum. 2002 Jan;46(1):109-13.
44. Vignon E, Conrozier T, Piperno M, et al. Radiographic assessment of hip and knee osteoarthritis. Recommendations: recommended guidelines. Osteoarthritis Cartilage. 1999 Jul;7(4):434-6.
45. Lequesne MG, Laredo JD. The faux profil (oblique view) of the hip in the standing position. Contribution to the evaluation of osteoarthritis of the adult hip. Ann Rheum Dis. 1998 Nov;57(11):676-81.
46. Buckland-Wright JC. Quantitative radiography of osteoarthritis. Ann Rheum Dis. 1994 Apr;53(4):268-75.
47. Nevitt MC. Definition of hip osteoarthritis for epidemiological studies. Ann Rheum Dis. 1996 Sep;55(9):652-5.
48. Hellio Le Graverand MP, Mazzuca S, Duryea J, et al. Radiographic grading and measurement of joint space width in osteoarthritis. Rheum Dis Clin North Am. 2009 Aug;35(3):485-502.
49. Felson DT. Identifying different osteoarthritis phenotypes through epidemiology. Osteoarthritis Cartilage. 2010 May;18(5):601-4.
50. Bierma-Zeinstra SM, Verhagen AP. Osteoarthritis subpopulations and implications for clinical trial design. Arthritis Res Ther. 2011 Apr 5;13(2):213.
51. Kerkhof HJ, Meulenbelt I, Akune T, et al. Recommendations for standardization and phenotype definitions in genetic studies of osteoarthritis: the TREAT-OA consortium. Osteoarthritis Cartilage. 2011 Mar;19(3):254-64.
52. Allen KD, Coffman CJ, Golightly YM, et al. Daily pain variations among patients with hand, hip, and knee osteoarthritis. Osteoarthritis Cartilage. 2009 Apr 17.
53. Neogi T, Nevitt MC, Yang M, et al. Consistency of knee pain: correlates and association with function. Osteoarthritis Cartilage. 2010 Oct;18(10):1250-5.
54. Peat G, Greig J, Wood L, et al. Diagnostic discordance: we cannot agree when to call knee pain 'osteoarthritis'. Fam Pract. 2005 Feb;22(1):96-102.

55. Creamer P, Lethbridge-Cejku M, Costa P, et al. The relationship of anxiety and depression with self-reported knee pain in the community: data from the Baltimore Longitudinal Study of Aging. Arthritis Care Res. 1999 Feb;12(1):3-7.
56. Hawker GA, Stewart L, French MR, et al. Understanding the pain experience in hip and knee osteoarthritis--an OARSI/OMERACT initiative. Osteoarthritis Cartilage. 2008 Apr;16(4):415-22.
57. van Dijk GM, Veenhof C, Spreeuwenberg P, et al. Prognosis of limitations in activities in osteoarthritis of the hip or knee: a 3-year cohort study. Arch Phys Med Rehabil. 2010 Jan;91(1):58-66.
58. Bellamy N, Buchanan WW, Goldsmith CH, et al. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol. 1988 Dec;15(12):1833-40.
59. Felson DT. Clinical practice. Osteoarthritis of the knee. N Engl J Med. 2006 Feb 23;354(8):841-8.
60. Grynpas MD, Alpert B, Katz I, et al. Subchondral bone in osteoarthritis. Calcif Tissue Int. 1991 Jul;49(1):20-6.
61. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. 1957 Dec;16(4):494-502.
62. Felson DT, Nevitt MC, Yang M, et al. A new approach yields high rates of radiographic progression in knee osteoarthritis. J Rheumatol. 2008 Oct;35(10):2047-54.
63. Nelson AE, Braga L, Renner JB, et al. Characterization of individual radiographic features of hip osteoarthritis in African American and White Women and Men: The Johnston County Osteoarthritis Project. Arthritis Care Res (Hoboken). 2010 Feb;62(2):190-7.
64. Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. Osteoarthritis Cartilage. 2007;15 Suppl A:A1-56.
65. Buckland-Wright JC, Ward RJ, Peterfy C, et al. Reproducibility of the semiflexed (metatarsophalangeal) radiographic knee position and automated measurements of medial tibiofemoral joint space width in a multicenter clinical trial of knee osteoarthritis. J Rheumatol. 2004 Aug;31(8):1588-97.
66. Conrozier T, Lequesne M, Favret H, et al. Measurement of the radiological hip joint space width. An evaluation of various methods of measurement. Osteoarthritis Cartilage. 2001 Apr;9(3):281-6.
67. Schmidt JE, Amrami KK, Manduca A, et al. Semi-automated digital image analysis of joint space width in knee radiographs. Skeletal Radiol. 2005 Oct;34(10):639-43.
68. Wong AK, Inglis D, Beattie KA, et al. Reproducibility of computer-assisted joint alignment measurement in OA knee radiographs. Osteoarthritis Cartilage. 2008 Oct 17.
69. Oka H, Muraki S, Akune T, et al. Fully automatic quantification of knee osteoarthritis severity on plain radiographs. Osteoarthritis Cartilage. 2008 Apr 17.
70. Ravaud P, Giraudeau B, Auleley GR, et al. Variability in knee radiographing: implication for definition of radiological progression in medial knee osteoarthritis. Ann Rheum Dis. 1998 Oct;57(10):624-9.
71. Auleley GR, Duche A, Drape JL, et al. Measurement of joint space width in hip osteoarthritis: influence of joint positioning and radiographic procedure. Rheumatology (Oxford). 2001 Apr;40(4):414-9.
72. Buckland-Wright JC, Wolfe F, Ward RJ, et al. Substantial superiority of semiflexed (MTP) views in knee osteoarthritis: a comparative radiographic study, without fluoroscopy, of standing extended, semiflexed (MTP), and schuss views. J Rheumatol. 1999 Dec;26(12):2664-74.
73. White DK, Zhang Y, Niu J, et al. Do worsening knee radiographs mean greater chance of severe functional limitation? The multicenter osteoarthritis study. Arthritis Care Res (Hoboken). 2010 May 6.
74. Hunter DJ. Lower extremity osteoarthritis management needs a paradigm shift. Br J Sports Med. 2011 Feb 6.
75. Felson DT, Naimark A, Anderson J, et al. The prevalence of knee osteoarthritis in the elderly. The Framingham Osteoarthritis Study. Arthritis Rheum. 1987 Aug;30(8):914-8.
76. Roemer FW, Guermazi A, Javaid MK, et al. Change in MRI-detected subchondral bone marrow lesions is associated with cartilage loss: the MOST Study. A longitudinal multicentre study of knee osteoarthritis. Ann Rheum Dis. 2009 Sep;68(9):1461-5.
77. Eckstein F, Maschek S, Wirth W, et al. One year change of knee cartilage morphology in the first release of participants from the Osteoarthritis Initiative progression subcohort: association with sex, body mass index, symptoms and radiographic osteoarthritis status. Ann Rheum Dis. 2009 May;68(5):674-9.
78. Wesseling J, Dekker J, van den Berg WB, et al. CHECK (Cohort Hip and Cohort Knee): similarities and differences with the Osteoarthritis Initiative. Ann Rheum Dis. 2009 Sep;68(9):1413-9.

Chapter 2

A systematic review of the association between radiographic and clinical osteoarthritis of hip and knee

MB Kinds[1,2]
PMJ Welsing[1,3]
EP Vignon[4]
JWJ Bijlsma[1]
MA Viergever[2]
ACA Marijnissen[1]
FPJG Lafeber[1]

[1]Rheumatology & Clinical Immunology, University Medical Center Utrecht
[2]Image Sciences Institute, University Medical Center Utrecht
[3]Julius Center for Health Sciences & Primary Care, University Medical Center Utrecht
[4]Rheumatology, Lyon-Sud University Hospital, Pierre-Bénite, France

Osteoarthritis & Cartilage 2011 Jul;19(7):768-778

Abstract

Objective

There is ongoing debate on whether an association between radiographic and clinical osteoarthritis (OA) exists. We hypothesized that the inconsistency in the detection of an association might be caused by different definitions of OA, by different radiographic protocols, and by scoring methods for radiographic damage and symptoms. The goal of this study was to evaluate which methodological criteria are important to detect an association between radiographic and clinical OA of hip and knee.

Methods

A literature search was performed with the keywords 'OA', 'hip', 'knee', 'radiographic', and 'clinical' and results were screened for relevant studies. Quality criteria for study characteristics and methodology were developed. Studies were classified according to these criteria and the presence of an association between radiographic and clinical OA was scored. The importance of methodological quality and patient characteristics on the presence of an association was evaluated.

Results

The literature search resulted in 39 studies describing an association between radiographic and clinical OA. The frequency of an association between radiographic and clinical OA outcome measures diminished when less quality criteria were fulfilled. Specifically the criterion for standardized outcome measures appeared important in the detection of an association. The association was not influenced by patient characteristics. Only four studies were identified that fulfilled all quality criteria and in these studies an association was found for the knee joint and an inconsistent association was found for the hip joint.

Conclusion

Methodological quality criteria are of importance to reveal an association between radiographic and clinical OA.

Introduction

Osteoarthritis (OA) is a slowly progressive degenerative joint disease, characterized by pain and functional disability. The larger joints are commonly affected and specifically involvement of the hip and knee joint has a great health (care) and economic burden. Diagnosis of OA is usually based on symptoms (clinical OA) and is confirmed by radiography[1]. An inconsistent association between radiographic and clinical OA hampers diagnosis however[2,3]. In clinical practice expression of disease varies significantly between patients, possibly implying the existence of different types of OA. Despite this inconsistency and the development of magnetic resonance imaging, with which a relation between pain and structural damage like bone marrow lesions and bone attrition was found[4,5], radiographs are still the gold standard for demonstrating structural changes since image acquisition is non-invasive, cheap, fast, and generally available[6,7].

Various outcome measures for radiographic and clinical OA are described in studies. Common outcomes for radiographic OA are Kellgren & Lawrence (K&L) grading[8] and in recent years actual measurement of joint space width (JSW) has been increasingly applied[9,10]. A limitation of radiographic evaluation is that, except for the direct evaluation of bone, the tissues involved in the OA process are either evaluated indirectly (cartilage) or not at all (synovium). In evaluation of clinical OA, the visual analogue scale (VAS) for pain and the Western Ontario and McMaster Universities OA Index (WOMAC)[11,12] scores for pain, stiffness, and function are validated and commonly used outcome measures.

To demonstrate an association between radiographic and clinical hip and/or knee OA the standardization of radiographic protocols might be important[13]. Also in recent years the importance of multiple radiographic knee views has been illustrated. Inclusion of patellofemoral (PF) radiographs improves the sensitivity to identify radiographic knee OA by symptoms (like pain)[14,15]. In the hip, joint space narrowing (JSN) is detected in more patients when faux profile radiographs are acquired complementary to anteroposterior (AP) radiographs[16,17].

The objective of this review is to evaluate whether the association between radiographic and clinical OA of the hip and/or knee is influenced by methodological quality and study characteristics. Therefore quality criteria are defined including OA definition for inclusion, acquisition of radiographs according to a standardized protocol, and the use of standardized outcome measures. In addition, the influence of patient characteristics on revealing an association is evaluated.

Methods

Identification of the literature

The literature was searched for studies on radiographic and clinical hip and/or knee OA. To identify studies a search was made in the PubMed search engine with the keywords:

1 ‘osteoarthritis’ or ‘osteoarthrosis’ or ‘OA’ or ‘arthritis’ or ‘arthrosis’;
2 ‘hip’ or ‘knee’;
3 ‘radiograph*’ or ‘radiolog*’ or ‘joint space’ or ‘osteophytes’ or ‘subchondral sclerosis’ or ‘bone mineral density’ or ‘BMD’ or ‘K&L’ or ‘Kellgren & Lawrence’ or ‘Altman’ or ‘Croft’;
4 ‘clinical osteoarthritis’ or ‘clinical outcome*’ or ‘clinical measure*’ or ‘practical outcome*’ or ‘practical measure*’ or ‘pain’ or ‘strength’ or ‘SF-36’ or ‘VAS’ or ‘Visual Analogue’ or ‘function*’ or ‘power’ or ‘symptoms’ or ‘disability’ or ‘quality of life’ or ‘WOMAC’ or ‘Western Ontario’ or ‘ADL’ or ‘stiffness’ or ‘Lequesne’ or ‘questionnaire’ or ‘daily activities’.

The search string was combined as #1 AND #2 AND (#3 OR #4) and was performed on Title/Abstract. Figure 1 depicts the flow chart for study inclusion in the present review.

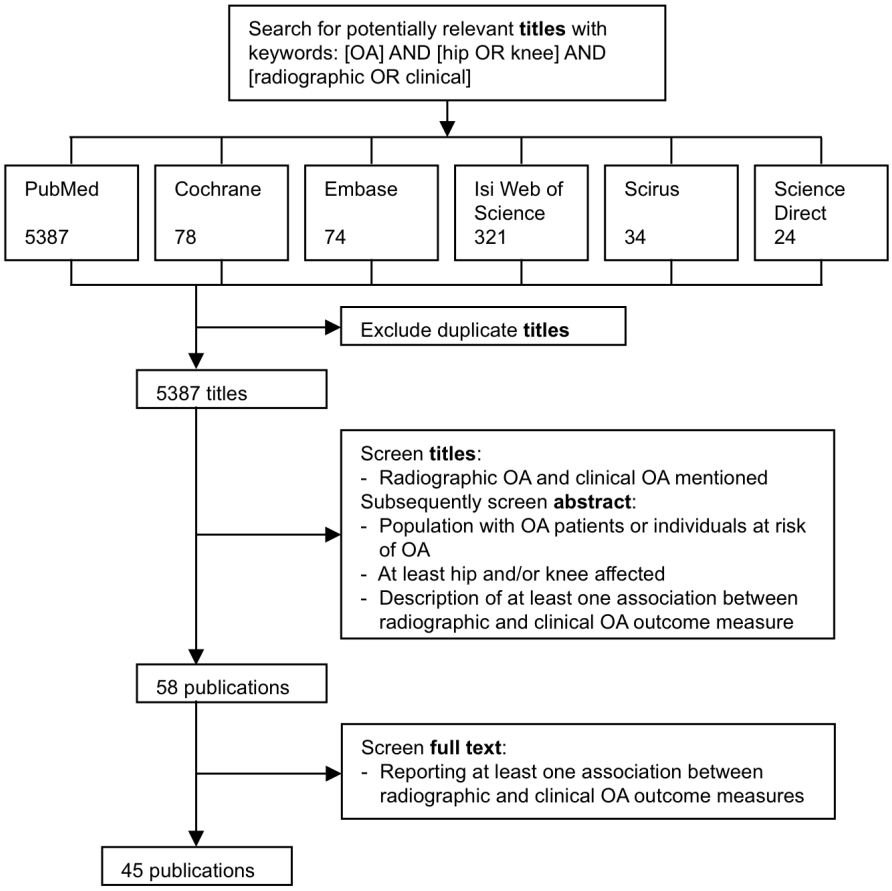

Figure 1 Flow chart of literature search

First, titles were screened for whether radiographic OA and clinical OA were mentioned in any way. Next, abstracts were screened for describing an association between a radiographic and a clinical outcome measure of OA in hip and/or knee joints. Finally, in case of doubt the full text publications were screened and studies were excluded if an association was not evaluated, and when prognosis was performed (i.e. prediction of the outcome over time instead of evaluating an association at the same time point). The search was performed in August 2009 and was limited to studies published in the English language that were added in PubMed from January 1990 since the WOMAC Index, a validated measure for clinical OA, dates from 1988 and is only used frequently from 1990. The search was repeated in the Cochrane Library, Embase, ISI Web of Knowledge, Scirus, and ScienceDirect databases. In this case search strings for #3 radiographic and #4 clinical OA were combined by 'AND' instead of 'OR' to narrow the initial broad approach since a great number of studies in the PubMed search was not relevant and all relevant publications would have been identified with this narrower search. Additionally PubMed was separately searched for authors from the field (e.g. Kellgren JH, Lawrence JS, and Buckland-Wright C). The search was extended, by screening the references of relevant publications identified.

Table 1 Quality criteria for study characteristics and methodology

Study characteristics

1) Description of study population
2) Size of population studied: number of patients > 100
3) Design:
   - Cross-sectional study; and/or
   - Longitudinal study with follow-up period of ≥6 months and a total drop-out ≤20% with information on reason of drop-out
4) Appropriate statistical analysis techniques and presentation of outcome measures as defined under 7): presentation of group percentage, mean, regression coefficient (ß), odds ratios (OR) and standard deviation (SD) or 95%confidence interval (95%CI)

Methodology

5) Inclusion based on osteoarthritis definition:
   - American College of Rheumatology (ACR) criteria for radiographic and/or clinical osteoarthritis; and/or
   - Kellgren & Lawrence grade of standardized radiographs; and/or
   - Osteoarthritis related pain with exclusion of arthritis
6) Radiographic protocol for hip and/or knee:
   - Hip:
     - Anteroposterior (AP) view of pelvis
   - Knee:
     - Posteroanterior (PA) semiflexed view of tibiofemoral joint (MTP); and
     - Skyline view of patellofemoral joint
7) Study describes at least 1 radiographic and 1 clinical outcome measure:
   - Radiographic:
     - Kellgren & Lawrence grade; and/or
     - Joint space width measured on continuous scale or scored on ordinal scale of at least 4 categories; and/or
     - Osteophytes measured on continuous scale or scored on ordinal scale of at least 4 categories
   - Clinical:
     - Pain measured on continuous scale (e.g. visual analogue scale; VAS) or ordinal scale of at least 4 categories (e.g. Lequesne Index); and/or
     - Western Ontario and McMaster Universities Osteoarthritis (WOMAC) Index: at least pain or physical function subscale

Quality criteria for study characteristics and methodology

Studies on (an association between) radiographic and clinical OA indeed reported numerous outcome measures of OA and therefore interpretation of the results of these studies was difficult. To enable evaluation of an association between radiographic and clinical OA in more homogeneous subgroups of patients, a list of quality criteria (table 1) was developed as suggested in previous studies[13,18,19] and as discussed with experts in the field (EV, JB, FL). The seven quality criteria were scored as fulfilled (+) or not fulfilled (-). Scoring was done twice for all studies, and three times for studies where disagreement existed between first and second scoring (MK).

The quality criteria consisted of general study characteristics and specific methodological quality of studies. Criteria 1) to 4) were developed to evaluate general study characteristics. An appropriate description of the study population was defined as at least the reporting of age and gender. The sample size was defined as sufficient if more than 100 patients were evaluated. The study design was scored as fulfilled when cross-sectional or longitudinal evaluation was described. Papers were not selected in case of prognosis. The criterion for sufficient statistical analysis and data presentation was fulfilled when the methods of statistical analysis were clearly stated and when the results were presented in common statistical measures with confidence intervals. The methodological quality was evaluated by criteria 5), 6), and 7). Criterion '5) inclusion' required predefined criteria to enable evaluation of patients that were indeed developing or suffering from OA. Patient inclusion was required to be based on structural damage (American College of Rheumatology (ACR) or K&L) or on OA symptoms (ACR or pain). Pain was not allowed to be due to arthritis to rule out inflammation related symptoms, since this review aimed on evaluating the association between structural damage and symptoms. Criterion '6) protocol' defined the standardization of radiographic protocols and multiple radiographic views. Since the AP pelvis view is most commonly used, this was required for the evaluation of radiographic hip OA. For evaluation of knee OA the posteroanterior semiflexed (metatarsophalangeal: MTP) view of the tibiofemoral (TF) joint was required since with this protocol the joint positioning was most accurate and reproducible[20-22]. A skyline radiograph was required since adding a PF view improves the sensitivity of symptoms for identifying radiographic knee OA[14,15]. Criterion '7) outcomes' concerned validated outcome measures on a scale of at least four categories for structural damage (radiographic OA) and clinical symptoms, like K&L [8] and WOMAC score (criteria: table 1)[11,12].

Analysis of the association

For all studies included in this review, the reported associations between outcome measures of radiographic and clinical OA were evaluated. Associations were scored as present (+) when a statistically significant association was reported between radiographic and clinical OA (as defined in the study) for all described outcome measures. An association was considered absent (-) when no statistically significant associations were present. An association was scored as inconsistent (+/-) when the comparison between radiographic

and clinical OA resulted in significant associations between some outcomes, and non-significant associations between other outcomes. In studies where both hip and knee joints were reported, the associations were evaluated only for the joint that fulfilled criterion '6) protocol (hip only or hip and knee). When a study reported multiple radiographic and clinical outcome measures and some of these fulfilled criterion '7) outcomes' (and some did not), only associations between these 'quality' outcome measures were assessed and studies were scored accordingly.

The strengths of associations were not reported since the estimates of strength of association and the specific comparisons were very diverse and were not always clearly described.

The influence of fulfilling criteria for study characteristics and methodological quality on the presence of an association was analyzed. First it was evaluated whether the frequency of an association (+), an inconsistent association (+/-), and no association (-) significantly changed when the number of fulfilled quality criteria diminished using analysis of variance (ANOVA). Second, the frequency of an association, an inconsistent association, and no association was evaluated for the specific methodological criteria. Chi-square tests were used to evaluate whether the frequencies of present, inconsistent, and absent associations were significantly different for the subgroup of studies that fulfilled a criterion (e.g. '5) inclusion') and for the subgroup of studies that did not fulfill that criterion. Concerning criterion '7) outcomes' the radiographic and clinical outcomes for OA were diverse and even in the studies that fulfilled this criterion comparisons could vary. Third, frequencies of associations were compared between subgroups of studies (using Chi-square tests). Frequencies of associations were compared between studies evaluating either the hip or knee joints. And it was evaluated whether the detection of an association was influenced by OA risk factors. Subgroups of studies with patient characteristics commonly described as risk factors for OA like older age, female gender, and high body mass index (BMI) were compared[23-25].

These risk factors were defined as a mean age of ≥65 years, inclusion of ≥75% females, and mean BMI of ≥30 kg/m^2. Also frequency of associations were compared between subgroups of studies in which respectively ≥80% and <80% of included patients had radiographic OA (K&L ≥II). And frequencies of associations were depicted for studies with patients with disease duration of ≥5 years and for population based studies (not directly comparable).

Finally the reported associations were described more specifically for the 'high quality' studies that fulfilled all quality criteria.

Results

Selection of the literature and fulfillment of quality criteria

The literature search resulted in the selection of 45 relevant publications[2,12,14,26-67]. Publications that described the same study population were[26-28/ 12,29/ 37,38/ 40,41/ 48,49], and these were summarized as one study each, resulting in 39 studies.

Table 2 lists the reported patient characteristics: age, percentage females, BMI, and disease severity. More specific study and methodological characteristics are also listed: the number of patients, inclusion criteria, the studied joint, the radiographic protocol, and the radiographic and clinical outcomes that were compared for an association.

Although the patients were commonly female and of older age, study populations were diverse concerning disease severity. The inclusion criteria were commonly based on clinical symptoms, however population based studies and studies in patients with established structural damage were also performed. Also the radiographic protocol and the outcome measures for evaluation of radiographic and clinical OA were substantially different between the selected studies.

Further table 2 lists for all publications the fulfillment of the seven criteria for study characteristics and methodological quality, marked as + or -. The publications were ordered; first according to the number of quality criteria fulfilled, subsequently for specific criteria not fulfilled (-), and finally by the year of publication. The associations between radiographic and clinical OA as reported in the studies, scored as present (+), inconsistent (+/-), or absent (-), are listed in the last column. Only four studies fulfilled all quality criteria and almost half of the studies did not fulfill three or four criteria.

Table 2 Patient, study and methodological characteristics of 39 studies (45 publications)

Number in references list	Author	Year	Age in years (mean±SD)	% females	Body Mass Index in kg/m^2 (mean±SD)	Severity (radiographic/disease duration)	Joint	Number of patients	Inclusion	Protocol	Outcomes radiographic OA	Outcomes clinical OA	1) Description of study population	2) Size of population: n >100	3) Study design	4) Statistics and data presentation	5) OA definition for inclusion	6) Radiographic protocol	7) Outcomes; radiographic and clinical	Association radiographic and clinical OA
26	Duncan[a]	2009	65.5 ±8.7	54	29.6±5.2	68% radOA	K	432	Pain 12mn	MTP-sky 45° sup-lat 45° sup	K&L/ Osteo 0-3	WOMAC 5 pt	+	+	+	+	+	+	+	
27	Duncan[b]	2008	65.5 ±8.7	54	29.6±5.2	68% radOA	K	777	Pain 12mn	MTP-sky 45° sup-lat 45° sup	K&L/ Osteo 0-3	WOMAC 5 pt	+	+	+	+	+	+	+	+
28	Duncan[c]	2007	65.2 ±8.6	55	29.6±5.2	68% radOA	K	745	Pain 12mn	MTP-sky 45° sup-lat 45° sup	K&L/ Osteo 0-3	WOMAC 5 pt	+	+	+	+	+	+	+	
12	Salaffi[a]	2005	67.8 ±9.1	55.1	27.5 (25.4-30.1)	90%ROA 5.1yr	H (K)	107	ACR clin	AP pelvis wb (AP semi wb)	K&L	WOMAC VAS/ SF-36 bp 5 pt	+	+	+	+	+	+	+	+/-
29	Salaffi[b]	2005	64.6	62.9	27.2 (25.1-28.2)	90%ROA 5.1yr	H (K)	105	ACR clin	AP pelvis wb (AP semi wb)	K&L	SF-36 bp 5 pt	+	+	+	+	+	+	+	
30	Günther	1998	60.5 ±9.7	52.4	27.3±4.1	5yr	H (K)	420	K&L	AP pelvis sup (AP ext wb-lat sup)	K&L	Pain VAS/ WOMAC	+	+	+	+	+	+	+	+/-
31	Dougados	1996	63 ±7	60	26±4	97%ROA 4.6yr	H	508	ACR clin	AP pelvis wb	JSW measure	Pain VAS/ Lequesne	+	+	+	+	+	+	+	-
32	Amaro	2007	68.4 ±9.4	39	NR	98%ROA	H	41	K&L	AP pelvis sup	K&L/ JSW measure	Lequesne	+	-	+	+	+	+	+	+
33	Rosemann	2006	56.6 ±12.4	44	26.4±3.5	9.9±5.8yr	H/ K	220	ACR	NR	K&L	WOMAC 5 pt	+	+	+	+	-	+	+	-
34	Szebenyi	2006	65.5 ±9.8	67	30.2±6.5	9.4±7.0yr	K	167	ACR clin&rad	AP ext stand-Lat 30° sup	K&L/ JSW, Osteo,Sc 0-1	Pain VAS/ WOMAC	+	+	+	+	+	-	+	+/-
35	Barker	2004	69.5 ±8.1	53.7	28.4/29.1/3 0.6	ROA, yr: 10.5±12	K	123	ACR	AP ext wb	K&L	WOMAC 5 pt	+	+	+	+	+	-	+	-
36	Bierma-Zeinstra	2002	66 ±9.6	73	NR	47.8% ROA	H	220	Pain	AP pelvis	JSW measure	Pain	+	+	+	+	+	-	-	+/-
37	Peat	2007	≥50	54	NR	70%ROA	K	819	Pain	MTP-Sky 45° sup-Lat 45° sup	K&L/ Osteo/ JSW/ Sc	Stand clin int Phys exam	+	+	+	+	+	-	-	+/-
38	Wood	2007	65.2 ±8.6	55	29.7±5.2	69%ROA	K	697	Pain	MTP-Sky 45° sup-Lat 45° sup	K&L/ Osteo/ JSW/ Sc	Stand clin int Phys exam	+	+	+	+	+	-	-	
39	Bruyere	2002	65.6 ±7.7	76	27.4±2.7	ACR OA	K	212	ACR clin & rad	AP ext wb	Mean, Min JSW measure	WOMAC	+	+	-	+	+	-	+	+/-
40	Birrell[a]	2001	63	68	26(24-30)	<12 mn pain	H	195	Pain	AP pelvis	Croft mod K&L/ Min JSW	Phys exam	+	+	+	-	+	+	-	
41	Birrell[b]	2000	63 ±11	68	NR	<12 mn pain	H	195	Pain	AP pelvis	Croft mod K&L/ Min JSW	Pain 0-10-function/ SF-36	+	+	+	-	+	+	-	+/-
42	Neogi	2009	62±8/ 68	60/ 62	31±6/29±5	Pain, risk OA	K	696/33 6	Pain-OA risk/pop	PA fix flex wb-Lat wb/PA wb	K&L/ Osteo 0-3/ JSN 0-3	Pain: freq-consistent-severe	+	+	+	+	-	-	+	+

43	Toivanen	2007	60.0 ±10.0	68.5	NR	24.6% ROA	K	130	>45yr	AP ext wb-Lat	K&L/ Ahlback/ Piperno	Pain VAS/ WOMAC VAS	+	+	+	+	-	-	+	+/-
44	Zhai	2006	62.7 ±7.1	56	28.7±5.3	Pop sample	K	500	50-79yr	AP semi stand (AP pelvis)	JSN,Osteo,Sc 0-3	WOMAC/ Pain 10 pt	+	+	+	+	-	-	+	-
45	Jordan	1997	63.0 ±10.5	66.9	28.8±6.1	37.8% ROA	K	1192	≥45yr	AP wb	K&L	HAQ disability 0-3	+	+	+	+	-	-	+	+/-
46	Birrell[c]	2005	>45	48	NR	Pop sample	H	1071	Pain	AP pelvis	Croft mod K&L/ Min JSW	Pain 0-1 and location	+	+	+	+	-	+	-	+/-
47	Baker	2004	67.9 ±6.1	59.6	26.0±4.0	35%ROA	K	2472	≥60	MTP(AP ext)-Sky 30° wb	K&L	Strength	+	+	+	+	-	+	-	+/-
48	Jacobsen[a]	2004	61	68	NR	Pop sample	H	3807	Pop sample	AP pelvis stand-Lat lumbar	Min JSW	Pain 4 questions 0-1	+	+	+	+	-	+	-	+/-
49	Jacobsen[b]	2004	65/ 62.5	63	NR	Pop sample	H	4151	Pop sample	AP pelvis stand-Lat lumbar	Min JSW	Pain 3 questions 0-1	+	+	+	+	-	+	-	
50	Reijman	2004	66.0 ±6.9	58	26.3±3.6	Pop sample	H	3585	>55yr	AP pelvis wb	K&L/ Croft/ Min JSW	Pain 0-1/ Stiffness/ Disability	+	+	+	+	-	+	-	+/-
51	Muraki	2009	74.0 ±6.4	64	23.0±3.7	82.5% ROA	K	2282	Pop sample	AP wb	K&L/ JSN lat med	Pain during last month	+	+	+	+	+	-	-	+/-
52	Petersson	1997	49(35 -54)	NR	26	15%ROA	K	185	Pain >3mn	PA ext wb	K&L/ Ahlback	Clinical findings	+	+	+	+	+	-	-	+/-
53	Guccione	1990	72.7 ±6.0	58	NR	63% no radOA	K	1416	Pop sample	NR	K&L	Phys disability	+	+	+	+	+	-	-	+/-
54	Salaffi[c]	1991	63.5 ±7.3	100	NR	7.6±2.9yr	K	61	ACR	AP stand/Lat stand	10 item Total Knee Score 0-4	AIMS/ McGill Pain Quest	+	-	+	-	+	-	+	-
55	Link	2003	63.7 ±11.5	60	NR	80%ROA	K	50	OA	AP-Lat-Sunrise	K&L	WOMAC 5 pt	+	-	+	+	-	-	+	-
56	Fox	1996	72.6 ±5.3	61	NR	RadOA	H	95	OA	AP hip	Croft, Lane/ JSN, Osteo 0-3	Pain/ Daily activities	+	-	+	+	-	-	+	-
57	Ersoz	2003	62.2 ±9.2	75	33.5±5.3	ACR OA	K	20	ACR	AP, Lat & Tangential stand	K&L	Range of motion	+	-	+	+	+	-	-	+/-
58	Lane	2004	71.8 ±5.2	100	NR	74.1% radOA	H	745	≥65yr	AP pelvis sup	Croft JSN 0-4/ Osteo 0-3/ Sc0-1	Pain/ Phys exam/ Disability	+	+	-	+	-	+	-	+
14	Lanyon	1998	62	65	NR	Pop sample	K	452	Pain	AP ext wb-Sky	Osteo 0-3/ JSW measure	Pain 2 questions 0-1	+	+	+	-	+	-	-	+/-
59	Chang	2007	69.4	97.2	26.8	TKR patients	K	107	TKR	AP stand-PA 45° stand-Lat 30°-PF	Ahlback	Pain/ Functional impairment	+	+	+	+	-	-	-	+/-
60	Lachance	2001	45.5/4 4.6	100	32.5±8/ 29.7±8	23%/ 9%ROA	K	323/ 506	Pop sample	AP wb	K&L	Pain 0-1/ Clinical 0-1	+	+	+	+	-	-	-	+/-
61	Odding	1998	69	60	NR	Pop sample	H/ K	2895	>55yr	wb	K&L	Locomotor disable/ HAQ-Pain 0-1	+	+	+	+	-	-	-	+/-
2	Dieppe	1997	62.6 ±11.7	68	26.9±4.7	RadOA	K	415	RadOA	AP stand-Lat 30° sup	Osteo 0-2/ Sc, JSN 0-1	Pain 4 pt/ Function	+	+	+	+	-	-	-	+/-
62	Cicuttini	1996	56.4 ±6.8	100	NR	OA twin study	K	250	Middle-aged	AP ext wb-Lat 30°-sky	Burnett Osteo,JSW 0-3	Pain questions	+	+	+	+	-	-	-	+/-
63	Donnelly	1996	45-65	100	NR	Pop sample	K	950	45-65yr	AP ext wb	Tibial spike/K&L/ Osteo,JSN 0-3	Pain 0-1	+	+	+	+	-	-	-	+/-
64	McAlindon	1993	70.2	56	NR	48% radOA	K	159	Pain	AP wb-Lat 30°	K&L/ JSN,Osteo 0-2/ Sc 0-1	Pain 0-1/ HAQ/ Strength	+	+	+	+	-	-	-	+/-
65	Hart	1991	54	100	NR	Pop sample	K	541	45-65yr	AP wb	K&L	Symptoms/ Clinical signs	+	+	+	+	-	-	-	+/-
66	Claessens	1990	>45	53.9	NR	Pop sample	K	2865	Pop sample	AP wb	K&L	Phys exam	-	+	+	+	-	-	-	+/-
67	Ledingham	1995	71	63	27.8 (17.5-42.4)	RadOA	K	188	Pain	AP stand-Lat 30° (AP pelvis sup)	K&L	Phys exam	+	+	+	-	-	-	-	+/-

Table 2 continued Quality criteria: += fulfilled, -= not fulfilled, number of fulfilled quality criteria separated by horizontal lines, association: += present, +/-= inconsistent -= absent; SD: standard deviation, OA: osteoarthritis, n: number of patients; NR: not reported; radOA: radiographic OA, ROA: radiographic OA defined as Kellgren & Lawrence grade ≥II, yr: years, ACR: American College of Rheumatology, mn: months, Pop: population sample, TKR: total knee replacement; H: hip, K: knee, (K): 6) not fulfilled for knee; clin: clinical, K&L: Kellgren & Lawrence, rad: radiographic; MTP: posteroanterior (PA) semiflexed view, sky: skyline, sup: supine, lat: lateral, AP: anteroposterior, wb: weight bearing, semi: semiflexed, ext: extension, stand: standing, fix flex: fixed flexion; Osteo: osteophytes, JSW: joint space width, Sc: sclerosis, Min: minimum, JSN: joint space narrowing, mod: modified, med: medial; WOMAC: Western Ontario and McMaster Universities OA index, pt: point, VAS: visual analogue scale, SF-36: Short Form Health Survey, bp: bodily pain, stand clin int: standardized clinical interview, phys exam: physical examination, freq: frequency, HAQ: Health Assessment Questionnaire, AIMS: arthritis impact measurement scale

Associations and the importance of quality criteria

An association between radiographic and clinical OA features was present (+) in only 10%, inconsistent (+/-) in 72%, and absent (-) in 18% of 39 studies.

Figure 2 shows, for the studies fulfilling 7 (all), 6, 5, and ≤4 quality criteria (4, 6, 13, and 16 studies respectively), the frequency of an association, an inconsistent association, and no association. An association was most often (25%) found in the four studies fulfilling all criteria. When the number of fulfilled criteria decreased, the frequency of an association diminished (6% for ≤4 criteria) and the frequency of inconsistent associations increased (75% for ≤4 criteria) (Chi-square test: p=0.67).

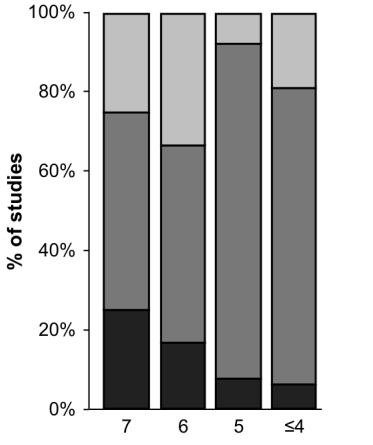

Figure 2 Frequency of an association (dark gray: +), an inconsistent association (middle gray: +/-), and no association (light gray: -) in 4, 6, 13, and 16 studies fulfilling 7 (all), 6, 5, and ≤4 quality criteria respectively

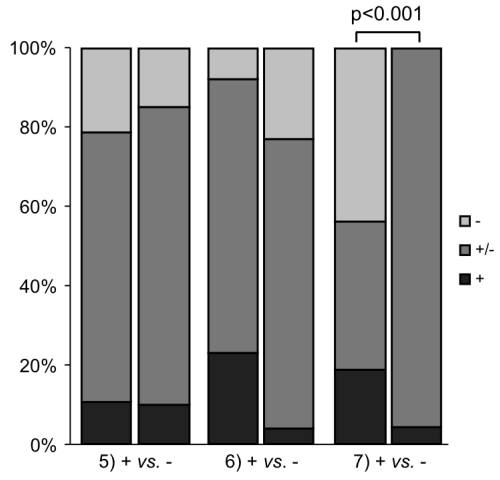

Figure 3 Frequency of an association (dark gray: +), an inconsistent association (middle gray: +/-), and no association (light gray: -) if methodological criteria are fulfilled (+) *vs.* not fulfilled (-). Criterion '5) inclusion': 18 *vs.* 21 studies; '6) protocol': 13 *vs.* 26 studies; and '7) outcomes': 16 *vs.* 23 studies. p-value depicts significance of Chi-square test

Specifically the methodological criteria consisting of '5) inclusion', '6) protocol', and '7) outcomes' were commonly not fulfilled (table 2). Figure 3 shows the frequency of a present, inconsistent, and absent association for the studies either fulfilling (+) or not fulfilling (-) the methodological criteria (regardless of fulfilling other criteria). An association was present in 11%, inconsistent in 67%, and absent in 22% of the 18 studies fulfilling criterion '5) inclusion' which was similar to the 21 studies not fulfilling this criterion (Chi-square test: p=0.78). Considering criterion '6) protocol' an association was more frequently present in the 13 studies fulfilling than in the 26 studies not fulfilling this criterion (p=0.12). The frequency of an association, an inconsistent association, and no association was significantly (p<0.001) different between the 16 studies that fulfilled criterion '7) outcomes' (19% +, 38% +/-, 44% -) and the 23 studies that did not fulfill this criterion (4% +, 96% +/-).
The study criteria 1) to 4) were not fulfilled in only one, five, two, and four studies respectively. For criterion '2) population size' however the associations in the 34 studies fulfilling the criterion (9% +, 79% +/-, and 12% -) were significantly different than in the five studies not fulfilling the criterion (20% +, 20% +/-, and 60% -) (p=0.016). In one of the three studies that did not fulfill criterion '3) study design' an association was present and in the other two studies the association was inconsistent. In the four studies not fulfilling criterion '4) statistics" no present associations were reported (0% +, 75% +/-, and 25% -).

Associations and the influence of patient characteristics

The frequency of an association, an inconsistent association, and no association was not significantly different between the 11 hip studies (18% +, 64% +/-, and 18% -) and the 26 knee studies (8% +, 77% +/-, and 15% -). The association was either inconsistent or absent in the two studies evaluating both hip and knee OA.
In figure 4 frequencies of associations (+, +/-, -) are depicted concerning different risk factors for OA. These frequencies are determined for a part of the studies, since detailed patient characteristics were not always thoroughly reported. The frequency of an association in studies fulfilling a specific risk factor was compared to the association in the studies not fulfilling this risk factor. For all risk factors, no statistically significant difference was found. However, the risk factor older age was reported in 16 studies (table 2) and in older study patients (≥65 years) compared to younger patients (<65 years) associations were more commonly present (19% compared to 0%) and less commonly absent (13% compared to 29%). The risk factors high percentage females and high BMI were only present in nine and two studies respectively. For example, 21 studies reported mean BMI; only in two studies patients were classified as obese (mean BMI ≥30 kg/m^2), in 16 studies patients had overweight (mean BMI 25-30 kg/m^2), and in only one study patients had normal BMI (in two studies several populations were studied who had either overweight or were obese).

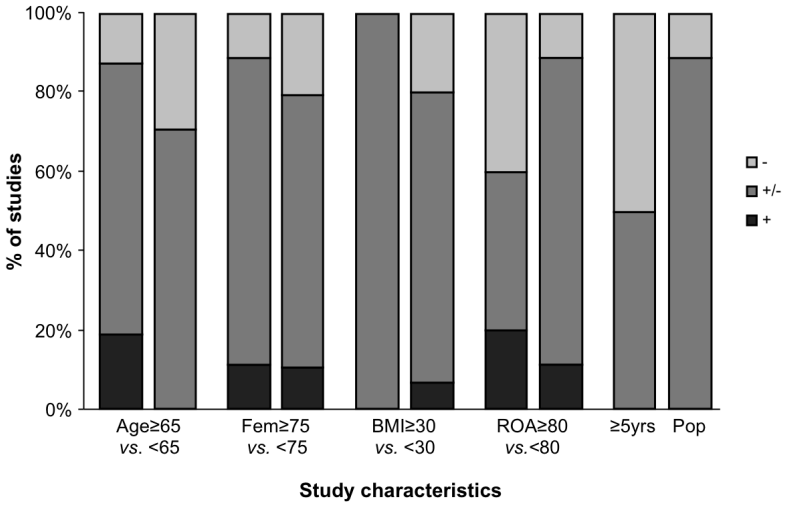

Figure 4 Frequency of an association (dark gray: +), an inconsistent association (middle gray: +/-), and no association (light gray: -) for subgroups of study characteristics. Age in years: 16 *vs.* 17 studies; Fem (=% females): 9 *vs.* 29 studies; BMI in kg/m^2: 2 *vs.* 15 studies; OA severity: ROA (=% radiographic OA, defined as KLG≥II): 5 *vs.* 9 studies, 6 studies with disease duration ≥5 yrs (=years) and 9 studies in Pop (=population sample)

Associations in studies fulfilling all criteria

Seven publications fulfilled all criteria for study characteristics and methodological quality. Three publications by Duncan[a-c] et al. described knee OA in The Clinical Assessment Study of the Knee [(CAS)K][26-28] and two publications by Salaffi[a,b] et al. described cross-sectional evaluation of the same study[12,29] resulting in four 'high quality' studies. Study and patient characteristics are listed in table 2 (top panel) and more specific study characteristics and the association between radiographic and clinical OA are listed in table 3.

The study by Duncan[a-c] et al. described an association (+) between K&L or osteophytes as radiographic measures of OA and mean WOMAC scores as clinical measures. The WOMAC subscale scores for pain, stiffness, and function were significantly different between normal knees, mild radiographic OA, and moderate/severe radiographic OA.[26-28]

The other 'high quality' studies only indirectly described the association between radiographic and clinical OA. The primary aim was to evaluate either the health impact of OA[12,29], the generalized character of radiographic OA[30], or radiographic progression[31].

Table 3 Study characteristics and associations of seven high quality publications

Ref	Study	Joint	n	Inclusion	Definition OA	Radiographic out	Clinical out	Results radiographic *vs.* clinical OA	Ass
26	Duncan[a] 2009	Knee: PF comp	432	Pain within last 12 months	K&L ≥II sky or lateral osteophytes	**K&L (0-IV) Osteophyte (0-3)**	**WOMAC pain, stiffness, function subscales (5-pt Likert; (0-20), (0-8), (0-68), and dichotomized; normal/mild *vs.* moderate/severe/extreme)**	Association radiographic severity & WOMAC Increasing WOMAC score for radiographic subgroups no, mild, moderate/severe: (linear) ANOVA pain F=4.7 p=0.01, stiffness F=4.5 p=0.012, function F=6.1 p=0.002	+
27	Duncan[b] 2008	Knee: TF PF comp	777	Pain within last 12 months	K&L≥II PA and/or sky and/or osteophytes	**K&L (0-IV) Osteophyte (0-3)**	**WOMAC: pain, function subscales (5-pt Likert; (0-20), (0-68))**	Association radiographic severity, compartmental distribution & WOMAC pain, function Increasing WOMAC pain, function for radiographic subgroups no, mild, and moderate/severe; (linear) ANOVA for whole knee; TF; PF compartment pain F=13.8; F=9.0; F=12.7 all p<0.0001, function F=14.9 p<0.0001; F=7.1 p=0.001; F=15.9 p<0.0001	+
28	Duncan[c] 2007	Knee	745	Pain within last 12 months	K&L≥II PA and/or sky and/or osteophytes	**K&L (0-IV) Osteophyte (0-3)**	**WOMAC: pain, stiffness, function subscales (5-pt Likert; (0-20), (0-8), (0-68))** Graded chronic pain (I-IV) Pain, stiffness, disability Questionnaires	Association radiographic OA & WOMAC: OR highest *vs.* lowest of five categories: pain 0-2 *vs.* 11-20 OR=3.7(2.0-6.7); stiffness 0-1 *vs.* 4-10 OR=2.0(2.0-4.6), function 0-5 vs. 35-68 OR=2.8(1.6-5.0)	+
12	Salaffi[a] 2005	Hip (Knee)	107	ACR	ACR	**K&L (0-IV)**	**WOMAC: pain, stiffness, function subscales (VAS; (0-500), (0-200), (0-1700))** **SF-36 (bodily pain; 5-pt)**	Association K&L & WOMAC function; ANOVA (Kruskal-Wallis) H=13.61 p<0.005 No association KL & WOMAC pain, stiffness, SF-36 bodily pain; ANOVA NS	+/-
29	Salaffi[b] 2005	Hip (Knee)	105	ACR	ACR	**K&L (0-IV)**	**SF-36 (bodily pain; 5-pt)**	Association KL & SF-36 bodily pain; ANOVA (Kruskal-Wallis) p<0.0001	+
30	Günther 1998	Hip (Knee)	420	K&L III-IV	K&L≥II	**K&L (0-IV) Lane (JSN, osteophytes; 0-3,** sclerosis, cysts, deformity; 0-1)	**Pain (VAS; 0-100) WOMAC: function subscale** Lequesne, Danielson score Funktionsfragebogen Hannover	Inconsistent association K&L & pain VAS In bilateral radiographic OA 49% pain vs. in unilateral OA 25% pain	+/-
31	Dougados 1996	Hip	508	ACR Pain daily during 1 month in last 3	JSN >0.6mm	**JSW (measure) K&L (0-IV)** Femoral head migration (pattern) Bilaterality (ACR) Polyarticular involve (Landsbury index)	**Pain (VAS; 0-100) Lequesne index** functional disability (0-24) Mobility impairment (measured as intercondylar distance)	No association JSW & clinical OA Regression; R^2=0.4% (p=0.44) pain, functional disability	-

Ref: number in reference list, n: number of patients, OA: osteoarthritis, out; outcome, *vs.*: versus, Ass: association **bold**: outcome measures according to criterion '7) outcomes', association between these outcomes is reported, PF: patellofemoral, K&L: Kellgren & Lawrence grade, WOMAC: Western Ontario & McMaster Universities index, pt: point, TF: tibiofemoral, OR: odds ratio, NS: not significant, JSN: joint space narrowing, ACR: American College of Rheumatology, SF-36: Short Form (36) Health Survey, JSW: joint space width, involve: involvement, R^2: explained variance

Salaffi[a] et al. described a significant association (p<0.005) between K&L and WOMAC function of the hip. However, no association was found for the other WOMAC subscales[12]. Günther et al.[30] described pain in 49% of cases with bilateral radiographic hip OA (defined as K&L≥II) compared to 25% of cases with unilateral radiographic hip OA. In these three publications similar associations were reported for the knee joint, although criterion '6) protocol' was not fulfilled for the knee radiographs. In the Dougados[31] study multiple linear regression analyses performed at baseline showed that all the clinical measures (pain, functional disability) explained only 0.4% of the variation in radiographic hip OA (p=0.4 for model). In summary, in the above three studies on hip OA, the associations were inconsistent (+/-) or absent (-).

Discussion

Only a limited number of studies evaluated an association between radiographic and clinical outcome measures of OA, despite the importance of studying this association for the further understanding of disease onset and progression. The importance of developing quality criteria for study characteristics and methodology is emphasized by the diminishing frequency of present associations with the decreasing number of fulfilled criteria. The fulfillment of all quality criteria as defined in this review resulted in the definition of only four 'high quality' studies out of the 39 selected studies. A significant association was scored for the 'high quality' knee study and inconsistent associations were scored for the studies evaluating hip OA.

The frequency of an association in the present review might have been underestimated due to the definition of present and inconsistent associations. An association was only scored present when all comparisons between radiographic and clinical OA outcomes were statistically significant and studies could have been misclassified as inconsistent based on this definition. For example, when multiple outcomes were compared and one comparison was not significant, the association was defined as inconsistent[12,47]. Also the association was scored as inconsistent when the association was not significant for all evaluated joint compartments[34,57]. On the other hand, an overestimation of associations might also have occurred. It is known that positive results are more easily accepted for publication (publication bias) and might be emphasized in publication (reporting bias). Also confounding might influence the association between radiographic and clinical OA and the association with quality of studies. Correction for confounding was reported by adjusted odds ratios (ORs) or analysis of variance (e.g. age, gender, BMI) in 17 of the selected studies. Associations were commonly inconsistent and less frequently absent (12% +, 82% +/-, and 6% -).

Further, the detection of an association might be hindered by variation between patients since the association between radiographic OA and pain is stronger within individuals than between individuals[42]. Although this might imply that an association can be more easily

reported in longitudinal studies, the detection of an association for individual patients might be hampered by variation due to measurement error or poor reproducibility of radiographic positioning during follow-up (despite protocol)[68]. The assumption of an equal value of a cross-sectional and longitudinal design, as defined by quality criterion '3) study design' in the present review, might therefore be unjust. In most studies (34) only a cross-sectional association was evaluated, which was comparable to all studies. Of the six longitudinal studies, associations were inconsistent in five[2,39,67,69-71] and only present in one study[58]. Also, the manifestation of clinical and radiographic disease might not be clearly linear, but have a more intermittent character that might obscure an association. An approach could be to use both radiographic and clinical features of multiple joints to correct for characteristics of an individual to enable cross-sectional evaluation. However, to identify which clinical and radiographic outcomes are important in onset and progression of disease, longitudinal evaluation remains important.

Even though the frequencies of associations were similar for subgroups of studies fulfilling or not fulfilling criterion '5) inclusion', this criterion has major implications for clinical practice. In evaluation of disease several OA definitions were used for patient inclusion. For example, a wide variation exists in the definition of OA related pain[13]. Also the ACR defined separate classification criteria for knee OA either based on clinical features only or on both radiographic and clinical features[72,73]. This might result in the identification of different populations and thus different OA types and possibly hinders consistent diagnosis and prognosis of disease. This might explain the inconsistent associations in the present review since in some types of OA a clear association might exist while in other types this association is less evident or even absent. Strict inclusion criteria might imply the selection of a homogeneous study population in which no associations can be detected due to little variation in the outcomes however. In the 'high quality' hip study by Dougados et al.[31] for example most patients suffered from mild OA only; 3% had K&L I, 70% K&L II, 27% K&L III, and <1% K&L IV and no association was scored. This is in contrast with the more heterogeneous knee study population[26-28], with 32% of patients with no radiographic OA, 28% with mild OA and 40% with moderate/severe OA, in which stronger associations were found. Further, with increasing severity, radiographic damage might also occur in multiple joints. This can enhance the detection of an association between the more general clinical measures (like VAS and WOMAC) and radiographic measures, which are not joint-specific.

The lack of association in the 'high quality' hip studies might also be related to radiographic protocol. For the knee joint, criterion '6) protocol' proved of importance since associations were found with multiple radiographic knee views,[26-28] whereas no or inconsistent associations were commonly found in studies with only one radiographic knee view or with insufficient protocol (figure 3). For the hip the additional value of the faux profile radiograph of the hip joint[16,17] was not evaluated since this protocol is hardly applied in clinical practice. Therefore this additional view was not included in the quality criteria, although this might explain the lack of present associations in the 'high quality' hip studies.

The fulfillment of criterion '7) outcomes' required standardized outcomes on a scale of at least four categories. Theoretically, measurement of OA outcome measures on a continuous scale improves precision and sensitivity to change. Measurement error or small variations between patients in positioning for radiography however might hamper the precision of the measurement and thereby the association with clinical OA measures[74]. For example, no association was found in the 'high quality' hip study that used JSW as radiographic outcome[31]. Studies comparing the same outcomes for radiographic and clinical OA were sparse and associations were not consistently present when a certain outcome measure, like K&L or WOMAC pain, was evaluated. Therefore, although the frequency of an association was different in studies that fulfilled the criterion '7) outcomes', the role of specific outcomes was not identified as key players in onset or progression of the OA process.

The small number of studies reporting 'normal' values for OA risk factors hindered the evaluation of the influence of patient characteristics on the detection of an association. An obvious role for increasing age, female gender, and high BMI was hence not found.

In conclusion, only a limited number of studies evaluated the association between radiographic and clinical outcome measures of OA. The lower frequency of an association in studies of lower quality emphasizes the importance of criteria for methodological quality, and specifically the standardization of outcome measures.

References

1. Hunter DJ, Felson DT. Osteoarthritis. Bmj. 2006 Mar 18;332(7542):639-42.
2. Dieppe PA, Cushnaghan J, Shepstone L. The Bristol 'OA500' study: progression of osteoarthritis (OA) over 3 years and the relationship between clinical and radiographic changes at the knee joint. Osteoarthritis Cartilage. 1997 Mar;5(2):87-97.
3. Hochberg MC, Lawrence RC, Everett DF, et al. Epidemiologic associations of pain in osteoarthritis of the knee: data from the National Health and Nutrition Examination Survey and the National Health and Nutrition Examination-I Epidemiologic Follow-up Survey. Semin Arthritis Rheum. 1989 May;18(4 Suppl 2):4-9.
4. Torres L, Dunlop DD, Peterfy C, et al. The relationship between specific tissue lesions and pain severity in persons with knee osteoarthritis. Osteoarthritis Cartilage. 2006 Oct;14(10):1033-40.
5. Felson DT, Niu J, Guermazi A, et al. Correlation of the development of knee pain with enlarging bone marrow lesions on magnetic resonance imaging. Arthritis Rheum. 2007 Sep;56(9):2986-92.
6. Buckland-Wright JC. Quantitative radiography of osteoarthritis. Ann Rheum Dis. 1994 Apr;53(4):268-75.
7. Nevitt MC. Definition of hip osteoarthritis for epidemiological studies. Ann Rheum Dis. 1996 Sep;55(9):652-5.
8. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. 1957 Dec;16(4):494-502.
9. Buckland-Wright JC, Ward RJ, Peterfy C, et al. Reproducibility of the semiflexed (metatarsophalangeal) radiographic knee position and automated measurements of medial tibiofemoral joint space width in a multicenter clinical trial of knee osteoarthritis. J Rheumatol. 2004 Aug;31(8):1588-97.
10. Conrozier T, Lequesne M, Favret H, et al. Measurement of the radiological hip joint space width. An evaluation of various methods of measurement. Osteoarthritis Cartilage. 2001 Apr;9(3):281-6.
11. Bellamy N, Buchanan WW, Goldsmith CH, et al. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol. 1988 Dec;15(12):1833-40.
12. Salaffi F, Carotti M, Grassi W. Health-related quality of life in patients with hip or knee osteoarthritis: comparison of generic and disease-specific instruments. Clin Rheumatol. 2005 Feb;24(1):29-37.
13. Bedson J, Croft PR. The discordance between clinical and radiographic knee osteoarthritis: a systematic search and summary of the literature. BMC Musculoskelet Disord. 2008;9:116.
14. Lanyon P, O'Reilly S, Jones A, et al. Radiographic assessment of symptomatic knee osteoarthritis in the community: definitions and normal joint space. Ann Rheum Dis. 1998 Oct;57(10):595-601.
15. McAlindon TE, Snow S, Cooper C, et al. Radiographic patterns of osteoarthritis of the knee joint in the community: the importance of the patellofemoral joint. Ann Rheum Dis. 1992 Jul;51(7):844-9.
16. Conrozier T, Bochu M, Gratacos J, et al. Evaluation of the 'Lequesne's false profile' of the hip in patients with hip osteoarthritis. Osteoarthritis Cartilage. 1999 May;7(3):295-300.
17. Lequesne MG, Laredo JD. The faux profil (oblique view) of the hip in the standing position. Contribution to the evaluation of osteoarthritis of the adult hip. Ann Rheum Dis. 1998 Nov;57(11):676-81.
18. Belo JN, Berger MY, Reijman M, et al. Prognostic factors of progression of osteoarthritis of the knee: a systematic review of observational studies. Arthritis Rheum. 2007 Feb 15;57(1):13-26.
19. van Dijk GM, Dekker J, Veenhof C, et al. Course of functional status and pain in osteoarthritis of the hip or knee: a systematic review of the literature. Arthritis Rheum. 2006 Oct 15;55(5):779-85.
20. Mazzuca SA, Brandt KD, Buckwalter KA, et al. Field test of the reproducibility of the semiflexed metatarsophalangeal view in repeated radiographic examinations of subjects with osteoarthritis of the knee. Arthritis Rheum. 2002 Jan;46(1):109-13.
21. Buckland-Wright JC, Wolfe F, Ward RJ, et al. Substantial superiority of semiflexed (MTP) views in knee osteoarthritis: a comparative radiographic study, without fluoroscopy, of standing extended, semiflexed (MTP), and schuss views. J Rheumatol. 1999 Dec;26(12):2664-74.
22. Buckland-Wright C. Which radiographic techniques should we use for research and clinical practice? Best Pract Res Clin Rheumatol. 2006 Feb;20(1):39-55.
23. Cooper C, Snow S, McAlindon TE, et al. Risk factors for the incidence and progression of radiographic knee osteoarthritis. Arthritis Rheum. 2000 May;43(5):995-1000.
24. Felson DT, Zhang Y, Hannan MT, et al. Risk factors for incident radiographic knee osteoarthritis in the elderly: the Framingham Study. Arthritis Rheum. 1997 Apr;40(4):728-33.
25. Maillefert JF, Gueguen A, Monreal M, et al. Sex differences in hip osteoarthritis: results of a longitudinal study in 508 patients. Ann Rheum Dis. 2003 Oct;62(10):931-4.
26. Duncan R, Peat G, Thomas E, et al. Does isolated patellofemoral osteoarthritis matter? Osteoarthritis Cartilage. 2009 Sep;17(9):1151-5.
27. Duncan R, Peat G, Thomas E, et al. How do pain and function vary with compartmental distribution and severity of radiographic knee osteoarthritis? Rheumatology (Oxford). 2008 Nov;47(11):1704-7.

28. Duncan R, Peat G, Thomas E, et al. Symptoms and radiographic osteoarthritis: not as discordant as they are made out to be? Ann Rheum Dis. 2007 Jan;66(1):86-91.
29. Salaffi F, Carotti M, Stancati A, et al. Health-related quality of life in older adults with symptomatic hip and knee osteoarthritis: a comparison with matched healthy controls. Aging Clin Exp Res. 2005 Aug;17(4):255-63.
30. Gunther KP, Sturmer T, Sauerland S, et al. Prevalence of generalised osteoarthritis in patients with advanced hip and knee osteoarthritis: the Ulm Osteoarthritis Study. Ann Rheum Dis. 1998 Dec;57(12):717-23.
31. Dougados M, Gueguen A, Nguyen M, et al. Radiological progression of hip osteoarthritis: definition, risk factors and correlations with clinical status. Ann Rheum Dis. 1996 Jun;55(6):356-62.
32. Amaro A, Amado F, Duarte JA, et al. Gluteus Medius Muscle Atrophy is Related to Contralateral and Ipsilateral Hip Joint Osteoarthritis. Int J Sports Med. 2007 Dec;28(12):1035-9.
33. Rosemann TT, Joos SS, Koerner TT, et al. Comparison of AIMS2-SF, WOMAC, x-ray and a global physician assessment in order to approach quality of life in patients suffering from osteoarthritis. BMC Musculoskelet Disord. 2006 Jan 26;7(1):6.
34. Szebenyi B, Hollander AP, Dieppe P, et al. Associations between pain, function, and radiographic features in osteoarthritis of the knee. Arthritis Rheum. 2006 Jan;54(1):230-5.
35. Barker K, Lamb SE, Toye F, et al. Association between radiographic joint space narrowing, function, pain and muscle power in severe osteoarthritis of the knee. Clin Rehabil. 2004 Nov;18(7):793-800.
36. Bierma-Zeinstra SM, Oster JD, Bernsen RM, et al. Joint space narrowing and relationship with symptoms and signs in adults consulting for hip pain in primary care. J Rheumatol. 2002 Aug;29(8):1713-8.
37. Peat G, Thomas E, Duncan R, et al. Estimating the probability of radiographic osteoarthritis in the older patient with knee pain. Arthritis Rheum. 2007 May 25;57(5):794-802.
38. Wood LR, Peat G, Thomas E, et al. Knee osteoarthritis in community-dwelling older adults: are there characteristic patterns of pain location? Osteoarthritis Cartilage. 2007 Jun;15(6):615-23.
39. Bruyere O, Honore A, Rovati LC, et al. Radiologic features poorly predict clinical outcomes in knee osteoarthritis. Scand J Rheumatol. 2002;31(1):13-6.
40. Birrell F, Croft P, Cooper C, et al. Predicting radiographic hip osteoarthritis from range of movement. Rheumatology (Oxford). 2001 May;40(5):506-12.
41. Birrell F, Croft P, Cooper C, et al. Health impact of pain in the hip region with and without radiographic evidence of osteoarthritis: a study of new attenders to primary care. The PCR Hip Study Group. Ann Rheum Dis. 2000 Nov;59(11):857-63.
42. Neogi T, Felson D, Niu J, et al. Association between radiographic features of knee osteoarthritis and pain: results from two cohort studies. Bmj. 2009;339:b2844.
43. Toivanen AT, Arokoski JP, Manninen PS, et al. Agreement between clinical and radiological methods of diagnosing knee osteoarthritis. Scand J Rheumatol. 2007 Jan-Feb;36(1):58-63.
44. Zhai G, Blizzard L, Srikanth V, et al. Correlates of knee pain in older adults: Tasmanian Older Adult Cohort Study. Arthritis Rheum. 2006 Apr 15;55(2):264-71.
45. Jordan J, Luta G, Renner J, et al. Knee pain and knee osteoarthritis severity in self-reported task specific disability: the Johnston County Osteoarthritis Project. J Rheumatol. 1997 Jul;24(7):1344-9.
46. Birrell F, Lunt M, Macfarlane G, et al. Association between pain in the hip region and radiographic changes of osteoarthritis: results from a population-based study. Rheumatology (Oxford). 2005 Mar;44(3):337-41.
47. Baker KR, Xu L, Zhang Y, et al. Quadriceps weakness and its relationship to tibiofemoral and patellofemoral knee osteoarthritis in Chinese: the Beijing osteoarthritis study. Arthritis Rheum. 2004 Jun;50(6):1815-21.
48. Jacobsen S, Sonne-Holm S, Soballe K, et al. Radiographic case definitions and prevalence of osteoarthrosis of the hip: a survey of 4 151 subjects in the Osteoarthritis Substudy of the Copenhagen City Heart Study. Acta Orthop Scand. 2004 Dec;75(6):713-20.
49. Jacobsen S, Sonne-Holm S, Soballe K, et al. The relationship of hip joint space to self reported hip pain. A survey of 4.151 subjects of the Copenhagen City Heart Study: the Osteoarthritis Substudy. Osteoarthritis Cartilage. 2004 Sep;12(9):692-7.
50. Reijman M, Hazes JM, Pols HA, et al. Validity and reliability of three definitions of hip osteoarthritis: cross sectional and longitudinal approach. Ann Rheum Dis. 2004 Nov;63(11):1427-33.
51. Muraki S, Oka H, Akune T, et al. Prevalence of radiographic knee osteoarthritis and its association with knee pain in the elderly of Japanese population-based cohorts: The ROAD study. Osteoarthritis Cartilage. 2009 Sep;17(9):1137-43.
52. Petersson IF, Boegard T, Saxne T, et al. Radiographic osteoarthritis of the knee classified by the Ahlback and Kellgren & Lawrence systems for the tibiofemoral joint in people aged 35-54 years with chronic knee pain. Ann Rheum Dis. 1997 Aug;56(8):493-6.

53. Guccione AA, Felson DT, Anderson JJ. Defining arthritis and measuring functional status in elders: methodological issues in the study of disease and physical disability. Am J Public Health. 1990 Aug;80(8):945-9.
54. Salaffi F, Cavalieri F, Nolli M, et al. Analysis of disability in knee osteoarthritis. Relationship with age and psychological variables but not with radiographic score. J Rheumatol. 1991 Oct;18(10):1581-6.
55. Link TM, Steinbach LS, Ghosh S, et al. Osteoarthritis: MR imaging findings in different stages of disease and correlation with clinical findings. Radiology. 2003 Feb;226(2):373-81.
56. Fox KM, Hochberg MC, Resnik CS, et al. Severity of radiographic findings in hip osteoarthritis associated with total hip arthroplasty. J Rheumatol. 1996 Apr;23(4):693-7.
57. Ersoz M, Ergun S. Relationship between knee range of motion and Kellgren-Lawrence radiographic scores in knee osteoarthritis. Am J Phys Med Rehabil. 2003 Feb;82(2):110-5.
58. Lane NE, Nevitt MC, Hochberg MC, et al. Progression of radiographic hip osteoarthritis over eight years in a community sample of elderly white women. Arthritis Rheum. 2004 May;50(5):1477-86.
59. Chang CB, Han I, Kim SJ, et al. Association between radiological findings and symptoms at the patellofemoral joint in advanced knee osteoarthritis. J Bone Joint Surg Br. 2007 Oct;89(10):1324-8.
60. Lachance L, Sowers M, Jamadar D, et al. The experience of pain and emergent osteoarthritis of the knee. Osteoarthritis Cartilage. 2001 Aug;9(6):527-32.
61. Odding E, Valkenburg HA, Algra D, et al. Associations of radiological osteoarthritis of the hip and knee with locomotor disability in the Rotterdam Study. Ann Rheum Dis. 1998 Apr;57(4):203-8.
62. Cicuttini FM, Baker J, Hart DJ, et al. Association of pain with radiological changes in different compartments and views of the knee joint. Osteoarthritis Cartilage. 1996 Jun;4(2):143-7.
63. Donnelly S, Hart DJ, Doyle DV, et al. Spiking of the tibial tubercles--a radiological feature of osteoarthritis? Ann Rheum Dis. 1996 Feb;55(2):105-8.
64. McAlindon TE, Cooper C, Kirwan JR, et al. Determinants of disability in osteoarthritis of the knee. Ann Rheum Dis. 1993 Apr;52(4):258-62.
65. Hart DJ, Spector TD, Brown P, et al. Clinical signs of early osteoarthritis: reproducibility and relation to x ray changes in 541 women in the general population. Ann Rheum Dis. 1991 Jul;50(7):467-70.
66. Claessens AA, Schouten JS, van den Ouweland FA, et al. Do clinical findings associate with radiographic osteoarthritis of the knee? Ann Rheum Dis. 1990 Oct;49(10):771-4.
67. Ledingham J, Regan M, Jones A, et al. Factors affecting radiographic progression of knee osteoarthritis. Ann Rheum Dis. 1995 Jan;54(1):53-8.
68. Mazzuca SA, Brandt KD, Buckwalter KA. Detection of radiographic joint space narrowing in subjects with knee osteoarthritis: longitudinal comparison of the metatarsophalangeal and semiflexed anteroposterior views. Arthritis Rheum. 2003 Feb;48(2):385-90.
69. Miller ME, Rejeski WJ, Messier SP, et al. Modifiers of change in physical functioning in older adults with knee pain: the Observational Arthritis Study in Seniors (OASIS). Arthritis Rheum. 2001 Aug;45(4):331-9.
70. Thorstensson CA, Andersson ML, Jonsson H, et al. Natural course of knee osteoarthritis in middle-aged subjects with knee pain: 12-year follow-up using clinical and radiographic criteria. Ann Rheum Dis. 2009 Dec;68(12):1890-3.
71. Thorstensson CA, Petersson IF, Jacobsson LT, et al. Reduced functional performance in the lower extremity predicted radiographic knee osteoarthritis five years later. Ann Rheum Dis. 2004 Apr;63(4):402-7.
72. Altman R, Alarcon G, Appelrouth D, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. Arthritis Rheum. 1991 May;34(5):505-14.
73. Altman R, Asch E, Bloch D, et al. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. Arthritis Rheum. 1986 Aug;29(8):1039-49.
74. Mazzuca SA, Brandt KD, Dieppe PA, et al. Effect of alignment of the medial tibial plateau and x-ray beam on apparent progression of osteoarthritis in the standing anteroposterior knee radiograph. Arthritis Rheum. 2001 Aug;44(8):1786-94.

Chapter 3

Detection of progression of radiographic joint damage in case of very early osteoarthritis: sensitivity to change of quantitative analysis compared to qualitative grading

MB Kinds[1,2]
ACA Marijnissen[1]
SMA Bierma-Zeinstra[3,4]
JWJ Bijlsma[1]
KL Vincken[2]
MA Viergever[2]
FPJG Lafeber[1]
PMJ Welsing[1,5]

[1]Rheumatology & Clinical Immunology, University Medical Center Utrecht
[2]Image Sciences Institute, University Medical Center Utrecht
[3]General Practice, University Medical Center Rotterdam
[4]Orthopaedics, University Medical Center Rotterdam
[5]Julius Center for Health Sciences & Primary Care, University Medical Center Utrecht

Submitted for publication

Abstract

Objective

For more tailored treatment of osteoarthritis it is of value to identify different subpopulations early in the disease. Objective of this study is to evaluate whether the sensitivity to detect progression of radiographic features, which may add to this identification, can be improved by quantitative measurement (using Knee Images Digital Analysis; KIDA), compared to qualitative grading (according to the Altman atlas).

Methods

Of individuals with early signs related to osteoarthritis (Cohort Hip & Cohort Knee; CHECK) symptomatic knees (n=1082) were selected. Standardized baseline and two-year follow-up radiographs were evaluated for joint space narrowing, osteophyte formation, and bone density changes using KIDA and Altman. Sensitivity to change was determined by calculating the standardized response mean (SRM). For all distinct KIDA parameters the smallest detectable difference was calculated to define radiographic changes at the individual level. The percentage of knees that changed was compared between KIDA measurement and Altman grading. Also agreement between both methods was evaluated.

Results

Studying radiographic progression in knees with early signs related to osteoarthritis showed, for all KIDA and Altman parameters, a small SRM and radiographic change in a small percentage of knees. The sensitivity to detect radiographic progression was similar for KIDA measurement and Altman grading. However, agreement between the Altman and KIDA method was limited (kappa ≤0.20).

Conclusion

Although sensitivity to change is limited, similar for KIDA measurement and Altman grading, this may not exclude that measurement of separate features might be useful to distinguish subpopulations of osteoarthritis later in the disease.

Introduction

Osteoarthritis (OA) is a joint disease characterized by pain and disability. Structural changes like articular cartilage damage, osteophyte formation, and subchondral bone changes are assumed to underlie, at least in part, these symptoms[1,2]. OA diagnosis is primarily based on clinical complaints, excluding other underlying pathologies, and is confirmed by radiography[3]. Despite the discrepancy between radiographic and clinical OA[4], radiographic changes are recognized as an important feature in progression of disease. Although new imaging techniques such as magnetic resonance imaging have clear advantages in research settings, radiographs are still the gold standard in clinical practice for demonstrating structural changes[5,6].

More thorough evaluation of radiographs might identify parameters of OA that are important in onset and progression of disease and might enable more consistent disease definition in clinical diagnosis and follow-up. Commonly the severity of radiographic OA of the knee is qualitatively evaluated by Kellgren & Lawrence (K&L) grading[7]. With K&L grading, a combination of structural changes is assumed to occur in a certain order. Evaluation of progression of distinct radiographic parameters might have additional value and might enable the identification of subpopulations (different phenotypes) of OA (e.g. those with primary cartilage damage compared to those with primary bone changes), in a way that is easily applicable in clinical practice. Qualitative grading of distinct radiographic parameters is possible by use of the Altman atlas[8]. A drawback of this method is that (ordinal) grading is rough and it is generally appreciated that radiographic changes take up to one or two years to become evident[9,10]. Specifically, structural changes early in the disease process, when treatment (tailored to specific phenotypes) may have the best chance of success, are difficult to track by qualitative grading methods. Quantitative measurement aims at more precise measurement and higher sensitivity to change. By measuring joint space width, which is already frequently applied[11,12], changes were detected more easily than when qualitative grading was used[13,14]. Whether quantitative measurement allows for the detection of small differences for other distinct parameters, like osteophytes and bone density, has not been studied.

Therefore, the objective of the present study is to evaluate whether quantitative measurements by use of Knee Images Digital Analysis (KIDA)[15] results in a higher sensitivity for radiographic changes (during two-year follow-up) than qualitative grading by the Altman atlas. Participant with early signs related to OA are evaluated for the distinct radiographic parameters joint space narrowing[11], osteophyte formation, and subchondral bone density in the separate knee joint compartments.

Methods

Study design and participants

The Cohort Hip & Cohort Knee (CHECK) is a prospective ten-year follow-up study on OA initiated by the Dutch Arthritis Association. Individuals (n=1002) with pain and/or stiffness of hip and/or knee, age 45-65 years, and without a previous visit or with a first visit no longer than six months ago to the general practitioner for these complaints, were included in 10 participating hospitals in the Netherlands (CHECK details[16]). The medical ethics committees of all participating hospitals approved the study and all participants gave their written informed consent according to the Helsinki declaration.

Radiographic procedures

Knee radiographs of all participants were acquired in each hospital by different technicians according to a predefined protocol. Posteroanterior weight bearing semiflexed views were taken of both knees separately without fluoroscopy according to Buckland-Wright[17,18]. Technicians were trained for implementation of the protocol and an overall coordinator performed regular quality visits on compliance and, if needed, procedures were corrected to the original protocol.

For the present study the baseline and two-year follow-up radiographs of all knees that were indicated as painful at baseline were evaluated, resulting in a total of 1082 knees. The actual number of analyzed knees can be lower for each of the parameters since KIDA measurement and Altman grading could be hampered by radiographic quality, e.g. osteophytes could not always be thoroughly identified and specifically bone density measurement required good contrast (and a clearly visible aluminum step wedge for KIDA).

Knee Images Digital Analysis (KIDA)

Parameters of radiographic OA were quantitatively measured on a continuous scale by KIDA[15]. In short: the joint space width (JSW; in mm) was determined in the lateral and medial compartment separately, by calculating the mean of four predefined locations in each compartment. Osteophyte area (in mm^2) was determined at the lateral and medial femur and lateral and medial tibia separately. Bone density (in mmAl) was determined in the femur and tibia separately, and gray values were normalized by using an aluminum reference wedge. The values of JSW were expressed as negative values. This enabled straightforward evaluation of whether OA progression occurred in the KIDA parameters, since for all parameters an increase in size represented an increase in OA severity. Inter- and intra-observer variation for KIDA measurements were proven to be low [15] and all baseline and two-year follow-up knee radiographs were analyzed by one experienced observer (ML) in random order unaware of the patient characteristics. The intra-observer variation, tested by random reanalysis of 108 radiographs several months later, showed strong correlations between two observations in this study. The Intraclass Correlation Coefficient (ICC) was 0.73 and 0.95 for lateral and medial JSW, 0.83, 0.83, 0.94, and 0.78 for osteophyte area at

the lateral and medial femur and lateral and medial tibia, and 0.99 for bone density in the femur and tibia.

Altman grading

Joint space narrowing, osteophytes, and subchondral sclerosis (increased bone density) were graded qualitatively according to the Altman atlas which was considered the gold standard for this study[8]. Radiographs from baseline and two-year follow-up were graded in pairs with known sequence, by five observers. In short: joint space narrowing for the medial and lateral compartment was graded from 0-3, osteophytes of the medial and lateral compartment of femur and tibia were each graded from 0-3, and subchondral sclerosis (increased bone density) of femur and tibia were graded as absent or present (0-1). Inter-observer variation in a subset of radiographs resulted, as expected for an ordinal scale, in relatively low ICC of 0.30 for lateral joint space narrowing, 0.61 for medial joint space narrowing, 0.24 and 0.45 for osteophytes at the lateral and medial femur, and 0.78 and 0.72 for osteophytes at lateral and medial tibia.

KIDA measurement and Altman grading each provided additional but different parameters. To enable comparison, the parameters that were similar between both methods were evaluated only.

Kellgren & Lawrence (K&L) grading

The severity of OA of the whole knee joint was also evaluated by K&L grading[7] to obtain an external standard. One observer graded the baseline and two-year follow-up knee radiographs in pairs with known sequence.

Statistical analyses

The sensitivity to change was compared between KIDA measurement and Altman grading by calculating the standardized response mean (SRM)[19]. Although SRM was originally not developed as a measure for qualitative data, a similar application has been reported before[13]. The SRM is defined as the mean change from baseline to two-year follow-up divided by the standard deviation (SD) of this change.

For individual knee joints the radiographic change of the distinct KIDA parameters was defined as a change larger than the smallest detectable difference (SDD) to distinguish random error in the measurement from a real change[20]. The SDD is defined by 1.96 times the SD of the difference in repeated measurement. For this purpose more than 300 knees without any joint damage over two years was used. Selection was based on an Altman grade zero at baseline and at two-year follow-up for all of the individual Altman parameters in the distinct joint compartments. Data from the two radiographs of these 300 knee joints were used to assess the SDD for each of the KIDA parameters. If the selected knees were not changing (no real tissue structure change) over time, the difference between the two KIDA measurements should on average be zero (as was the case).

The percentage of symptomatic knees in CHECK that demonstrated an actual structural change on radiographs from baseline to two-year follow-up was calculated according to KIDA measurement (based on a change larger than SDD) and to Altman grading (defined as, at least, one grade change) for the distinct parameters. To evaluate whether the percentages of knees with a progression in OA and with a decrease in OA severity were significantly different between both methods McNemar tests[21] were used.
Further, it was evaluated whether agreement existed between radiographic change on the distinct parameters according to KIDA measurement and Altman grading using cross-tabulations and calculation of kappa values. Agreement was present when knees were classified similarly with both methods either as increase, decrease, or no change in OA severity.

Results

Baseline characteristics

Of the participants (n=692) with one or two symptomatic knees at baseline 80% was female, mean age was 56±5 (SD) years, median [25-75th percentile] body mass index was 26 [24-28], and median pain intensity (0-10 scale) was 3 [2-5]. Note, as intended, that this cohort concerns an early phase of OA since at baseline K&L was 0 in 78%, I in 18%, II in 3%, and III in 0.5% of the 1082 knees.

Sensitivity to change

Standardized Response Mean (SRM)

The SRM was not evidently greater for the separate radiographic parameters of KIDA measurement as compared to the corresponding parameters of Altman grading (table 1).

Table 1 Standardized Response Mean (SRM) for KIDA measurement and Altman grading

		KIDA change T2y-T0			Altman change T2y-T0		
parameter	**n knees**	**mean**	**SD**	**SRM**	**mean**	**SD**	**SRM**
Joint Space Narrowing							
Lateral	1082	-0.26	1.55	0.17	0.03	0.27	0.11
Medial	1081	0.21	0.63	0.34	0.10	0.39	0.25
Osteophytes							
Femur Lateral	1042	0.40	2.86	0.14	0.08	0.34	0.24
Femur Medial	1035	0.21	2.21	0.09	0.04	0.23	0.18
Tibia Lateral	1037	0.39	2.76	0.14	0.11	0.41	0.27
Tibia Medial	1035	0.55	3.10	0.18	0.08	0.40	0.20
Bone Density							
Femur	732	1.25	5.38	0.23	0.01	0.08	0.12
Tibia	733	2.34	6.59	0.36	0.01	0.11	0.09

T2y: two-year follow-up, T0: baseline

Smallest Detectable Difference (SDD) for KIDA parameters

To define radiographic change in individual knees, the SDD was determined for the distinct KIDA parameters. Table 2 depicts for each KIDA parameter, in the selection of unchanged knees, the mean and SD at baseline, the mean difference and the SD of the difference between baseline and follow-up, and the SDD.

Table 2 Smallest Detectable Difference (SDD) for KIDA parameters in selection of unchanged knees (Altman parameters grade 0)

		KIDA difference (T2y-T0)		
parameter	**n knees**	**mean**	**SD**	**SDD**
Joint Space Narrowing (mm)				
Lateral	313	-0.19	1.34	2.63
Medial	313	0.19	0.49	0.96
Osteophyte area (mm^2)				
Femur Lateral	301	0.34	2.49	4.89
Femur Medial	301	0.04	0.66	1.30
Tibia Lateral	301	0.11	1.73	3.39
Tibia Medial	301	0.02	2.13	4.17
Bone Density (mmAl)				
Femur	213	0.48	5.57	10.92
Tibia	213	1.40	6.61	12.96

T0: baseline, T2y: two-year follow-up

The difference between baseline and follow-up was around zero on average (indeed no progression of damage), while SD (and thus SDD) was quite large. For joint space narrowing the SDD was as expected smallest in the medial compartment. For osteophytes the SDD was smallest in the medial femur and for bone density smallest in the femur.

Radiographic change

The radiographic change from baseline to two-year follow-up as measured with KIDA is depicted by plotting values against the baseline KIDA value for the parameters medial joint space narrowing, osteophyte area of the medial tibia, and bone density of the tibia (as representatives) in figure 1A-C. The SDD value is depicted to illustrate the portion of knees that changed or remained unchanged.

Table 3 depicts for all distinct radiographic parameters of KIDA and Altman the percentage of knees with an increase (progression) or a decrease in OA severity during two-year follow-up.

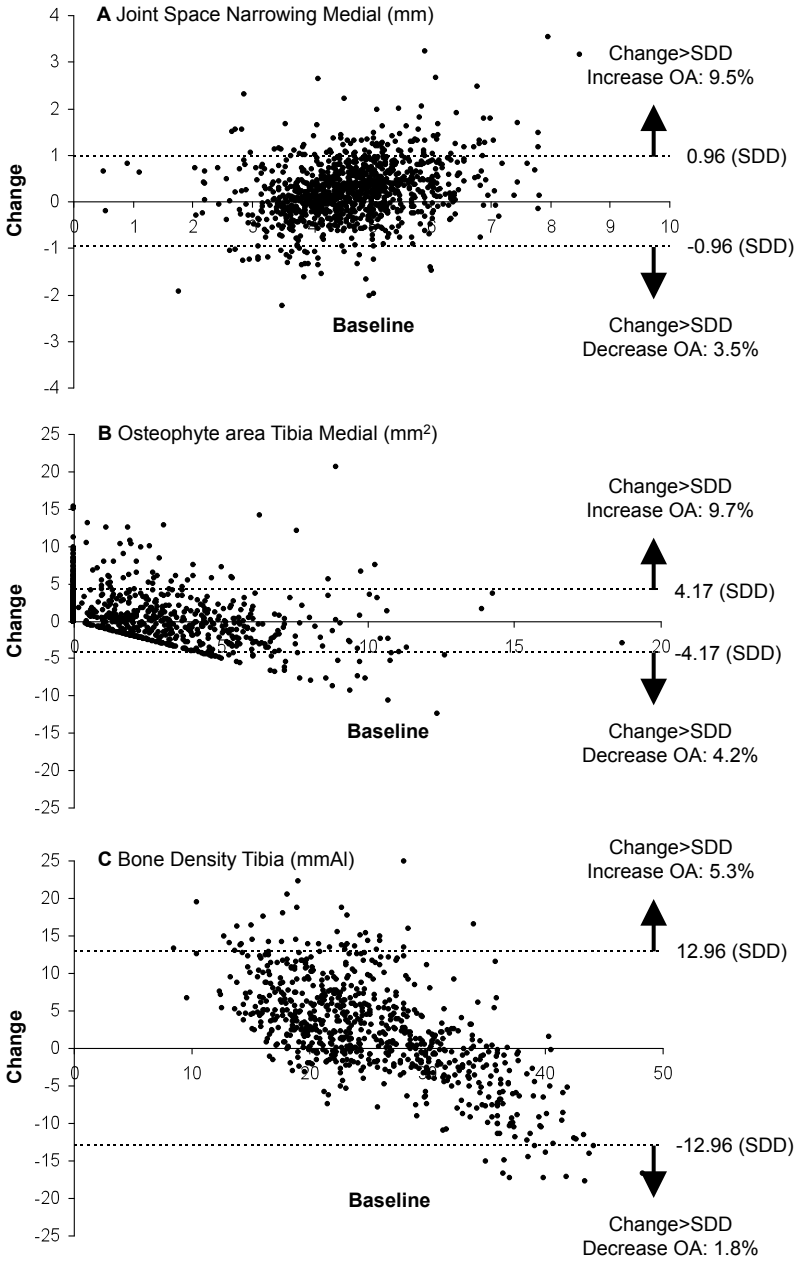

Figure 1 KIDA measurement: change (two-year follow-up – baseline) plotted against baseline value (each dot depicts a knee) for **A** Joint Space Narrowing Medial (mm) of 1081 knees; **B** Osteophyte area Tibia Medial (mm^2) of 1035 knees; **C** Bone Density Tibia (mmAl) of 733 knees

Table 3 Percentage of knees with OA increase or decrease between baseline and two-year follow-up defined according to KIDA measurement and Altman grading

		Increase OA		Decrease OA	
parameter	**n knees**	**KIDA**	**Altman**	**KIDA**	**Altman**
Joint Space Narrowing					
Lateral	1082	**2.6%**	**4.3%**	**7.5%**	**1.1%**
Medial	1081	9.5%	11.5%	3.5%	2.2%
Osteophytes					
Femur Lateral	1042	6.2%	7.4%	**4.7%**	**0.6%**
Femur Medial	1035	**10.1%**	**3.8%**	**5.8%**	**0.4%**
Tibia Lateral	1037	**7.8%**	**12.1%**	**4.1%**	**1.8%**
Tibia Medial	1035	9.7%	10.1%	4.2%	2.7%
Bone Density					
Femur	732	**2.2%**	**0.7%**	**2.9%**	**0.0%**
Tibia	732	**5.3%**	**1.2%**	**1.8%**	**0.0%**

In **bold**: % of knees with increase or decrease significantly different (McNemar $p<0.05$) between KIDA measurement and Altman grading

Clearly, evaluating OA early in the disease process implies that only a small percentage of knees progressed. An increase in OA severity was more frequently identified than a decrease for all distinct Altman parameters, as expected (paired observation with known sequence). But also for KIDA parameters (baseline and two-year follow-up assessment are independent) an increase in OA severity was more frequently found than a decrease (see also as representatives figure 1A-C), except for the parameters bone density of the femur and lateral joint space narrowing. The latter is simply explained by medial compartmental OA that commonly exists, with wedging of the joint due to narrowing of the joint space at the medial side and widening of the joint space at the lateral side (artificially). A decrease in OA severity was found more frequently with KIDA measurement than with Altman grading. For the parameters lateral joint space narrowing, osteophytes at the lateral and medial femur, and bone density of femur and tibia, overall KIDA showed a greater sensitivity to change than Altman (increase and decrease vs. no change, all $p<0.05$).

An increase in OA severity was most commonly found for medial joint space narrowing, in accordance with the higher prevalence of medial (compared to lateral) compartmental knee OA, which was not significantly different for both methods (quantitative with KIDA and qualitative with Altman).

Progression (increase OA) was significantly (**bold** in table 3) more frequently identified with KIDA measurement than with Altman grading for osteophytes at the medial femur (10.1% compared to 3.8% of knees) and for bone density of the femur and tibia. For the parameters lateral joint space narrowing and osteophytes at the lateral tibia a greater percentage of knees was found to increase according to Altman grading than to KIDA measurement. For the parameters medial joint space narrowing and osteophytes at the lateral femur and medial tibia no significant difference was found between the methods.

For K&L, in 9.6% of knees an increase and in 0.1% a decrease in OA severity was found (K&L data from n=1043 symptomatic knees available, paired observation with known sequence; data not shown).

Agreement between radiographic change according to KIDA, Altman, and K&L

The level of agreement between the classification increase, no change, and decrease according to KIDA and Altman was evaluated; cross-tabulations are depicted in table 4.

Table 4 Agreement between KIDA measurement and Altman grading in radiographic change (two-year follow-up – baseline)

parameter	change	(n)	%	(n)	%	(n)	%
Joint Space Narrowing							
Lateral (1082 knees)			***Altman***				
		Increase OA		No change		Decrease OA	
	Increase OA	(5)	***0%***	(23)	2%	(0)	0%
KIDA	No change	(40)	4%	(923)	***86%***	(10)	1%
	Decrease OA	(1)	0%	(78)	7%	(2)	***0%***
Medial (1081 knees)					***Altman***		
		Increase OA		No change		Decrease OA	
	Increase OA	(29)	***3%***	(71)	7%	(3)	0%
KIDA	No change	(91)	9%	(833)	***77%***	(16)	1%
	Decrease OA	(4)	0%	(29)	3%	(5)	***0%***
Osteophytes							
Femur Lateral (1042 knees)					***Altman***		
		Increase OA		No change		Decrease OA	
	Increase OA	(16)	***2%***	(49)	4%	(0)	0%
KIDA	No change	(58)	6%	(864)	***83%***	(6)	1%
	Decrease OA	(3)	0%	(46)	4%	(0)	***0%***
Femur Medial (1035 knees)					***Altman***		
		Increase OA		No change		Decrease OA	
	Increase OA	(15)	***1%***	(90)	9%	(0)	0%
KIDA	No change	(18)	2%	(850)	***82%***	(2)	0%
	Decrease OA	(6)	1%	(52)	5%	(2)	***0%***
Tibia Lateral (1037 knees)					***Altman***		
		Increase OA		No change		Decrease OA	
	Increase OA	(20)	***2%***	(60)	6%	(1)	0%
KIDA	No change	(101)	10%	(796)	***77%***	(17)	2%
	Decrease OA	(4)	0%	(37)	3%	(1)	***0%***
Tibia Medial (1035 knees)					***Altman***		
		Increase OA		No change		Decrease OA	
	Increase OA	(21)	***2%***	(78)	8%	(1)	0%
KIDA	No change	(82)	8%	(785)	***76%***	(25)	2%
	Decrease OA	(2)	0%	(39)	4%	(2)	***0%***
Bone Density							
Femur (732 knees)					***Altman***		
		Increase OA		No change		Decrease OA	
	Increase OA	(0)	***0%***	(16)	2%	(0)	0%
KIDA	No change	(5)	1%	(690)	***94%***	(0)	0%
	Decrease OA	(0)	0%	(21)	3%	(0)	***0%***
Tibia (733 knees)			***Altman***				
		Increase OA		No change		Decrease OA	
	Increase OA	(0)	***0%***	(39)	5%	(0)	0%
KIDA	No change	(9)	1%	(672)	***92%***	(0)	0%
	Decrease OA	(0)	0%	(12)	2%	(0)	***0%***

%: percentage of available knees per parameter, ***bold italic***: agreement between KIDA and Altman

For all parameters agreement existed in a large percentage of knees, since most knees were classified as unchanged according to KIDA measurement and according to Altman grading (76% to 94% for the distinct parameters).
A consequent increase in OA severity according to both methods was only found in a small percentage of all knees however (0% to 3%, dependent on the parameter). In a substantial number of knees disagreement existed between the methods (6% to 22%; sum of the values in boxes) e.g. for medial joint space narrowing (second panel) disagreement existed in 20% of knees: in 7% of knees an increase and in 3% a decrease in OA severity was found with KIDA while Altman grading remained unchanged, and in 9% of knees an increase and in 1% of knees a decrease was found with Altman while KIDA remained unchanged.
Kappa was calculated for the level of agreement between KIDA measurement and Altman atlas on the radiographic change (increase, no change, and decrease in OA severity). For all parameters slight agreement was found between both methods. For joint space narrowing at the lateral compartment kappa is 0.06 (95% confidence interval (95%CI): -0.06–0.18) and at the medial compartment kappa is 0.20 (0.12–0.27). For osteophytes at the femur lateral, femur medial, tibia lateral, and tibia medial, kappa is 0.08 (-0.02–0.19), 0.14 (0.04–0.25), 0.07 (-0.02–0.16), and 0.08 (-0.01–0.17), respectively. For bone density at the femur and tibia, kappa is -0.01 (-0.26–0.24) and -0.02 (-0.23–0.19).
To enable evaluation of radiographic progression, the sensitivity to detect change with KIDA measurement and Altman grading was analyzed in the subgroup of knees classified as changed according to one or both methods (data not directly shown; extraction from table 4). In this subgroup, as expected based on table 4, also only a small percentage of knees progressed according to both methods (0%-12% for the distinct parameters) and thus in the largest percentage of knees disagreement existed. Knees were more commonly defined as either increasing or decreasing in OA severity with KIDA while remaining unchanged with Altman grading (40% to 88%; range for the distinct parameters) than ‘vice versa’ however (11%-48%).
Since only slight agreement was found between KIDA measurement and Altman grading, it was evaluated whether changes in either KIDA or Altman were in more agreement with changes in K&L (as an external standard). Radiographic change of the distinct KIDA and Altman parameters were compared with the change according to K&L (one score for the whole knee).
The level of agreement between KIDA and K&L and between Altman atlas and K&L was similar (and similar to the agreement between KIDA and Altman grading). Agreement existed in 78%-87% (range for the distinct parameters) for no radiographic change, 0%-1% for increase in OA severity, and disagreement existed in 13%-21% of knees when comparing KIDA and K&L. Similarly, when comparing Altman atlas with K&L, the agreement between knees not changing was 79%-91%. The agreement between knees increasing in OA severity was 0%-2%, and disagreement existed in 9%-19% of all knees.

Discussion

The sensitivity to detect progression of radiographic joint damage is similar for KIDA measurement and Altman grading when evaluated in knees with early signs related to osteoarthritis. Only in a small percentage of knees a radiographic change is identified with any of the methods. The sensitivity to change (SRM) is small using KIDA measurement and Altman grading and only a limited level of agreement exists between the two methods.

Importantly, although both KIDA measurement and Altman grading are relatively fast methods for the evaluation of distinct OA parameters and are applicable in clinical trials, the approach to evaluate radiographic change is substantially different. Altman grading was performed on paired radiographs with known sequence. In this evaluation method, changes in radiographic parameters are commonly not graded for decrease in OA severity since an increase in OA severity is anticipated[23]. The effect on lateral and medial joint space narrowing demonstrates this clearly. In contrast, the mathematical approach of KIDA enables precise measurement of radiographic parameters and both increase and decrease in OA severity can be measured since the observer is blinded for sequence. For example, according to KIDA measurement in a greater percentage of knees the joint space at the lateral compartment was found to increase (i.e. decrease OA severity) rather than to decrease. This can be explained by a change in knee joint alignment due to medial joint space narrowing causing widening of the lateral joint space. In contrast to KIDA, an increase in joint space was not graded using the Altman atlas since definition of this method is aimed on evaluating a decrease in joint space, and an increase was not anticipated.

The detection of radiographic change using KIDA measurement might have been hindered by the selection of knees for SDD calculation. Although Altman grades remained zero for all parameters, subtle radiographic changes might have occurred in these selected knees during follow-up. The differences between the baseline and follow-up KIDA measurement were on average around zero but all differences had the direction as expected with progression of disease (minimal increase in OA for all parameters, except for lateral joint space). This might have resulted in an overestimation of the SDD and thus the identification of radiographic change according to KIDA in a lower percentage of knees, which underestimates the sensitivity to change. On the other hand, SDD might have been underestimated in the subgroup of radiographs with Altman grade zero since the (random) error might be greater if more radiographic damage exists.

Importantly, differences between baseline and follow-up radiographs due to the radiographic acquisition[24] and differences in knee joint position[24,25] are likely to introduce variation in the objective measurement of KIDA parameters. This is in contrast with subjective Altman grading in which a certain degree of variation can be taken into account in case of sequential scoring. Already slight differences introduce variation independent of the actual radiographic change in the KIDA measurement, which enlarges the calculated SDD and can result in false positive or negative change scores, which is illustrated in figure 2A-D. The substantial percentage of knees (4.1% to 5.8%) in which a decrease in osteophyte area was

found during follow-up was probably also due to this difficulty with radiographic positioning. On the other hand, it can not be ruled out that decreases in osteophyte area actually occur, specifically early in the disease, as this was never studied. Irrespectively, to have advantage of the highly reproducible quantitative analyses, quality of acquisition appears to be of major importance for clinical application of KIDA in longitudinal radiographic evaluation. For future studies it might be worthwhile to identify pairs of radiographs with reproducible knee positioning during acquisition. However, this might result in only a small portion of radiographic pairs with good knee joint alignment, even in studies with standardized protocols for image acquisition[26].

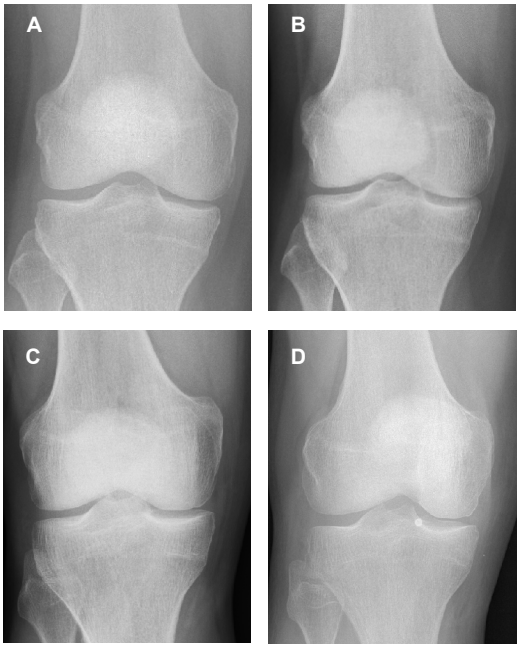

Figure 2 Examples of disagreement between KIDA measurement and Altman grading due to radiographic position. Evaluation of change in medial JSN from baseline (A, C) to two-year follow-up (B, D). **A-B** OA increase according to Altman and no change according to KIDA; **C-D** no change according to Altman and OA increase according to KIDA

Misclassification of radiographic change due to (random) error in the baseline and follow-up measurement, which is a difficulty in all measurement methods, probably explains part of the disagreement between KIDA measurement and Altman grading. For Altman grading reproducibility is relatively low when compared to KIDA. The difficulty with defining radiographic change is supported by similar percentages of radiographic change and lack of agreement with K&L as an external standard (both when compared to Altman grading and KIDA measurement). Also for K&L grading reproducibility is relatively low (ICC: 0.67 to 0.85 for 60 CHECK radiographs).

Importantly, in the present study, radiographic parameters of OA were evaluated in participants with early signs of knee OA. In these knee joints, painful at baseline, only a small percentage of knees changed and the absolute changes were small. E.g. only little increase in OA severity was found for bone density since this parameter is expected not to be profound early in the disease. The value of quantitative measurement of separate parameters with KIDA might prove of additional value when evaluated further in the OA process and in case of more reproducible image acquisition. As such, subpopulations of OA with higher risk for OA progression[27] might be identified slightly earlier in the disease. Irrespective of the method used evaluation of separate parameters might enable the identification of (independent) progression of specific radiographic features.

Acknowledgements

We acknowledge Marja Lafeber (ML) for the KIDA measurements.

References

1. McAlindon T, Dieppe P. Osteoarthritis: definitions and criteria. Ann Rheum Dis. 1989 Jul;48(7):531-2.
2. Peat G, Thomas E, Duncan R, et al. Clinical classification criteria for knee osteoarthritis: performance in the general population and primary care. Ann Rheum Dis. 2006 Oct;65(10):1363-7.
3. Hunter DJ, Felson DT. Osteoarthritis. Bmj. 2006 Mar 18;332(7542):639-42.
4. Dieppe PA, Cushnaghan J, Shepstone L. The Bristol 'OA500' study: progression of osteoarthritis (OA) over 3 years and the relationship between clinical and radiographic changes at the knee joint. Osteoarthritis Cartilage. 1997 Mar;5(2):87-97.
5. Buckland-Wright JC. Quantitative radiography of osteoarthritis. Ann Rheum Dis. 1994 Apr;53(4):268-75.
6. Nevitt MC. Definition of hip osteoarthritis for epidemiological studies. Ann Rheum Dis. 1996 Sep;55(9):652-5.
7. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. 1957 Dec;16(4):494-502.
8. Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. Osteoarthritis Cartilage. 2007;15 Suppl A:A1-56.
9. Boegard TL, Rudling O, Petersson IF, et al. Joint space width of the tibiofemoral and of the patellofemoral joint in chronic knee pain with or without radiographic osteoarthritis: a 2-year follow-up. Osteoarthritis Cartilage. 2003 May;11(5):370-6.
10. Vignon E, Conrozier T, Hellio Le Graverand MP. Advances in radiographic imaging of progression of hip and knee osteoarthritis. J Rheumatol. 2005 Jun;32(6):1143-5.
11. Buckland-Wright JC, Ward RJ, Peterfy C, et al. Reproducibility of the semiflexed (metatarsophalangeal) radiographic knee position and automated measurements of medial tibiofemoral joint space width in a multicenter clinical trial of knee osteoarthritis. J Rheumatol. 2004 Aug;31(8):1588-97.
12. Conrozier T, Lequesne M, Favret H, et al. Measurement of the radiological hip joint space width. An evaluation of various methods of measurement. Osteoarthritis Cartilage. 2001 Apr;9(3):281-6.
13. Gossec L, Jordan JM, Mazzuca SA, et al. Comparative evaluation of three semi-quantitative radiographic grading techniques for knee osteoarthritis in terms of validity and reproducibility in 1759 X-rays: report of the OARSI-OMERACT task force Extended report. Osteoarthritis Cartilage. 2008 Apr 14.
14. Mazzuca SA, Brandt KD, Katz BP, et al. Comparison of quantitative and semiquantitative indicators of joint space narrowing in subjects with knee osteoarthritis. Ann Rheum Dis. 2006 Jan;65(1):64-8.
15. Marijnissen AC, Vincken KL, Vos PA, et al. Knee Images Digital Analysis (KIDA): a novel method to quantify individual radiographic features of knee osteoarthritis in detail. Osteoarthritis Cartilage. 2008 Feb;16(2):234-43.
16. Wesseling J, Dekker J, van den Berg WB, et al. CHECK (Cohort Hip and Cohort Knee): similarities and differences with the Osteoarthritis Initiative. Ann Rheum Dis. 2009 Sep;68(9):1413-9.
17. Buckland-Wright C. Protocols for precise radio-anatomical positioning of the tibiofemoral and patellofemoral compartments of the knee. Osteoarthritis Cartilage. 1995 Sep;3 Suppl A:71-80.
18. Buckland-Wright JC, Wolfe F, Ward RJ, et al. Substantial superiority of semiflexed (MTP) views in knee osteoarthritis: a comparative radiographic study, without fluoroscopy, of standing extended, semiflexed (MTP), and schuss views. J Rheumatol. 1999 Dec;26(12):2664-74.
19. Benichou OD, Hunter DJ, Nelson DR, et al. One-year change in radiographic joint space width in patients with unilateral joint space narrowing: data from the Osteoarthritis Initiative. Arthritis Care Res (Hoboken). 2010 Jul;62(7):924-31.
20. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986 Feb 8;1(8476):307-10.
21. Dwyer AJ. Matchmaking and McNemar in the comparison of diagnostic modalities. Radiology. 1991 Feb;178(2):328-30.
22. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 1988;2nd edition.
23. Botha-Scheepers S, Watt I, Breedveld FC, et al. Reading radiographs in pairs or in chronological order influences radiological progression in osteoarthritis. Rheumatology (Oxford). 2005 Nov;44(11):1452-5.
24. Ravaud P, Giraudeau B, Auleley GR, et al. Assessing smallest detectable change over time in continuous structural outcome measures: application to radiological change in knee osteoarthritis. J Clin Epidemiol. 1999 Dec;52(12):1225-30.
25. Mazzuca SA, Brandt KD, Dieppe PA, et al. Effect of alignment of the medial tibial plateau and x-ray beam on apparent progression of osteoarthritis in the standing anteroposterior knee radiograph. Arthritis Rheum. 2001 Aug;44(8):1786-94.
26. Mazzuca SA, Brandt KD, Buckwalter KA, et al. Field test of the reproducibility of the semiflexed metatarsophalangeal view in repeated radiographic examinations of subjects with osteoarthritis of the knee. Arthritis Rheum. 2002 Jan;46(1):109-13.
27. Bierma-Zeinstra SM, Verhagen AP. Osteoarthritis subpopulations and implications for clinical trial design. Arthritis Res Ther. 2011 Apr 5;13(2):213.

Chapter 4

Radiographic features of knee and hip osteoarthritis represent characteristics of an individual, in addition to severity of osteoarthritis

MB Kinds[1,2]
KL Vincken[2]
EP Vignon[3]
S ten Wolde[4]
JWJ Bijlsma[1]
PMJ Welsing[1,5]
ACA Marijnissen[1]
FPJG Lafeber[1]

[1]Rheumatology & Clinical Immunology, University Medical Center Utrecht
[2]Image Sciences Institute, University Medical Center Utrecht
[3]Rheumatology, Lyon-Sud University Hospital, Pierre-Bénite, France
[4]Rheumatology, Kennemer Gasthuis, Haarlem
[5]Julius Center for Health Sciences & Primary Care, University Medical Center Utrecht

Scandinavian Journal of Rheumatology; in press

Abstract

Objective

To evaluate to what extent radiographic features of knees and hips that are normally related to osteoarthritis, represent characteristics of an individual in addition to osteoarthritis severity.

Methods

A cohort of individuals (n=1002) with very early signs of hip and knee osteoarthritis (CHECK) was studied. Baseline radiographs were evaluated by digital analyses (Holy's and KIDA software), providing separate quantitative measures of radiographic osteoarthritis features. Additionally, conventional Kellgren & Lawrence (K&L) grading was performed. Digital parameters were evaluated for correlations within participants between contralateral (left *vs.* right hip and left *vs.* right knee), ipsilateral (e.g. left hip *vs.* left knee), and diagonal joints (e.g. left hip *vs.* right knee). Analyses were performed separately for participants with K&L grade 0-I and those with evident radiographic osteoarthritis (K&L grade II-III). Regression analyses determined whether demographic characteristics were related to radiographic features.

Results

Correlations between digital parameters and K&L grade were moderate, and within each K&L grade large variation was found. Within participants strong correlations were found for digital parameters between joints in individuals with K&L grade 0-I (R=0.60-0.89), strongest for contralateral comparison, but no statistically significant correlations were found for participants with K&L grade II-III. The demographic characteristics age, gender, height, and weight were to a limited extent, (R^2=0.01-0.20) but statistically significant, related to radiographic characteristics.

Conclusion

Using digital analyses of radiographic osteoarthritis, strong correlations between joints within participants were found. These correlations diminished when osteoarthritis became evident. This has implications for monitoring joint damage in (very) early osteoarthritis with digital analyses.

Introduction

Osteoarthritis is a slowly developing joint disease characterized by pain and disability. Structural changes like articular cartilage damage, osteophyte formation, synovial inflammation, and subchondral bone changes are assumed to originate these symptoms[1,2]. However, there is a discrepancy between radiographic and clinical features of osteoarthritis, hampering definition, diagnosis, and evaluation of progression[3]. Despite this discrepancy and the progressive development of magnetic resonance imaging, which enables the detection of specific structural changes before becoming radiographically evident[4], radiographs are still considered the gold standard for demonstrating structural changes. This is because the image acquisition method is non-invasive, cheap, fast, and generally available[5,6].

Radiographic osteoarthritis is commonly scored by use of a Kellgren & Lawrence (K&L) grade[7]. Drawback of such a grading is that it provides only a qualitative (ordinal) score of a combination of separate structural aspects. In general it takes up to one or two years before progression of a single radiographic grade becomes evident[8,9].

Quantitative measurement of separate radiographic features of osteoarthritis by digital image analysis theoretically enables more precise measurement of disease severity with higher sensitivity to change. Therefore, measurement of joint space width (JSW) is frequently used as an individual quantitative parameter[10,11]. Also quantitative measures of osteophyte area and bone density may allow for the detection of small differences[12].

The availability of these digital tools tempted us to evaluate to what extent radiographic features of knees and hips, which are normally related to osteoarthritis, represent characteristics of an individual in addition to severity of osteoarthritis.

Methods

Study population and radiographic procedures

To enable the detection of characteristics (joint anatomy) of individuals[13,14] in addition to severity of the disease, a cohort with very early signs of osteoarthritis is needed. Moreover, standardized acquisition of radiographs according to a protocol in such a cohort is a prerequisite since variation in quantitative separate radiographic parameters of osteoarthritis is strongly dependent on variation in positioning[15-17] and radiographic procedure[17,18] during acquisition of the images.

CHECK (Cohort Hip & Cohort Knee) provided a large study population of 1002 participants (aged 45-65 years) with very early characteristics of hip and/or knee osteoarthritis. Individuals were included based on pain and/or stiffness of hip and/or knee, and without a visit or with a first visit no longer than six months ago to the general practitioner for these complaints. Individuals with K&L grade IV and individuals with pathological conditions other than osteoarthritis explaining the complaints were excluded. The medical ethics committees

of all participating centers approved the study and all participants gave their written informed consent according to the declaration of Helsinki before entering the study[19]. The baseline assessment of this cohort is used for the present study.
Participants were included in 10 hospitals in the Netherlands. Standardized radiographs according to a predefined protocol were taken in each hospital (in general clinical practice). A central data manager appointed within CHECK performed regular quality checks on the procedure. For both knees the posterior anterior weight bearing semiflexed view of the tibiofemoral joint according to Buckland-Wright[20,21] was used at 55 peak kilo voltage, 5 milliampere second, and 100 cm source image distance. An aluminum standard was added as a reference for bone density during the knee radiographs[12]. For the hip joints a weight bearing anterior posterior view of the pelvis[11,18] was acquired at 80 kilo voltage and with 100 cm source image distance. Radiographs were either obtained digitally (57% of participants) or as film-screen (analogue: 43%). Film-screen radiographs were scanned (Vidar Diagnostic pro plus) at the University Medical Center Utrecht and digitized at a resolution of 300 dots per inch. The CHECK study protocol provided radiographs of hips and knees as standardized as possible for such a large cohort.

Evaluation of radiographs

Radiographs of hips and knees were first scored according to K&L[7] by eight different experienced observers unaware of the source of the radiographs. To evaluate inter-observer reliability all observers scored 60 radiographs that were representative of the spectrum of radiographic severity of osteoarthritis in CHECK. The Intraclass Correlation Coefficient was good, ranging from 0.67 to 0.85 for the four joints.

Holy's software

Digital image analyses of the hips was performed by use of Holy's software-β19/20™ (Lyon-Sud University Hospital, Pierre-Bénite, France)[22,23]. Hip radiographs were analyzed at the department of rheumatology of Lyon-Sud University Hospital in random order by two observers (MK and AM) unaware of the source of the radiograph. Approximately a quarter of hips (557 of the 2004) were analyzed by both observers to determine inter-observer variation, resulting in good ICC's of 0.77 for minimum JSW (joint space width, and 0.80 for mean JSW. For these radiographs the average value of both analyses was used. In short, Holy analysis includes three steps: First, joint space contours are automatically determined by outlining the femoral head and the inferior margin of the acetabulum. Second, the measurement area is demarcated, starting at the lateral side of the joint and moving medially with a standard angle of 65°. In cases where medial boundary of the acetabulum is unclear, a smaller angle and thus a smaller measurement area is manually selected (20% of the radiographs). Finally, contours of the femoral head and acetabulum within the 65° angle are fitted by the program. When the program gives no good fit, manual interference is possible to reach better fit (8% of the radiographs). The technique provides, in a

standardized and (semi-) objective way, quantitative measurement of minimum and mean JSW for each hip.

Knee Images Digital Analysis (KIDA)

For digital analyses of knee radiographs, Knee Images Digital Analysis (KIDA) was used[12]. Inter-observer reproducibility was proven to be high[12] and all knee radiographs were analyzed by one experienced observer (ML) in random order unaware of the source of radiographs. The intra-observer variation, tested by random reanalyzes of 105 radiographs several months later, showed good correlations between two observations (Pearson's R=0.85, R=0.90, R=0.90, and R=0.99 for minimum JSW, mean JSW, osteophyte area, and bone density). Of these radiographs, the initially obtained value was used. KIDA provides, in a (semi-) objective and standardized way, quantitative measurement of multiple radiographic joint characteristics. In short, KIDA consists of six steps: 1) Identification of aluminum bone density standard; 2) Identification of the joint by placing a framework with lines at the lateral and medial margin of the joint, superior margin of the tibia, and inferior margin of the femoral condyles; 3) Defining bone cartilage interface and subchondral area by placing four circles in each compartment (lateral femur, medial femur, lateral tibia, medial tibia); 4) Defining tibial eminence height by placing a circle on both eminences; 5) Defining osteophyte margins by placing a circle at the margin of each compartment; 6) Indication of minimum JSW by marking location with the smallest distance between femur and tibia. For measurement of bone density, the gray scales of the circles in step 3) are related to the bone density standard (for details see[12]).

For the present study, KIDA parameters consisted of minimum JSW and mean JSW completed with total osteophyte area, and bone density. These parameters were selected since digital hip parameters measured were minimum and mean JSW, and K&L grading has osteophytes and sclerosis as features additionally to joint space narrowing[7]. Moreover, narrowing of JSW is the most accepted digital parameter for defining osteoarthritis progression[24,25]. Osteophyte area and bone density are relatively new digital parameters in characterizing joint degeneration in osteoarthritis[12].

Statistical analyses

It is hypothesized that early in the process of disease a large variation between participants within each K&L grade is present since quantitative measures of separate radiographic features represent joint characteristics of an individual in addition to characteristics of osteoarthritis. Therefore, first the average and variation (standard deviation: SD) of the digital parameters was determined per K&L grade and was compared between K&L grades using the independent sample t-tests. Also the Pearson correlation coefficient between K&L grades and the digital parameters was calculated (using for osteophyte area 'log [osteophyte area +1]'; necessary to obtain a more normal distribution).

Based on 'normal anatomy', in participants with no or doubtful radiographic osteoarthritis (K&L grade 0-I) a relation between separate parameters within participants (i.e. between

joints) is hypothesized. To evaluate characteristics within participants, the average difference between the quantitative measurement of digital parameters of the contralateral hips and knees (i.e. difference between left hip *vs.* right hip, and left knee *vs.* right knee), ipsilateral joints (left hip *vs.* left knee, and right knee *vs.* right hip), and diagonal joints (left hip *vs.* right knee, and right hip *vs.* left knee) was determined with it's variations (SD). This difference (and SD) was expected to be smaller in participants with K&L grade 0-I (no radiographic osteoarthritis) compared to participants with K&L grade II-III (radiographic osteoarthritis), which was defined according to K&L grade of the worst joint (e.g. a participant with K&L grade II for the left knee and 0 for the right knee was defined as having radiographic osteoarthritis). Subsequently, Pearson correlations for contralateral joints, and additionally for ipsilateral and diagonal joints were determined within participants[26]. Since contralateral joints are evaluated for the hips on one radiograph and on separate radiographs for the knees, quantitative parameters for the contralateral hips are expected to be more similar than for the knees and more similar than for ipsilateral and diagonal comparisons. It was evaluated whether the association between joints within participants was most evident in participants without radiographic osteoarthritis (K&L grade 0-I) and whether a disturbance of radiographic anatomy (and with that a weaker association) occurred in case of radiographic osteoarthritis (K&L grade II-III).

Furthermore, joint characteristics of an individual might be related to demographic characteristics (e.g. height, age, weight). The relation between digital parameters and the demographic variables age, gender, height, and weight was studied using linear regression analyses. These analyses were performed in the group with K&L grade 0-I to minimize the influence of radiographic joint damage, with the average values of the digital parameters of all four joints as dependent variable. Also separate regression analyses for the knees and hips were performed to see if the effect of demographic variables on digital parameters was different between knees and hips.

Statistical analyses were performed using Statistical Package for the Social Sciences software (version 15.0), $p<0.05$ was considered statistically significant.

Results

Baseline characteristics

Average age of CHECK (Cohort Hip & Cohort Knee) participants at baseline was 56 (±5) years and median pain intensity was 3 (0-10 scale). Based on K&L grade on radiographs participants were indeed evaluated early in the osteoarthritis process: 79% had grade 0, 15% I, 5% II, and 1% III of the hip, and 68%, 25%, 6%, and 1% had K&L grade 0 to III of the knee, respectively. The American College of Rheumatology criteria for clinical osteoarthritis were fulfilled in 24% and 76% of participants with hip and knee symptoms, respectively. Detailed CHECK characteristics were described by Wesseling et al.[19].

Available data

The 4 joints of all 1002 participants were evaluated, however not all data were available and of sufficient quality. Missing data were due to missing or incomplete radiographic series (162 of 4008 radiographs), inadequate quality for determination of K&L grade (26 of 4008 radiographs), and insufficient quality as defined by limitations in executing all steps for digital analysis or poor contrast (717 of 4008 radiographs). With these restrictions all quality radiographs available were used for the different comparisons (n values: figure legends and tables). Demographics and disease characteristics were not statistically significantly different for the participants with quality radiographs when compared to the entire cohort (table 1).

Table 1 Demographic and disease characteristics in (subgroups of) CHECK participants

	All	Hip rad	Knee rad	Hips contra	Knees contra
Demographic					
n participants (rad)	1002	932 (1860)	717 (1243)	928	527
Age in years	56 (5)	56 (5)	56 (5)	56 (5)	56 (5)
Female gender	79%	79%	78%	79%	79%
BMI in kg/m^2	26 [23-28]	25 [23-28]	25 [23-28]	26 [23-28]	25 [23-28]
Pain					
Intensity (0-10)	3 [2-5]	3 [2-5]	3 [2-5]	3 [2-5]	3 [2-5]
Hip only	17%	17%	18%	17%	19%
Knee only	41%	41%	41%	41%	39%
Hip & knee	42%	42%	41%	42%	42%
K&L grade					
Hip					
0	79.4%	79.8%	78.4%	79.8%	77.9%
I	14.7%	14.1%	15.4%	14.1%	14.4%
II	5.3%	5.5%	5.7%	5.5%	7.0%
III	0.6%	0.6%	0.6%	0.5%	0.6%
Knee					
0	68.3%	67.9%	69.3%	68.1%	70.2%
I	24.9%	25.5%	24.4%	25.5%	23.5%
II	6.0%	5.8%	5.6%	5.6%	5.5%
III	0.8%	0.8%	0.7%	0.8%	0.8%

Mean (standard deviation) or median [25-75th percentile] depicted, rad: radiographs, contra: contralateral pairs, BMI: Body Mass Index, K&L: Kellgren & Lawrence

Relation between digital parameters (Holy and KIDA) and K&L grade

All separate osteoarthritis features of hips (Holy; Hospitalier Lyon-Sud) and knees (KIDA; Knee Images Digital Analysis) correlated statistically significantly with K&L (Kellgren & Lawrence) grade. Figure 1 shows the relation between K&L grade and minimum joint space width (JSW) of the hip, minimum JSW of the knee, and knee osteophyte area. Variation in these parameters within each K&L grade is shown. For the hip, minimum JSW showed a moderate correlation (1A), slightly stronger than for mean JSW (R=-0.24, data not shown). Correlations for JSW of the knees were lower than for the hips, both for minimum (1B) and mean JSW (R=-0.19, data not shown). As expected, since osteophytes are a key feature in K&L grading, for the knee the strongest correlation was found for osteophyte area (1C). A small correlation was found between K&L grade and bone density of the knee (R=0.10, data

not shown). The average values of the digital parameters were mostly statistically significantly different between the different K&L grades, but more importantly the variation in the digital measures was large, even in K&L grade 0 and I (figure 1).

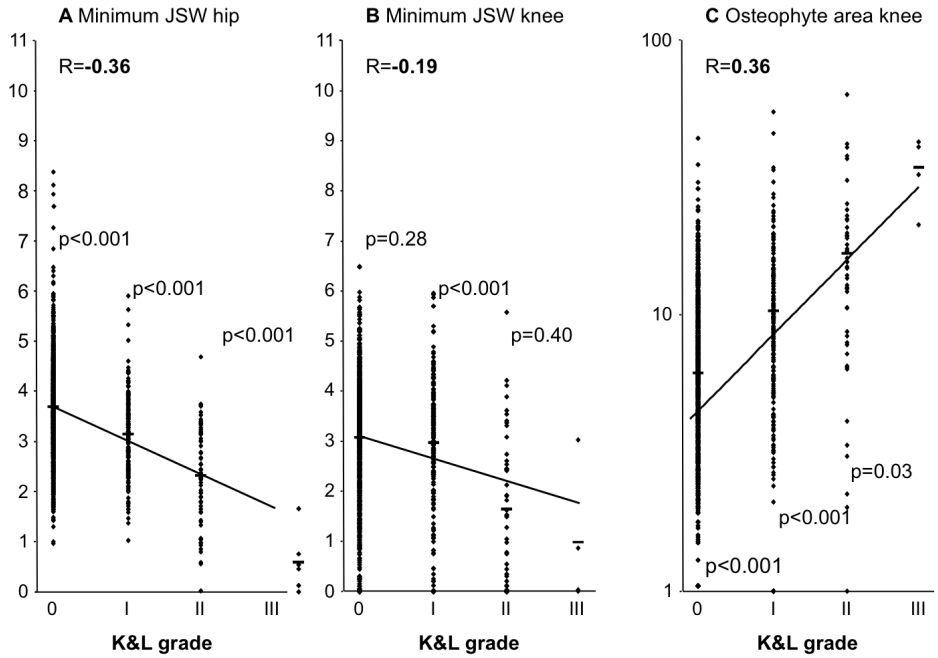

Figure 1 Significant correlation (R in **bold**) between K&L grade and quantitative digital parameters:
A minimum JSW of 1860 hip radiographs; **B** minimum JSW and **C** osteophyte area of 1243 knee radiographs. R= Pearson correlation coefficient, p: significance level between different K&L grades (independent sample t-test), osteophyte area: on log-scale

The average (± standard deviation; SD) minimum JSW (mm) was 3.67 (0.84) for hips with K&L grade 0, 3.13 (0.79) for K&L grade I, 2.30 (0.96) for K&L grade II, and 0.59 (0.59) for K&L grade III. For the knee joints averages (±SD) were 3.06 (1.15), 2.96 (1.30), 1.63 (1.45), 0.99 (1.41) for minimum JSW, and 5.89 (4.81), 10.31 (7.28), 16.54 (11.86), and 34.03 (9.67) for osteophyte area (+1 on log scale), for K&L grade 0, I, II, and III. respectively. A comparable variation was found for mean JSW of hip and knee, and bone density of the knee (data not shown).

Relation between joints within participants

To evaluate characteristics within participants, both the variation and the correlation between joints were evaluated. Table 2 shows the variation within participants between the

contralateral knees and hips, the ipsilateral joints, and the diagonal joints, respectively. The variation for all digital parameters as the SD of the difference within participants was depicted separately for the subgroups with K&L grade 0-I and II-III. The average difference was close to zero, specifically in the contralateral joints and in the subgroup with K&L grade 0-I. The variation in the digital measures within participants was expectedly larger in participants with higher K&L grades. Importantly, the variation for the digital parameters within participants was smaller than the variation between participants (compare SD in table 2 and figure 1). For example, the variation (SD) in minimum JSW of the hip was 0.84 mm for K&L grade 0 and the variation in the difference in minimum JSW between contralateral hips within participants was 0.39 mm (K&L grade 0-I).

Table 2 Difference and variation (SD) between joints within participants in digital parameters

		Minimum JSW (mm)		**Mean JSW** (mm)		**Osteophytes** (Total area-mm^2)		**Bone density** (Mean-mmAl)	
Comparison	**n pairs**	**Diff**	**(SD)**	**Diff**	**(SD)**	**Diff**	**(SD)**	**Diff**	**(SD)**
Contralateral									
Hip									
K&L 0-I	872	0.45	(0.39)	0.37	(0.33)				
K&L II-III	56	1.17	(0.88)	0.89	(0.69)				
Knee									
K&L 0-I	494	0.60	(0.58)	0.42	(0.39)	3.29	(3.83)	1.90	(1.90)
K&L II-III	33	1.60	(1.32)	1.33	(1.43)	10.83	(10.43)	1.75	(2.37)
Ipsilateral									
K&L 0-I	1071	1.13	(1.00)	0.88	(0.71)				
K&L II-III	95	1.74	(1.20)	1.45	(0.98)				
Diagonal									
K&L 0-I	1069	1.14	(1.00)	0.88	(0.72)				
K&L II-III	96	1.65	(1.25)	1.40	(0.99)				

Diff: difference calculated as average of absolute differences, SD: standard deviation of difference, K&L: Kellgren & Lawrence grade

Correlations between joints within participants are provided in figures 2 and 3 for contralateral comparison and in table 3 for ipsilateral and diagonal comparisons. For minimum and mean JSW of the hip joints (figure 2A-B) and knee joints (figure 3A-B), and osteophyte area and bone density of the knees (figure 3C-D) separately for the participants with K&L grade 0-I (black dots) and those with K&L grade II-III (triangles) the relation between left and right joint is depicted and Pearson correlation coefficients are shown.

Correlations of separate digital parameters between joints within participants were statistically significant, and generally much larger than correlations of these separate digital parameters with K&L grade (although they are not directly comparable). Correlations were higher between the contralateral joints than between the ipsilateral and diagonal joints. These correlations were all stronger for the K&L grade 0-I subgroup than for the K&L grade II-III subgroup, except for bone density, which showed strong correlations in both subgroups.

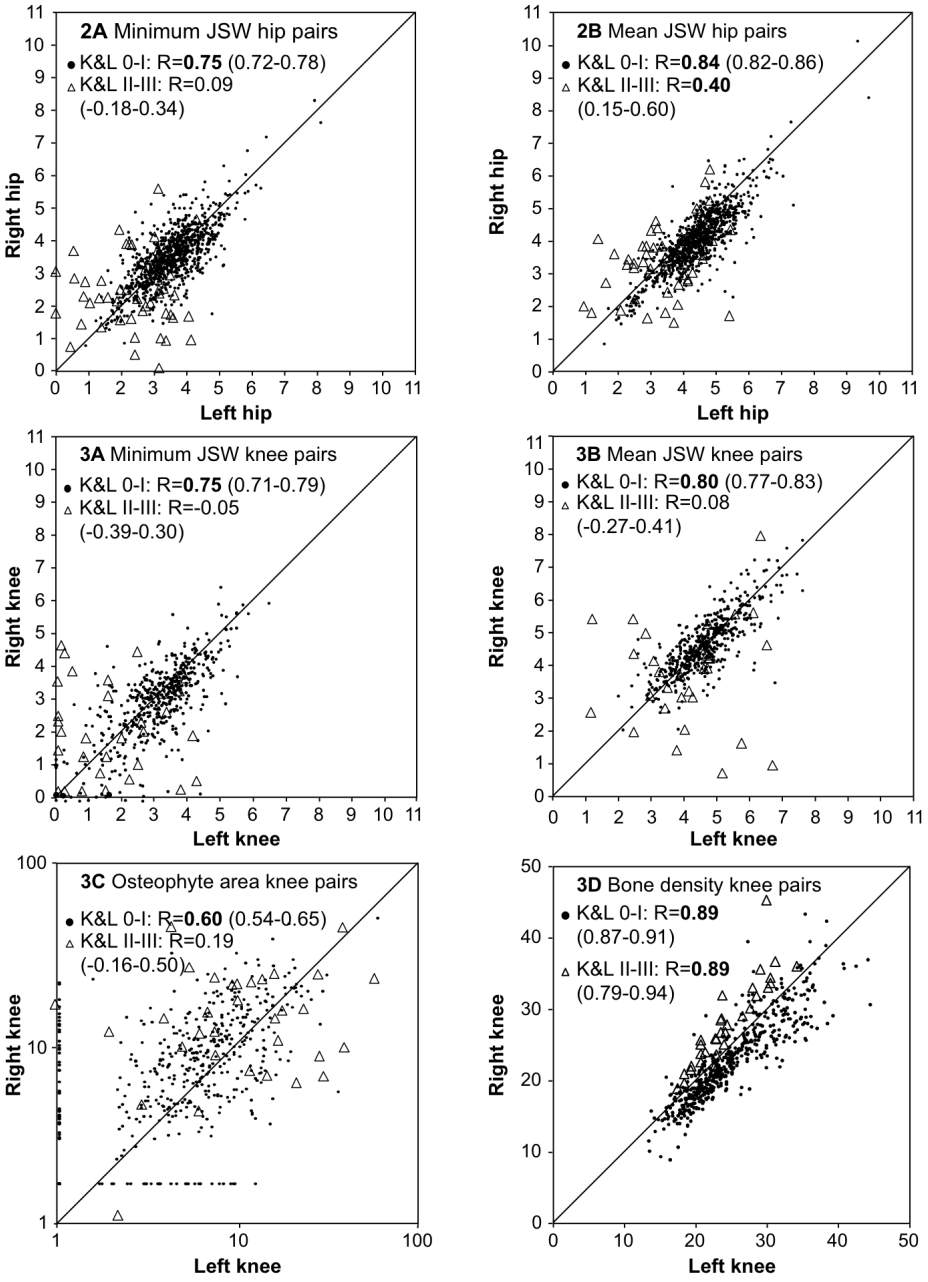

Figure 2 Contralateral comparison of 928 left and right hips within participants: **A** Minimum JSW (mm); **B** Mean JSW (mm). **Figure 3** Contralateral comparison of 527 knees: **A** Minimum JSW (mm); **B** Mean JSW (mm); **C** Osteophyte area (mm^2+1; log scale); **D** Bone density (mmAl). Dots depict participants with K&L grade 0- I, triangles depict participants with K&L II-III (based on most severe of both joints). R= Pearson correlation coefficient (95% confidence interval) per subgroup; R in **bold** statistically significant

Table 3 Correlation between joints within participants

		Minimum JSW (mm)		Mean JSW (mm)	
Comparison	**n pairs**	**R**	**(95%CI)**	**R**	**(95%CI)**
Ipsilateral					
K&L 0-I	1071	**0.09**	**(0.03-0.15)**	**0.29**	**(0.23-0.34)**
K&L II-III	95	**-0.22**	**(-0.40- -0.02)**	0.02	(-0.18-0.22)
Diagonal					
K&L 0-I	1069	**0.08**	**(0.02-0.14)**	**0.29**	**(0.23-0.34)**
K&L II-III	96	-0.18	(-0.37-0.02)	0.06	(-0.14-0.26)

R: Pearson correlation coefficient, 95%CI: 95% confidence interval, K&L: Kellgren & Lawrence grade, values in **bold** are statistically significant

Are age, gender, height, and weight related to radiographic joint characteristics?

For regression analyses quantitative measurements of the digital parameters of all four joints of 494 participants were available. A small to moderate part of the variance in the digital parameters could be explained (R^2 from 0.01 to 0.20) by demographic variables. For the average of all four joints, height (regression coefficient (ß)=0.03, 95% confidence interval (CI)=0.02-0.04) and weight (ß=0.006, CI=0.001-0.012) were significantly related to mean JSW and the explained variance was 0.18. For minimum JSW the model consisted of height only (ß=0.02, CI=0.01-0.03), resulting in R^2=0.07. Separately for hips only, height was associated with mean JSW (ß=0.02, CI=0.01-0.03, R^2=0.04) and minimum JSW (ß=0.01, CI=0.005-0.02, R^2=0.01). For the knees, models consisted of height (ß=0.03, CI=0.02-0.04) and gender (ß=-0.48, CI=-0.69- -0.26) for mean JSW (R^2= 0.20), and height only (ß=0.03, CI=0.01-0.04) for minimum JSW (R^2=0.04). Finally, only weight on its own was related to osteophyte area (ß=0.006, CI=0.004-0.009, R^2=0.08). The combination of weight (ß=0.08, CI=0.04-0.12), gender (ß=-2.35, CI=-3.62- -1.08), and age (ß=-0.14, CI=-0.23- -0.05) were related to bone density with an R^2 of 0.11.

Discussion

This study demonstrates that, within each K&L grade, quantitative parameters of radiographic osteoarthritis show large variation, clearly in participants with no or doubtful osteoarthritis (K&L grade 0 and I). Within participants with K&L grade 0-I, radiographic features are correlated between joints, not only contralateral but also ipsilateral and diagonal between hips and knees mutually. This suggests that radiographic features of knee and hip osteoarthritis represent characteristics of an individual, in addition to severity of osteoarthritis.

The correlation with K&L grade was strongest for osteophyte area for the knee and minimum JSW of the hip, which is in agreement with the definition of osteoarthritis according to K&L. In the knee an increase in K&L grade from 0 to I and from I to II corresponds with increasing

osteophyte severity only, and the involvement of JSW starts at grade III. Since only a very limited percentage of CHECK participants suffer from K&L grade III, specifically osteophytes have an evident contribution. For the hip joint only JSW is measured with digital analysis, resulting in a correlation between K&L grade and minimum JSW that is comparable to that of knee osteophytes. This correlation might be explained by the definition according to K&L, in which increasing osteoarthritis severity in the hip involves not primarily osteophytes but both JSW and osteophytes.

The considerable variation between individuals might largely be explained by a natural variation, independent of osteoarthritis features[27]. This is supported by the lower variation in the digital parameters within participants than between participants, and by the correlations between contralateral, ipsilateral, and diagonal joint characteristics within participants. A role for demographic variables like age, gender, height, and weight on radiographic features is possible[25]. For example, a large variation in the JSW measurement between a tall and a short participant is reasonable, while the joint characteristics within a participant are highly comparable. Indeed, the individual radiographic parameters can, although only for a small portion, be explained by individual demographic characteristics with height being most commonly associated with the different radiographic parameters.

It is reasonable to assume that part of the variation is not due to individual joint characteristics, but due to variations in image acquisition, despite optimal standardization. Position (alignment) of the pelvis and the tibial plateau is subject to variation[18,28] and is thus of influence on data obtained by digital analysis. Thorough standardization of radiographic procedures is important to decrease variation and to increase comparability within and between participants[29]. For large cohorts in clinical trials and regular clinical practice it is not feasible to perform fluoroscopy during acquisition of the images, however.

Nevertheless, the finding that correlations were also significant for ipsilateral and diagonal comparison of hips and knees, and that radiographic characteristics correlate with demographics, supports that the relation between joint characteristics within participants is not solely based on image acquisition.

The weaker correlations and the greater variation in digital parameters in participants with K&L grade II-III, suggest a disruption of ‘normal anatomy’ during development of osteoarthritis in a single joint. This ‘normal anatomy’ should be taken into account when using detailed quantitative (continuous) radiographic parameters. Since differences exist between individuals, the absolute measures of quantitative parameters do not per se indicate whether characteristics represent early osteoarthritis or just joint anatomy of an individual. In cross-sectional studies, early in the osteoarthritis process, the relation between joints within an individual might be helpful in correcting for radiographic joint characteristics. A ratio between the separate radiographic characteristics of the affected joint and the contralateral or ipsilateral ‘control’ joint might be used for early evaluation of development of osteoarthritis in a joint. It might even be more helpful to use the relation of the affected joint with a reference joint that is not commonly affected by osteoarthritis, e.g. the shoulder. It needs further evaluation however whether a relation exists for such a reference joint with

joints that are commonly affected by osteoarthritis within an individual. Very early in the disease process the use of a reference joint might lead to increased sensitivity to change for digital analyses compared to conventional grading methods. However, also this needs proof by longitudinal analyses. When osteoarthritis severity progresses the imbalance between joints will disappear due to the generalized character of disease, as is more commonly described for K&L grade[26,30], making such an approach irrelevant.
In conclusion, when evaluating joint damage specifically very early in the disease process, using detailed quantitative measures of radiographic osteoarthritis, individual characteristics in addition to osteoarthritis characteristics should be taken into account.

Acknowledgements

We acknowledge Marja Lafeber (ML) for the KIDA measurements.

References

1. McAlindon T, Dieppe P. Osteoarthritis: definitions and criteria. Ann Rheum Dis. 1989 Jul;48(7):531-2.
2. Peat G, Thomas E, Duncan R, et al. Clinical classification criteria for knee osteoarthritis: performance in the general population and primary care. Ann Rheum Dis. 2006 Oct;65(10):1363-7.
3. Dieppe PA, Cushnaghan J, Shepstone L. The Bristol 'OA500' study: progression of osteoarthritis (OA) over 3 years and the relationship between clinical and radiographic changes at the knee joint. Osteoarthritis Cartilage. 1997 Mar;5(2):87-97.
4. Javaid MK, Lynch JA, Tolstykh I, et al. Pre-radiographic MRI findings are associated with onset of knee symptoms: the most study. Osteoarthritis Cartilage. 2009 Nov 11.
5. Nevitt MC. Definition of hip osteoarthritis for epidemiological studies. Ann Rheum Dis. 1996 Sep;55(9):652-5.
6. Buckland-Wright JC. Quantitative radiography of osteoarthritis. Ann Rheum Dis. 1994 Apr;53(4):268-75.
7. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. 1957 Dec;16(4):494-502.
8. Boegard TL, Rudling O, Petersson IF, et al. Joint space width of the tibiofemoral and of the patellofemoral joint in chronic knee pain with or without radiographic osteoarthritis: a 2-year follow-up. Osteoarthritis Cartilage. 2003 May;11(5):370-6.
9. Vignon E, Conrozier T, Hellio Le Graverand MP. Advances in radiographic imaging of progression of hip and knee osteoarthritis. J Rheumatol. 2005 Jun;32(6):1143-5.
10. Buckland-Wright JC, Ward RJ, Peterfy C, et al. Reproducibility of the semiflexed (metatarsophalangeal) radiographic knee position and automated measurements of medial tibiofemoral joint space width in a multicenter clinical trial of knee osteoarthritis. J Rheumatol. 2004 Aug;31(8):1588-97.
11. Conrozier T, Lequesne M, Favret H, et al. Measurement of the radiological hip joint space width. An evaluation of various methods of measurement. Osteoarthritis Cartilage. 2001 Apr;9(3):281-6.
12. Marijnissen AC, Vincken KL, Vos PA, et al. Knee Images Digital Analysis (KIDA): a novel method to quantify individual radiographic features of knee osteoarthritis in detail. Osteoarthritis Cartilage. 2008 Feb;16(2):234-43.
13. Daysal GA, Goker B, Gonen E, et al. The relationship between hip joint space width, center edge angle and acetabular depth. Osteoarthritis Cartilage. 2007 Dec;15(12):1446-51.
14. Dougados M, Gueguen A, Nguyen M, et al. Radiological progression of hip osteoarthritis: definition, risk factors and correlations with clinical status. Ann Rheum Dis. 1996 Jun;55(6):356-62.
15. Jacobsen S, Sonne-Holm S, Soballe K, et al. The relationship of hip joint space to self reported hip pain. A survey of 4.151 subjects of the Copenhagen City Heart Study: the Osteoarthritis Substudy. Osteoarthritis Cartilage. 2004 Sep;12(9):692-7.
16. Mazzuca SA, Brandt KD, Dieppe PA, et al. Effect of alignment of the medial tibial plateau and x-ray beam on apparent progression of osteoarthritis in the standing anteroposterior knee radiograph. Arthritis Rheum. 2001 Aug;44(8):1786-94.
17. Ravaud P, Giraudeau B, Auleley GR, et al. Assessing smallest detectable change over time in continuous structural outcome measures: application to radiological change in knee osteoarthritis. J Clin Epidemiol. 1999 Dec;52(12):1225-30.
18. Auleley GR, Duche A, Drape JL, et al. Measurement of joint space width in hip osteoarthritis: influence of joint positioning and radiographic procedure. Rheumatology (Oxford). 2001 Apr;40(4):414-9.
19. Wesseling J, Dekker J, van den Berg WB, et al. CHECK (Cohort Hip and Cohort Knee): similarities and differences with the Osteoarthritis Initiative. Ann Rheum Dis. 2009 Sep;68(9):1413-9.
20. Buckland-Wright C. Protocols for precise radio-anatomical positioning of the tibiofemoral and patellofemoral compartments of the knee. Osteoarthritis Cartilage. 1995 Sep;3 Suppl A:71-80.
21. Buckland-Wright JC, Wolfe F, Ward RJ, et al. Substantial superiority of semiflexed (MTP) views in knee osteoarthritis: a comparative radiographic study, without fluoroscopy, of standing extended, semiflexed (MTP), and schuss views. J Rheumatol. 1999 Dec;26(12):2664-74.
22. Conrozier T, Jousseaume CA, Mathieu P, 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 Sep;37(9):961-8.
23. Conrozier T, Tron AM, Mathieu P, et al. Quantitative assessment of radiographic normal and osteoarthritic hip joint space. Osteoarthritis Cartilage. 1995 Sep;3 Suppl A:81-7.
24. Vignon E, Conrozier T, Piperno M, et al. Radiographic assessment of hip and knee osteoarthritis. Recommendations: recommended guidelines. Osteoarthritis Cartilage. 1999 Jul;7(4):434-6.
25. Brandt KD. Osteophytes in osteoarthritis. Clinical aspects. Osteoarthritis Cartilage. 1999 May;7(3):334-5.
26. Sayre EC, Jordan JM, Cibere J, et al. Quantifying the Association of Radiographic Osteoarthritis in Knee or Hip Joints with Other Knees or Hips: The Johnston County Osteoarthritis Project. J Rheumatol. 2010 Apr 15.

27. Jacobsen S, Sonne-Holm S, Soballe K, et al. Factors influencing hip joint space in asymptomatic subjects. A survey of 4151 subjects of the Copenhagen City Heart Study: the Osteoarthritis Substudy. Osteoarthritis Cartilage. 2004 Sep;12(9):698-703.
28. Botha-Scheepers S, Dougados M, Ravaud P, et al. Effect of medial tibial plateau alignment on serial radiographs on the capacity to predict progression of knee osteoarthritis. Osteoarthritis Cartilage. 2008 Feb;16(2):272-6.
29. Mazzuca SA, Brandt KD, Buckwalter KA, et al. Field test of the reproducibility of the semiflexed metatarsophalangeal view in repeated radiographic examinations of subjects with osteoarthritis of the knee. Arthritis Rheum. 2002 Jan;46(1):109-13.
30. Dahaghin S, Bierma-Zeinstra SM, Reijman M, et al. Does hand osteoarthritis predict future hip or knee osteoarthritis? Arthritis Rheum. 2005 Nov;52(11):3520-7.

Chapter 5

Separate quantitative features of early radiographic knee osteoarthritis: development over five years and relation with symptoms in the CHECK cohort

MB Kinds[1,2]
ACA Marijnissen[1]
JWJ Bijlsma[1]
M Boers[3]
FPJG Lafeber[1]
PMJ Welsing[1,4]

[1]Rheumatology & Clinical Immunology, University Medical Center Utrecht
[2]Image Sciences Institute, University Medical Center Utrecht
[3]Epidemiology & Biostatistics, VU University Medical Center Amsterdam
[4]Julius Center for Health Sciences & Primary Care, University Medical Center Utrecht

Submitted for publication

Abstract

Objective

To evaluate whether quantitative measurement of knee radiographs enables identification of different domains of joint damage and to evaluate the relation of radiographic features with each other and with clinical characteristics over five-year follow-up in very early osteoarthritis (OA).

Methods

Knee radiographs from the Cohort Hip & Cohort Knee (CHECK) were evaluated with Knee Images Digital Analysis (KIDA). Principal component analysis aided the definition of specific radiographic features. These features were evaluated for development, and were related to clinical outcome using baseline, two-year and five-year follow-up radiographs.

Results

Joint space width (JSW: minimum, medial, and lateral), varus angle, osteophyte area, eminence height, and bone density were identified as radiographic features. The features progressed in radiographic severity at different times in follow-up: early (medial JSW, osteophyte area), late (minimum and lateral JSW, eminence height), and both early and late (varus angle, bone density). The separate radiographic features were correlated to each other, and correlations varied between different time points. The JSW features were most strongly related to each other (up to r=0.82), but also e.g. osteophyte area and bone density were correlated (largest r=0.33). The relations with clinical outcome varied over time, but relations were most commonly found for osteophyte area and JSW.

Conclusion

In early OA, separate radiographic features were identified that progressed at different rates between time points. The relations between the radiographic features and with clinical outcome varied over time. This implies that longitudinal evaluation of separate features will give further insight in OA progression.

Introduction

Osteoarthritis (OA) is the most common joint disorder, characterized by pain, functional disability, and limited quality of life. Structural changes affect the whole joint and comprise cartilage damage, osteophyte formation, changes in subchondral density, synovial inflammation, and affection of soft tissue such as ligaments and muscles[1]. Diagnosis of OA, especially in an early phase of the disease, is difficult due to the lack of sensitive and specific diagnostic criteria[2,3]. The inconsistent association that is commonly found between clinical symptoms and radiographic characteristics that represent structural damage hampers definition of such criteria[4-6]. The detection of an association might be improved by measuring separate features of radiographic OA. Also evaluation of separate features over time, from a very early phase of the disease, might provide more insight into the development and progression of structural damage. For example the detailed evaluation of such features might reveal a sequence in the development of specific radiographic aspects during the course of the disease, and the existence of specific relations between aspects of radiographic damage over time.

Knee Images Digital Analysis (KIDA)[7] has been developed to measure separate parameters of radiographic OA damage of the knee in more detail on a continuous scale (quantitative) as opposed to the existing ordinal methods such as Kellgren & Lawrence (K&L) grading[8] and the Altman atlas[9]. The use of KIDA measurement aims at valid and sensitive evaluation of OA progression for application in clinical studies. For evaluation of the onset and progression of OA it is of importance to investigate whether specific radiographic features can be identified using the KIDA measurements. The aim of this study is to evaluate radiographic OA development over time from an early phase of the disease using specific radiographic features based on KIDA measurements, and to evaluate how these features relate to each other and to clinical characteristics of OA.

Methods

Cohort Hip & Cohort Knee (CHECK)

The Cohort Hip & Cohort Knee (CHECK) is a prospective ten-year follow-up study on OA in ten participating hospitals in the Netherlands, initiated by the Dutch Arthritis Association. Individuals (n=1002) with pain and/or stiffness of hip and/or knee, age 45-65 years, and without a previous visit or with a first visit no longer than six months ago to the general practitioner for these complaints were included in the cohort[10]. The course of complaints and radiographic damage were monitored to identify markers for diagnosis and progression of disease. For the present study data from baseline (T0), two-year follow-up (T2y), and five-year follow-up (T5y) were used.

Knee Images Digital Analysis (KIDA)

In CHECK posteroanterior weight bearing semiflexed views are acquired without fluoroscopy according to a standardized protocol (Buckland-Wright)[11,12]. In the present study, the separate radiographs of both knees acquired at T0, T2y, and T5y were evaluated. Separate radiographic parameters were quantitatively measured by use of Knee Images Digital Analysis (KIDA)[7]. Minimum joint space width (JSW in mm) was measured as the smallest distance between femur and tibia. Medial JSW and lateral JSW were defined as the mean of four predefined locations in each compartment. The varus angle (in degrees) between the femur and tibia was determined in the frontal plane using the intersection points that determine the bone and cartilage interface; a positive value represents (more) varus and a negative value represents valgus alignment. Height of the lateral and medial tibial eminence was measured in mm. Osteophyte area (in mm^2) was determined at the lateral and medial femur and lateral and medial tibia. Bone density (in mmAl equivalents) was determined at four predefined locations in the lateral and medial femur and tibia, by normalizing the gray values of the subchondral bone region to those of an aluminum reference step wedge that was present on all radiographs. The KIDA measurements were performed by one experienced observer (ML) in random order blinded to individual and disease characteristics and time point of evaluation. The number of analyzed knees may slightly vary for each of the radiographic parameters since KIDA measurement can be hampered by poor radiographic quality, despite standardized procedures. E.g. osteophyte area can not always be thoroughly outlined and specifically bone density measurement requires good contrast and a clearly visible aluminum reference wedge.

Clinical characteristics

In CHECK clinical characteristics of OA are collected yearly by use of questionnaires and physical examination. In the present study clinical characteristics at the same time points as the radiographic features were used (T0, T2y, and T5y). Clinical OA at the joint level was expressed by the presence of pain during physical examination, for the left and right knee separately. Clinical OA at the participant level was expressed by use of pain and functional limitation scores of the Western Ontario and McMaster Universities Osteoarthritis index (WOMAC).

Statistical analyses

A principal component analysis aided decisions on how to best combine the KIDA parameters into domains that represent specific separate radiographic OA characteristics. In a principal component analysis a large set of parameters, in our case the measurement of fourteen KIDA parameters, is reduced into a smaller set of components by detecting structure in the relationships between parameters and thereby identifying underlying domains (designated as separate radiographic features). In our analysis the number of extracted components was determined using the Kaiser criterion to extract only factors with so called ‘eigenvalues’ greater than 1, and using an Equamax rotation matrix of the factor

loadings[13]. Results of the component analyses performed at T0, T2y, and T5y separately were compared and discussed with experts in the field (FL, AM) to arrive at the definition of separate radiographic features that were considered relevant.

The development of the radiographic features over time (T0, T2y, and T5y) was described separately for left and right knees, as measurements in the same participant are not independent. Also there might be a difference in progression between the left and right knees.

To investigate the relation between the features and to evaluate whether a specific time sequence of development exists, a correlation matrix was calculated. Correlations were determined between radiographic features at T0, T2y, and T5y at the same time point (concurrently) but also with time lags (non-concurrent) between the features (e.g. JSW at T0 and osteophyte area at T2y). If all radiographic features occur concurrently, correlations are expected to be strongest when determined concurrently and weaker when determined non-concurrently. If a specific radiographic feature generally occurs first (as assumed for osteophytes based on K&L grading) non-concurrent correlations of this radiographic feature with the other radiographic features will not be lower (or even higher) than the concurrent correlations.

Cross-sectional univariate and multivariate regression analyses were used to study the relation of the radiographic features with clinical outcome at T0, T2y, and T5y (dependent variable). A manual backward stepwise selection procedure was used to arrive at a final model. Next to radiographic features, demographic and clinical characteristics that were possibly associated with clinical outcome were evaluated. These comprised age, gender and erythrocyte sedimentation rate (ESR in mm/hour) at T0, and body mass index (BMI in kg/m^2) at the evaluated time point (T0, T2y, and T5y). In the T2y and T5y analyses, the presence at baseline of a painful joint, WOMAC pain and functional limitations were also evaluated as potential confounders.

Analyses were performed both at the level of the joint and at the level of the participant. On the joint level the dependent variable was the presence (1) or absence (0) of pain in a specific knee (left or right). The features were evaluated as the measured (calculated) value for each knee, and also as the difference between a knee and the contralateral knee of a participant to take into account that the absolute value of the radiographic features might partly be a characteristic of the individual[14]. This analysis was performed by logistic regression in which the dependency of the left and right knee within individuals was considered, using a random intercept (‘PROC GLIMMIX’ in SAS; Statistical Analysis Software).

On the participant level the dependent variables were the WOMAC pain and functional limitation scores. The radiographic features comprised the sum of the left and right knee (per participant) to represent the total burden of the radiographic features, and also the absolute difference between the left and right knee to take into account the joint characteristics of individuals into account[14]. This analysis was performed by linear regression in a subgroup of participants with involvement of the knees only, by excluding participants with pain or

radiographic involvement (defined K&L grade ≥II) in the hip at T0. This was done to reduce potential confounding by OA in other joints.
Analyses were performed using SPSS (Statistical Package for the Social Sciences) version 15.0 and SAS version 9.1.3, a p-value <0.05 was considered statistically significant.

Results

Most participants in CHECK were female (79%), with mean age (±standard deviation) 56±5 years and ESR 8 [5-13; median and 25-75th percentile] at T0. BMI was 26 [23-28] at T0, 25 [23-28] at T2y and 26 [23-28] at T5y. The WOMAC (0-100 scale with 100 being worst condition) pain score was 25 [10-35] at T0, 20 [10-35] at T2y, and 20 [5-35] at T5y. The WOMAC function score was 21 [10-35], 19 [7-32], and 21 [9-37] at T0, T2y, and T5y respectively. Knee pain was present in 65% of knees at T0, in 56% at T2y, and in 51% at T5y respectively. The K&L grade at T0 was 0 in 81%, I in 16%, II in 3%, and III in 0.4% of knees.

Identification of separate radiographic features

KIDA measurements were available for 1713 of 2004 knees. Principal component analyses identified five similar components at T0 and T5y. At T2y two components were identified, but when five components were forced the definition of components was similar to the other time points. Table 1 depicts the rotated component matrix at T5y.

Table 1 Rotated component matrix at T5y

	Components				
KIDA parameters	**Medial JSW**	**Lateral JSW**	**Osteophyte**	**Eminence**	**Bone density**
Minimum JSW	***0.85***	-0.08	-0.11	-0.24	-0.05
Medial JSW	***0.90***	-0.13	-0.08	0.18	0.06
Lateral JSW	0.09	***0.93***	-0.01	0.26	0.05
Varus angle	-0.47	***0.86***	0.05	0.08	-0.02
Osteophyte femur lateral	-0.05	0.32	***0.60***	0.02	0.10
Osteophyte femur medial	-0.34	-0.01	***0.56***	0.05	0.06
Osteophyte tibia lateral	0.02	-0.08	***0.73***	0.27	0.00
Osteophyte tibia medial	-0.09	-0.01	***0.77***	-0.05	0.03
Eminence height lateral	-0.09	0.18	0.10	***0.86***	0.05
Eminence height medial	0.03	0.12	0.03	***0.86***	0.11
Bone density femur lateral	0.05	0.08	0.04	0.17	***0.93***
Bone density femur medial	-0.16	0.09	0.06	0.11	***0.93***
Bone density tibia lateral	0.14	-0.11	0.08	0.07	***0.93***
Bone density tibia medial	-0.06	0.07	0.07	-0.01	***0.96***

Factor loading per KIDA parameter for five components, ***bold italic***: identified components, JSW: joint space width, osteophyte: osteophyte area, eminence: eminence height

The five extracted components could be labeled as the domains: ‘medial JSW’, ‘lateral JSW’, ‘osteophyte’ (area), ‘eminence’ (height), and ‘bone density’. The factor loading for

minimum JSW was considerable in the 'medial JSW' component and the factor loading for varus angle was considerable in the 'lateral JSW' component. Nevertheless, since the minimum JSW is commonly reported as a separate measure for OA severity[15,16] that we considered important, 'minimum JSW' was defined as a separate radiographic feature. Also 'varus angle' was considered an important separate feature, which also has a different measurement unit (degrees) than the other features. In the components defined as 'osteophyte area', 'eminence', and 'bone density', in which more KIDA parameters were combined, the factor loadings of the most prominent parameters were of comparable magnitude. Therefore, to enable a straightforward interpretation of radiographic OA features in clinical practice, the radiographic features were defined as follows:

- Minimum JSW (mm): value as measured by KIDA;
- Medial JSW (mm): value as measured by KIDA;
- Lateral JSW (mm): value as measured by KIDA;
- Varus angle (degrees): value as measured by KIDA (+: varus and -: valgus);
- Osteophyte area (mm^2): sum of lateral and medial femur, and lateral and medial tibia;
- Eminence height (mm): sum of lateral and medial eminence height;
- Bone density (mmAl): mean of lateral and medial femur, and lateral and medial tibia.

The number of knee radiographs available for determination of the radiographic features differed for the three time points: at T0 922 left knees and 929 right knees, at T2y 920 left and 913 right knees, and at T5y 859 left and 854 right knees were available.

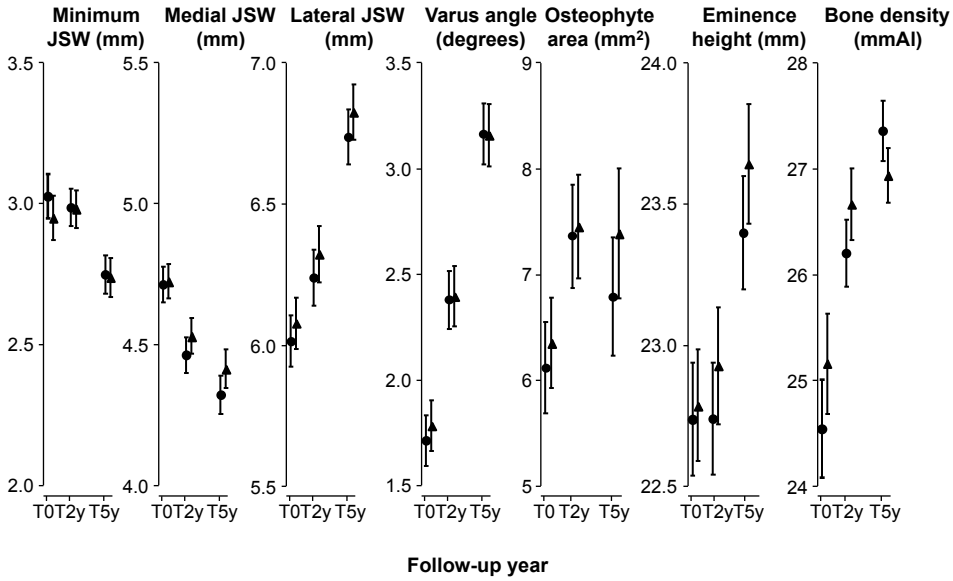

Figure 1 Mean values for separate radiographic features (with 95% confidence interval) at T0, T2y, and T5y for left (dots) and right knees (triangles)

Development over time of separate radiographic features

Overall, during follow-up the knees with early signs of OA in CHECK revealed a statistically significant increase in radiographic severity of OA between T0 and T2y as well as between T2y and T5y on all radiographic features (Figure 1).

Only the changes between T0 and T2y in minimum JSW and the increase in eminence height in the left knees were not statistically significant. The decrease over time in minimum JSW and medial JSW is considered to be in accordance with the increase in lateral JSW and varus angle. Changes in medial JSW and osteophyte area were most evident between T0 and T2y. In contrast, changes in minimum JSW, lateral JSW, and eminence height were most evident between T2y and T5y. Varus angle (and bone density) showed changes during both follow-up periods.

Relation between separate radiographic features

The seven radiographic features were significantly correlated, both at concurrent and non-concurrent time points, and results were similar for the right and left knee. Correlations were more outspoken between the three JSW measurements (minimum, medial, and lateral) and varus angle (r at concurrent time points: between lateral JSW and varus angle around 0.8, and between medial JSW and varus angle -0.4 to -0.6), than between these four features and the three other radiographic features (osteophyte area, eminence height, and bone density). Still, considerable correlations were found between e.g. lateral JSW and the radiographic features osteophyte area, eminence height, and bone density (concurrent r: 0.1 to 0.4). Osteophyte area was most strongly related to bone density (concurrent r: 0.1 to 0.3), and was correlated to eminence height, lateral JSW, and varus angle with comparable strength (concurrent r: around 0.15).

Radiographic joint damage is considered to be represented by a smaller minimum and medial JSW, and larger lateral JSW and varus angle (see figure 1) due to primary medial compartment narrowing and subsequent lateral compartment widening. As such the negative correlations between minimum/medial JSW and lateral JSW/varus angle were expected, as were the positive relations between minimal and medial JSW and between lateral JSW and varus angle. However, the positive correlation between the medial and lateral JSW was not expected but decreased over time (concurrent r: around 0.27 at T0, around 0.15 at T2y, and not significant at T5y).

The correlations were commonly strongest at the same time point (concurrently), but concurrent correlations seemed to increase over time. This was the case for the correlation of varus angle with minimum as well as medial JSW (r: from -0.2 at T0 to -0.5 at T5y for minimum JSW, and from -0.4 at T0 to -0.6 at T5y for medial JSW). Also the correlations between eminence height and minimum as well as lateral JSW increased (r: -0.1 to -0.2 and 0.3 to 0.4 from T0 to T5y, respectively).

Significant but generally weaker correlations were found at non-concurrent time points. Interestingly, the minimum JSW at T5y correlated more strongly with osteophyte area (r: around -0.15) than the minimum JSW at the other time points (r: commonly not significant),

irrespective of the time point of osteophyte area. This might imply that osteophyte development occurred in the early phase in contrast to cartilage loss, which is in accordance with the general development of the radiographic features over time in figure 1. Further, the correlation between eminence height and osteophyte area was strongest when eminence height was measured at T5y and osteophyte area was measured at T0 (r: around 0.25), and this correlation decreased over time (r: around 0.15 when osteophyte area was measured at T5y). This is also considered in accordance with the development of osteophytes before eminence height (figure 1).

Relation of separate radiographic features with clinical outcome

Separate regression analyses were performed at T0, T2y, and T5y to evaluate the association between radiographic features and clinical outcome.

Joint level

Osteophyte area and JSW features were associated with the presence of knee pain at all three time points (table 2). Interestingly, the model at T0, but not at T2y and T5y, mainly included the difference between the contralateral knees for the radiographic features although associations were sometimes counterintuitive. Pain at T0 was a prominent predictor for a painful knee at T2y and T5y together with specific radiographic features.

Table 2 Multivariate regression models at T0, T2y, and T5y with presence of knee pain as dependent variable

Variable	OR	(95%CI)	p
Radiographic feature T0			
Osteophyte area	1.43	(1.09–1.88)	0.01
Minimum JSW (difference)	0.81	(0.71–0.91)	0.001
Medial JSW (difference)	0.70	(0.49–1.00)	0.05
Lateral JSW (difference)	1.38	(0.99–1.94)	0.06
Varus angle (difference)	0.74	(0.57–0.96)	0.02
Eminence height (difference)	1.08	(1.02–1.13)	0.007
Demographic			
ESR T0	1.02	(1.01–1.04)	0.006
Radiographic feature T2y			
Lateral JSW	0.87	(0.76–0.99)	0.04
Varus angle	1.17	(1.06–1.29)	0.002
Osteophyte area	1.42	(1.06–1.91)	0.02
Demographic & clinical			
Pain presence T0	4.53	(3.57–5.75)	<0.0001
Radiographic feature T5y			
Minimum JSW	0.79	(0.70–0.90)	0.0004
Osteophyte area	1.47	(1.08–2.01)	0.02
Bone density	1.04	(1.00–1.07)	0.04
Demographic & clinical			
Female gender T0	1.57	(1.11–2.21)	0.01
BMI T5y	1.04	(1.01–1.07)	0.01
Pain presence T0	2.68	(2.09–3.43)	<0.0001

OR: odds ratio, 95%CI: 95% confidence interval, p: significance level, difference: knee - contralateral knee

Participant level

In the multivariate analyses of the subset of 336 participants with only knee pain and no hip affection at T0, only a few radiographic features were found to be associated with WOMAC pain score (table 3). Again, early in the disease (at T0 and T2y) the models included differences between the contralateral knees in radiographic features (and no sum scores of the radiographic features), with a larger difference representing more pain. At T5y sum scores of the radiographic features were included in the model; smaller minimum JSW (sum) and larger osteophyte area (sum) were associated with more pain.

Table 3 Multivariate regression models at T0, T2y, and T5y with WOMAC pain (0-100 scale as dependent variable

Variable	ß	(95%CI)	p
Radiographic feature T0			
Medial JSW (difference)	4.36	(1.26–7.45)	0.01
Demographic			
BMI T0	0.80	(0.35–1.25)	0.0006
Radiographic feature T2y			
Bone density (difference)	1.14	(0.21–2.07)	0.02
Demographic			
BMI T2y	0.79	(0.31–1.26)	0.001
ESR T0	0.34	(0.05–0.62)	0.02
Radiographic feature T5y			
Minimum JSW (sum)	-1.28	(-2.26– -0.29)	0.01
Osteophyte area (sum)	3.49	(0.73–6.24)	0.01
Demographic			
BMI T5y	0.81	(0.39–1.23)	0.0002

ß: regression coefficient, 95%CI: 95% confidence interval, p: significance level, difference: knee - contralateral knee, sum: left + right knee

Table 4 Multivariate regression models at T0, T2y, and T5y with WOMAC function (0-100 scale) as dependent variable

Variable	ß	(95%CI)	p
Radiographic feature T0			
Osteophyte area (sum)	3.89	(0.52–7.26)	0.02
Bone density (difference)	-1.08	(-2.09– -0.06)	0.04
Demographic			
BMI T0	0.92	(0.32–1.53)	0.003
Radiographic feature T2y			
Osteophyte area (difference)	-8.26	(-15.53– -1.00)	0.03
Demographic			
BMI T2y	1.33	(0.85–1.80)	<0.0001
ESR T0	0.40	(0.12–0.68)	0.01
Radiographic feature T5y			
Minimum JSW (sum)	-1.22	(-1.99– -0.45)	0.002
Demographic & clinical			
BMI T5y	0.53	(0.18–0.87)	0.003
WOMAC function T0	0.66	(0.55-0.77)	<0.0001

ß: regression coefficient, 95%CI: 95% confidence interval, p: significance level, difference: knee - contralateral knee, sum: left + right knee

Also few radiographic features were associated with WOMAC function score (table 4). At T0 the sum of the osteophyte area was associated with WOMAC function, but at T2y, unexpectedly a higher difference in osteophyte area was related to less functional disability. At T5y the sum of the minimum JSW, a feature that was found to progress later in the disease (see figure 1), was the only variable that was associated with outcome.

Discussion

In this cohort of participants with very early symptoms related to osteoarthritis, radiographic features were defined as minimum JSW, medial JSW, lateral JSW, varus angle, osteophyte area, eminence height, and bone density, and were found to be related to each other and partly to clinical outcome. The relation between these radiographic features and the relation of these features with clinical outcome was found to change during progression of disease.

The use of measurement of JSW is already commonly applied to evaluate radiographic OA (progression)[15,17,18], and has been used to evaluate the relation between radiographic and clinical OA characteristics[19,20]. Although the measurement of osteophytes and angle has also been described[21], the application of such measures in a clinical study has not been reported. Moreover, these measures are commonly used in established OA (e.g. K&L grade ≥II)[22,23] and not in such an early OA cohort.

The early progression of medial JSW and osteophyte area (between T0 and T2y) in these participants with early (symptoms of) OA is in agreement with the assumed sequence of these features in K&L grading. However, in the present study varus angle and bone density show promise as early markers of radiographic damage as well.

Interestingly, *widening* of the lateral joint space was identified as a characteristic of progression of early radiographic OA, especially between T2y and T5y in this study. As such, the current focus in clinical trials on narrowing of the medial joint space only, which is the most commonly affected compartment in case of OA, may be unjust and ignores relevant information regarding disease progression provided by the radiograph.

The value of bone density on radiographs may need reappraisal. In the Altman atlas[9], bone density is scored roughly, as either present or absent for different joint compartments. However, our results demonstrate a gradual increase in bone density over time. Grading OA severity on more levels, as is the case for joint space narrowing in the Altman atlas (0-3 scale instead of absent-present), could lead to improvement of grading of radiographic OA severity.

Although the detection of an association between radiographic and clinical characteristics is known to be difficult[4,6], specific radiographic features were found to be associated with clinical outcome in this early OA study. The finding that less radiographic features were associated with WOMAC scores than with presence of knee pain might be due to the evaluation on participant level instead of joint level. When studying clinical OA at joint level, the association is studied more directly with less interference of other (systemic) factors. The

association of the differences between the contralateral knees in radiographic features with clinical outcome mainly at T0 and not at the other time points could suggest that this difference in radiographic features between knees in an individual is a very sensitive measure in the early phase of the disease where only one of the knee joints is commonly affected. However, the regression analyses sometimes showed counterintuitive associations, since also smaller differences in features were found to be associated with worse clinical outcome. Therefore, the interpretation of these differences needs further evaluation.

Osteophyte area was commonly identified as a feature of radiographic osteoarthritis that was associated with clinical outcome, but this was not that consistent for the other radiographic features. The importance of these other features appears limited or might depend on the phase of the disease regarding clinical outcome.

Moreover, several subtypes or phenotypes of OA might exist with specific combinations of (progression in) radiographic features and clinical outcome in subgroups of individuals. In the present study participants might be affected differently, and studying this group as a whole might have hampered the detection of an association of specific radiographic features with clinical outcome[24]. Specific phenotypes might need specific treatment, which is important in clinical trial design. The specific radiographic features as defined and described in this study might aid in the definition of such phenotypes.

In conclusion, using KIDA parameters the following specific and separate radiographic features of knee OA could be identified: minimum, medial and lateral JSW, varus angle, osteophyte area, eminence height, and bone density. All features progressed over time, some mainly in an early phase and some later. Radiographic features were related to each other, and specific features were related to clinical outcome, which possibly depends on the phase of the disease. The identification of these features, their mutual relation, and the relation with clinical outcome might be of use in identifying specific relevant phenotypes of OA.

Acknowledgements

We acknowledge Marja Lafeber (ML) for the KIDA measurements.

References

1. Hunter DJ, Felson DT. Osteoarthritis. Bmj. 2006 Mar 18;332(7542):639-42.
2. Hart DJ, Spector TD. The classification and assessment of osteoarthritis. Baillieres Clin Rheumatol. 1995 May;9(2):407-32.
3. McAlindon T, Dieppe P. Osteoarthritis: definitions and criteria. Ann Rheum Dis. 1989 Jul;48(7):531-2.
4. Bedson J, Croft PR. The discordance between clinical and radiographic knee osteoarthritis: a systematic search and summary of the literature. BMC Musculoskelet Disord. 2008;9:116.
5. Belo JN, Berger MY, Reijman M, et al. Prognostic factors of progression of osteoarthritis of the knee: a systematic review of observational studies. Arthritis Rheum. 2007 Feb 15;57(1):13-26.
6. Kinds MB, Welsing PM, Vignon EP, et al. A systematic review of the association between radiographic and clinical osteoarthritis of hip and knee. Osteoarthritis Cartilage. 2011 Jul;19(7):768-78.
7. Marijnissen AC, Vincken KL, Vos PA, et al. Knee Images Digital Analysis (KIDA): a novel method to quantify individual radiographic features of knee osteoarthritis in detail. Osteoarthritis Cartilage. 2008 Feb;16(2):234-43.
8. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. 1957 Dec;16(4):494-502.
9. Altman RD, M. B., Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. Osteoarthritis Cartilage. 2007;15 Suppl A:A1-56.
10. Wesseling J, Dekker J, van den Berg WB, et al. CHECK (Cohort Hip and Cohort Knee): similarities and differences with the Osteoarthritis Initiative. Ann Rheum Dis. 2009 Sep;68(9):1413-9.
11. Buckland-Wright C. Protocols for precise radio-anatomical positioning of the tibiofemoral and patellofemoral compartments of the knee. Osteoarthritis Cartilage. 1995 Sep;3 Suppl A:71-80.
12. Buckland-Wright JC, Wolfe F, Ward RJ, et al. Substantial superiority of semiflexed (MTP) views in knee osteoarthritis: a comparative radiographic study, without fluoroscopy, of standing extended, semiflexed (MTP), and schuss views. J Rheumatol. 1999 Dec;26(12):2664-74.
13. Hill T, Lewicki P. Statistics: Methods and Applications. 2007.
14. Kinds MB, Vincken KL, Vignon E, et al. Radiographic features of knee and hip osteoarthritis represent characteristics of an individual, in addition to severity of osteoarthritis. Scand J Rheumatol. 2011;*In press*.
15. Boegard TL, Rudling O, Petersson IF, et al. Joint space width of the tibiofemoral and of the patellofemoral joint in chronic knee pain with or without radiographic osteoarthritis: a 2-year follow-up. Osteoarthritis Cartilage. 2003 May;11(5):370-6.
16. Schmidt JE, Amrami KK, Manduca A, et al. Semi-automated digital image analysis of joint space width in knee radiographs. Skeletal Radiol. 2005 Oct;34(10):639-43.
17. Neumann G, Hunter D, Nevitt M, et al. Location specific radiographic joint space width for osteoarthritis progression. Osteoarthritis Cartilage. 2008 Nov 12.
18. Reichmann WM, Maillefert JF, Hunter DJ, et al. Responsiveness to change and reliability of measurement of radiographic joint space width in osteoarthritis of the knee: a systematic review. Osteoarthritis Cartilage. 2011 May;19(5):550-6.
19. Laxafoss E, Jacobsen S, Gosvig KK, et al. Case definitions of knee osteoarthritis in 4,151 unselected subjects: relevance for epidemiological studies : The Copenhagen Osteoarthritis Study. Skeletal Radiol. 2010 Sep;39(9):859-66.
20. Fukui N, Yamane S, Ishida S, et al. Relationship between radiographic changes and symptoms or physical examination findings in subjects with symptomatic medial knee osteoarthritis: a three-year prospective study. BMC Musculoskelet Disord. 2010;11:269.
21. Oka H, Muraki S, Akune T, et al. Fully automatic quantification of knee osteoarthritis severity on plain radiographs. Osteoarthritis Cartilage. 2008 Nov;16(11):1300-6.
22. Eckstein F, Le Graverand MP, Charles HC, et al. Clinical, radiographic, molecular and MRI-based predictors of cartilage loss in knee osteoarthritis. Ann Rheum Dis. 2011 Jul;70(7):1223-30.
23. Mazzuca SA, Brandt KD, Katz BP, et al. Risk factors for progression of tibiofemoral osteoarthritis: an analysis based on fluoroscopically standardised knee radiography. Ann Rheum Dis. 2006 Apr;65(4):515-9.
24. Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: an update with relevance for clinical practice. Lancet. 2011 Jun 18;377(9783):2115-26.

Chapter 6

Evaluation of separate quantitative radiographic features adds to the prediction of incident radiographic osteoarthritis in individuals with recent onset of knee pain: five-year follow-up in the CHECK cohort

MB Kinds[1,2]
ACA Marijnissen[1]
KL Vincken[2]
MA Viergever[2]
KW Drossaers-Bakker[3]
JWJ Bijlsma[1]
SMA Bierma-Zeinstra[4,5]
PMJ Welsing[1,6]
FPJG Lafeber[1]

[1]Rheumatology & Clinical Immunology, University Medical Center Utrecht
[2]Image Sciences Institute, University Medical Center Utrecht
[3]Arthritis Center Twente, University Twente & Medisch Spectrum Twente, Enschede
[4]General Practice, University Medical Center Rotterdam
[5]Orthopaedics, University Medical Center Rotterdam
[6]Julius Center for Health Sciences & Primary Care, University Medical Center Utrecht

Osteoarthritis & Cartilage, conditionally accepted

Abstract

Objective

Detailed radiographic evaluation might enable the identification of osteoarthritis (OA) earlier in the disease. This study evaluated whether and which separate quantitative features on knee radiographs of individuals with recent onset knee pain are associated with incidence of radiographic OA and persistence and/or progression of clinical OA during five-year follow-up.

Methods

From the Cohort Hip & Cohort Knee study participants with knee pain at baseline were evaluated. Radiographic OA development was defined as Kellgren & Lawrence (K&L) grade ≥II at five-year follow-up. Clinical OA was defined as persistent knee pain and as progression of WOMAC pain and function during follow-up. At baseline radiographic damage was determined by quantitative measurement of separate features using Knee Images Digital Analysis, and by K&L grading.

Results

Measuring osteophyte area (odds ratio (OR)=7.0) and minimum joint space width (OR=0.7), in addition to demographic and clinical characteristics, improved the prediction of radiographic OA five years later (area under curve (AUC)=0.74 *vs.* 0.64 without radiographic features). When the predictive score (based on multivariate regression coefficients) was larger than the cut-off for optimal specificity, the chance of incident radiographic OA was 54% instead of the prior probability of 19%. Evaluating separate quantitative features performed slightly better than K&L grading (AUC=0.70). Radiographic characteristics hardly added to prediction of clinical OA.

Conclusion

In individuals with onset knee pain, radiographic characteristics added to the prediction of radiographic OA development five years later. Quantitative radiographic evaluation in individuals with suspected OA is worthwhile when determining treatment strategies and designing clinical trials.

Introduction

Osteoarthritis (OA) is a disabling joint disease, which most commonly affects the knee joint. Symptoms of pain and functional limitations are assumed to be originated by structural changes like articular cartilage damage, osteophyte formation, synovial inflammation, and subchondral bone changes[1,2]. Radiography is the gold standard for demonstrating structural changes since image acquisition is non-invasive, cheap, fast, and generally available[1,3]. Radiographic OA is commonly graded according to Kellgren & Lawrence (K&L)[4]. A drawback is that this method only provides a qualitative (ordinal) score of a combination of structural aspects. It is generally appreciated that it takes at least a year before a change of one grade (scale 0-IV) becomes evident[5,6]. More detailed evaluation by quantitative measurement of separate features of joint damage might improve the association with clinical symptoms, which is currently not consistently found[7,8]. More importantly, it might enable the identification of initial tissue damage earlier in the disease, which is of value for the development of interventions to prevent structural damage[9].

In clinical practice patients visit a physician when they suffer from complaints that are possibly related to OA. Early in the disease process reliable diagnosis is difficult because structural damage can not yet be detected on radiographs using methods like K&L-grading. Also, pain often has an intermittent character and not all individuals suspected for the disease will eventually develop progressive OA. Higher age[10], higher body mass index (BMI)[11,12], and female gender[10] have been shown to be associated with the onset and progression of OA. In addition, the detection of early evident tissue damage by precise measurement on radiographs might predict the radiographic and or clinical course of disease[13,14]. In the hip joint the measurement of smaller joint space width was found to predict total hip replacement[15,16], but in the knee joint initial severity has not been found to be of evident additional value in the prediction of radiographic[17] and clinical progression[18].

The objective of the present study was to evaluate whether and which separate features of radiographic damage, measured quantitatively in knees with early symptoms related to OA, are associated with the incidence of radiographic OA and the persistence and/or progression of clinical OA during five-year follow-up.

Methods

Cohort Hip & Cohort Knee (CHECK)

Cohort Hip & Cohort Knee (CHECK) is a Dutch prospective multi-center ten-year follow-up study. Individuals (n=1002) with pain and/or stiffness of hip and/or knee, age 45-65 years, and without a previous visit or with a first visit no longer than six months ago to the general practitioner for these complaints were included. Individuals with pathological conditions other than OA explaining the complaints and individuals with K&L-grade IV were excluded[19]. The study procedures are in accordance with the standards of the medical ethics committees of

all ten participating hospitals and with the Helsinki Declaration of 1975 (as revised in 2000), and all participants gave their written informed consent.

Radiographic features baseline

Of both knees separately, standardized radiographs were acquired[20,21]. The baseline (T0) radiographs of both knees were evaluated for their predictive ability.

Radiographic parameters of knee OA were quantitatively measured by use of Knee Images Digital Analysis (KIDA)[22]. Key radiographic features were defined for evaluation in the present study, based on principal component analysis and on expert opinion (AM/JB/FL). The 'minimum JSW' (joint space width in mm) was measured as the smallest distance between femur and tibia. Also 'medial JSW' and 'lateral JSW' were determined by calculating the mean of four predefined locations (standardized, based on the joint dimensions[22]). The angle between the femur and tibia in the frontal plane was determined to represent the alignment of the joint ('varus angle' in degrees; positive value represents varus alignment). 'Osteophyte area' (in mm^2) was determined by summing the osteophyte area of the lateral and medial femur and tibia. 'Eminence height' was calculated as the sum of both eminences. 'Bone density' was determined as the mean of the bone density determined in the lateral and medial femur and tibia. Bone density was expressed in mmAl equivalents, by normalizing the gray values of the subchondral bone region to those of an aluminum reference wedge present on all radiographs[22]. The KIDA method is a mathematical interactive software tool to analyze knee radiographs and takes a few minutes per knee joint. Measurements were performed by one experienced observer (ML) in random order and blinded to any information (e.g. clinical characteristics). The intra-observer variation tested by random reanalysis of 108 radiographs several months later, revealed good intra-observer variability (ICC=0.73-0.99) for the different features.

The number of analyzed knees varied for the different radiographic features since KIDA measurement requires good radiographic quality. E.g. measurement of varus angle and eminence height was hampered in 10 (0.5% of 2004) knees, and osteophyte area could not be thoroughly outlined in 31 knees (1.5%). Specifically bone density measurement, which requires good contrast and a clearly visible aluminum reference wedge, was not always possible despite standardized procedures (28%). The baseline characteristics were not significantly different between participants without and with missing quality radiographs.

Radiographic OA at T0 was also assessed by the commonly used K&L-grading. K&L-grades were determined without knowledge of any other characteristics and reading was performed in pairs (T0 and two-year follow-up (T2y)) to obtain a reliable grade for T0 (T2y data not used in the present study).

OA development

The development of OA from T0 to five-year follow-up (T5y) was classified as 'poor' outcome (incidence, persistence, or progression) or 'good' outcome (no incidence, persistence, or progression) based on radiographic and clinical evaluation. OA development was evaluated in participants with complaints in at least one knee at study inclusion. For the different definitions of OA development separate analyses were performed with specific criteria for joint and participant exclusion (table 1).

Table 1 Definitions of OA development used as outcome in the different analyses

	Radiographic OA Incidence K&L	**Clinical OA Persistent pain**	**WOMAC pain**	**WOMAC function**
Inclusion of participants/knees				
Criteria: at T0 painful knee, and	K&L <II	Only painful knee	Hip: not painful and K&L <II	
(n)	(985 knees)	(1060 knees)	(286 participants)	(279 participants)
'poor' outcome				
Definition	K&L ≥II at T5y	Painful at T4y&T5y	Quintiles: highest 3/ move higher at T5y	
% of participants/knees	19%	48%	54%	56%

Radiographic OA: incidence

The incidence of radiographic OA ('poor' outcome) was defined on joint level (knees separately), as the development of a K&L-grade ≥II at T5y. Since knees needed to be susceptible for development of radiographic OA, knees with K&L-grade ≥II at T0 were excluded for these analyses. For each knee the T5y radiograph was graded according to K&L with the radiographs of T0 and T2y in view. The T5y radiograph was scored in another scoring session than the initial T0 grade, which was determined independently of the T5y outcome to prevent information bias.

Clinical OA: persistence

The development of 'poor' clinical outcome was evaluated on joint level and on participant level. For the definition on joint level all knee joints that were painful at T0 were used. The physician assessed this during examination of joint motion, for the left and right knee separately. Clinical persistence ('poor' outcome) was defined as still having a painful knee during physical examination both at four-year and at five-year follow-up (T4y and T5y), and otherwise the outcome was considered 'good'.

Clinical OA: progression

On participant level, the Western Ontario & McMaster Universities Osteoarthritis index (WOMAC) pain score and function score were used. For these analyses, confounding of the WOMAC scores by hip involvement was prevented by excluding participants with additional painful hip(s) and/or K&L-grade ≥II of the hip(s) at T0. WOMAC scores were standardized to a 0-100 scale with the maximum score representing the worst condition. Because WOMAC scores are recognized to be quite variable over time[23], the WOMAC 'baseline' value was

calculated as the mean of T0 and T1y, and the 'follow-up' score was calculated as the mean of T4y and T5y. Development (and persistence) of WOMAC pain and function values from 'baseline' to 'follow-up' were classified according to Sharma et al. using a quintile approach[24]. The clinical progression was defined as 'poor' when participants moved to a higher quintile or remained in the highest three quintiles, and the outcome was defined as 'good' when participants moved to a lower quintile or remained in the lowest two quintiles.

Statistical analyses

Separate binary logistic regression analyses were performed with 'good' (0) versus 'poor' (1) radiographic or clinical outcome as dependent variable. The radiographic characteristics defined as separate key features (KIDA) or K&L-grade were used as independent variables (for osteophyte area 'log [osteophyte area +1]' was calculated to obtain a more normal distribution). For analyses on joint level (K&L and persistent pain outcome) the value of a knee was used, and for analyses on participant level (WOMAC outcome) the sum of the left and right knee was evaluated to represent the total burden of disease.

Also, since radiographic characteristics might be dependent on characteristics of an individual (e.g. larger individuals have larger JSW and females have lower bone density), the difference-value (difference between knee and contralateral knee) was used as independent variable as well[25].

Furthermore, gender, age, BMI, and erythrocyte sedimentation rate (ESR in mm/hour), were added as independent variables. The latter was included because this parameter is frequently determined in this early stage of OA to exclude arthritic conditions. Also, dependent on the clinical outcome (persistence or progression), overall pain intensity (0-10 scale), WOMAC pain score, or WOMAC function score at baseline was added as independent variable.

Univariate and multivariate regression analyses were performed. In the multivariate analyses all variables were initially included and variables were removed manually using a backward stepwise selection procedure. Variables that were either statistically significantly ($p \leq 0.05$) related to the outcome or that changed the regression coefficient for one of the radiographic characteristics with >10% (confounding variables) were kept in the final model.

Separate models were constructed where instead of the separate quantitative features the K&L-grade was used. Models including radiographic characteristics (separate features or K&L-grade) were compared to models with only the demographic and clinical predictors. This approach aimed at representing clinical practice when a patient visits a physician with the first OA related symptoms. It was analyzed whether evaluation of (separate) radiographic characteristics, next to the assessment of basic demographic and clinical characteristics, adds to decision-making.

To evaluate the fit of the final models, Hosmer-Lemeshow tests were performed. Prognostic ability of the models was summarized using the area under the curve (AUC) of the receiver operating characteristic curve (ROC). The AUC-ROC provides a measure for the ability to

discriminate between 'good' and 'poor' outcome, where an AUC-ROC <0.70 was regarded as poor, 0.70-0.80 as fair, 0.80-0.90 as good, and ≥0.90 as excellent[26].
When the discriminative ability of the models was considered fair to good, also the sensitivity, specificity, and positive and negative predictive values (PPV and NPV) were calculated as prognostic statistics for different cut-off values. Therefore, the regression coefficients were corrected for over-fitting using the method of van Houwelingen and LeCessie[27] and the regression function was converted into a simple predictive score. Analyses were performed using SPSS 15.0, p-value ≤0.05 was considered statistically significant.

Results

Baseline characteristics

Of 1002 CHECK participants, 829 had pain in the knee(s) at study inclusion (1294 painful knees). Radiographs of good quality were available of 1060 (of the 1294) knee joints for analyses of persistent pain. Of these, data of 985 knees with K&L-grade <II at T0 were available for analyses of radiographic incidence. For analyses of WOMAC progression (participant level), data were available of 286 participants with no hip affection at T0.
Table 2 depicts demographic and clinical as well as radiographic characteristics at T0 of the respective datasets, separately for the participants with 'good' and 'poor' radiographic and clinical outcome.

Table 2 Demographic, clinical and radiographic characteristics at T0

	K&L (joint)		Persistent (joint)		WOMAC pain (pp)		WOMAC function (pp)	
Characteristics	**'good'**	**'poor'**	**'good'**	**'poor'**	**'good'**	**'poor'**	**'good'**	**'poor'**
n participants (k)	520 (796)	133 (189)	346 (551)	334 (509)	131	155	124	155
Age in years	56 (5)	56 (5)	56 (5)	56 (5)	56 (5)	56 (5)	56 (5)	56 (5)
Female gender	78%	86%	82%	80%	75%	81%	**73%**	**83%**
BMI in kg/m^2	**25 [23-28]**	**27 [24-31]**	25 [23-28]	26 [24-29]	**25 [23-27]**	**26 [24-29]**	**25 [23-27]**	**26 [24-29]**
ESR in mm/hour	8 [5-14]	8 [5-14]	8 [5-14]	8 [5-14]	**7 [4-11]**	**8 [5-14]**	**6 [4-11]**	**9 [5-15]**
Pain intensity	**3 [2-5]**	**4 [2-5]**	**3 [2-5]**	**4 [2-5]**	**3 [2-4]**	**3 [2-5]**	3 [2-5]	3 [2-5]
WOMAC pain	**20 [10-35]**	**25 [15-40]**	**20 [10-35]**	**25 [15-40]**	**20 [9-30]**	**20 [10-35]**	15 [10-30]	20 [10-31]
WOMAC function	**19 [9-34]**	**24 [13-38]**	**18 [10-31]**	**25 [13-38]**	**13 [6-25]**	**21 [10-32]**	13 [6-27]	20 [10-32]
K&L grade ≥II	0%	0%	2.5%	4.6%	3.2%	4.2%	3.0%	4.5%
Radiographic features								
Minimum JSW	**3.07 (1.1)**	**2.65 (1.4)**	**3.04 (1.2)**	**2.85 (1.3)**	2.99 (1.3)	2.99 (1.3)	**3.14 (1.2)**	**2.84 (1.3)**
Medial JSW	**4.80 (0.9)**	**4.49 (1.1)**	4.74 (0.9)	4.65 (1.0)	4.59 (1.0)	4.75 (1.1)	4.79 (1.1)	4.57 (1.0)
Lateral JSW	6.07 (1.4)	6.24 (1.5)	6.09 (1.3)	6.01 (1.4)	6.10 (1.4)	6.13 (1.4)	6.15 (1.3)	6.11 (1.5)
Varus angle	**1.68 (1.7)**	**2.32 (2.0)**	1.78 (1.8)	1.79 (1.8)	1.98 (1.9)	1.83 (2.0)	1.81 (1.9)	2.02 (1.9)
Osteophyte	**5.20 (4.9)**	**10.21 (9.1)**	**5.76 (5.6)**	**7.24 (7.6)**	6.97 (7.3)	7.55 (6.9)	6.83 (6.2)	7.75 (7.7)
Eminence	**22.7 (3.1)**	**23.2 (3.1)**	22.8 (3.0)	22.8 (3.2)	22.6 (3.1)	22.8 (3.3)	22.7 (3.1)	22.7 (3.2)
Bone density	**24.5 (6.1)**	**26.1 (6.2)**	24.5 (5.6)	25.0 (6.4)	25.2 (6.5)	24.3 (6.4)	25.4 (6.3)	24.2 (6.5)

(joint): defined at joint level, (pp): defined at participant level, (k): n knees, mean (standard deviation) or median [25-75th percentile] depicted, **bold**: significant difference between participants with 'good' and 'poor' outcome

Incidence of radiographic OA was observed in 189 of 985 knees (19% ‘poor’ outcome: K&L-grade ≥II at T5y). Persistent knee pain was observed in 509 of 1060 knees (48% ‘poor’ outcome). Clinical progression (‘poor’ outcome according to quintile approach) was observed in 155 of 286 (54%) and 155 (56%) of 279 participants evaluated for WOMAC pain and WOMAC function, respectively.

At T0, BMI was statistically significantly (Mann-Whitney test; **bold** in table 2) higher in participants classified as ‘poor' for K&L ($p<0.0001$), WOMAC pain ($p=0.006$), and WOMAC function ($p=0.01$) outcome. ESR was higher in case of ‘poor’ outcome for WOMAC pain ($p=0.01$) and WOMAC function ($p=0.001$). Pain intensity, WOMAC pain, as well as WOMAC function were relatively low at T0, but statistically significant higher (although slightly) in case of ‘poor’ K&L outcome ($p=0.03$, $p=0.02$, and $p=0.006$, respectively), persistent knee pain (all $p<0.0001$), WOMAC pain ($p=0.02$, $p<0.05$, and $p=0.001$), and WOMAC function outcome (WOMAC pain: $p=0.02$, function: $p=0.01$). At T0 only a small percentage of participants had K&L-grade ≥II, which was not significantly different between participants with ‘good’ and ‘poor’ outcome. For the evaluation of K&L outcome, knees were selected with K&L-grade <II at T0 and therefore K&L-grade ≥II was not present in this dataset. Most of the radiographic features were statistically significant different between the participants with ‘good’ and ‘poor’ K&L outcome. Minimum JSW (and osteophyte area) were significantly different between participants with ‘good’ and ‘poor’ persistent pain and WOMAC function outcome.

Predictors OA development

Radiographic OA: incidence

Table 3 depicts results of univariate and multivariate regression analyses with incidence of radiographic OA as dependent variable (K&L-grade ≥II at T5y). Odds ratios (OR) with 95% confidence interval (95%CI) and statistical significance (p) are depicted for the independent variables determined at T0. For the multivariate models the AUC-ROC is depicted.

Table 3 Regression analyses with radiographic OA (K&L grade ≥II) as dependent variable

	Univariate			Multivariate				
Variable	**OR**	**(95%CI)**	**p**	**Variable**	**OR**	**(95%CI)**	**p**	**AUC**
Demographic & clinical				**Demographic & clinical**				0.64*
Age	1.03	(1.00-1.06)	0.10					
Female gender	1.73	(1.10-2.72)	0.02	Female gender	1.68	(1.06-2.66)	0.03	
BMI	1.10	(1.06-1.14)	<0.0001	BMI	1.10	(1.06-1.14)	<0.0001	
ESR	1.00	(0.98-1.02)	0.76					
Pain intensity	1.08	(1.00-1.16)	0.05					
Radiographic				**Radiographic (demographic & clinical)**				
Feature				***Feature***				0.74*
Minimum JSW	0.75	(0.66-0.86)	<0.0001	Minimum JSW	0.74	(0.64-0.85)	<0.0001	
Medial JSW	0.67	(0.57-0.83)	<0.0001					
Lateral JSW	1.09	(0.97-1.22)	0.15					
Varus angle	1.22	(1.11-1.33)	<0.0001					
Osteophyte	6.30	(3.82-10.4)	<0.0001	Osteophyte	6.97	(4.11-11.8)	<0.0001	
Eminence	1.06	(1.00-1.12)	0.03					
Bone density	1.04	(1.01-1.08)	0.02					
Feature diff								
Minimum JSW	0.90	(0.74-7.08)	0.25					
Medial JSW	0.97	(0.84-1.12)	0.67					
Lateral JSW	0.86	(0.66-1.11)	0.86					
Varus angle	1.03	(0.94-1.14)	0.59					
Osteophyte	1.83	(1.17-2.86)	0.009					
Eminence	0.98	(0.91-1.07)	0.70					
Bone density	1.01	(0.92-1.11)	0.76					
				Demographic & clinical				
				Female gender	1.99	(1.20-3.31)	0.008	
				BMI	1.09	(1.05-1.14)	<0.0001	
K&L grade	3.54	(2.46-5.07)	<0.0001	***K&L grade***	4.74	(3.13-7.19)	<0.0001	0.70
K&L diff	0.86	(0.62-1.20)	0.38	***K&L diff***	0.56	(0.40-0.78)	0.001	
				Demographic & clinical				
				Female gender	1.90	(1.17-3.08)	0.009	
				BMI	1.09	(1.05-1.13)	<0.0001	

OR: odds ratio, (95%CI): 95% confidence interval, p: significance level, AUC: area under the receiver operating characteristic curve, diff: difference-value, note: female gender and BMI depicted for 3 multivariate models: demographic & clinical, and radiographic feature and K&L grade in addition to demographic & clinical, * statistically significantly different

For the multivariate models the fit was adequate (Hosmer-Lemeshow tests: p>0.05). The prognostic ability was clearly improved when radiographic characteristics were added to demographic and clinical characteristics. For the multivariate models using demographic and clinical characteristics, separate features, and K&L-grading as independent variables data of respectively 965, 904, and 955 (of 985) knees were available. The model with basic

demographic and clinical variables only revealed that female gender and higher BMI were statistically significant predictors of incidence of radiographic OA, which is in accordance with the literature[10-12]. Prognostic ability of this model was considered poor[26]: AUC-ROC=0.64 (95%CI: 0.59-0.68, p<0.0001). When key radiographic features obtained at T0 were added to the model the ability to predict radiographic outcome at T5y improved (AUC-ROC=0.74, 95%CI: 0.69-0.78, p<0.0001), resulting in fair prognostic ability. This AUC-ROC was statistically significantly higher than the AUC-ROC of the model with demographics and clinical variables only (p=0.007), as evaluated according to Hanley et al.[28] (in participants with complete data for both models). This model implied that knees with *smaller* minimum JSW (OR=0.74) and those with *larger* osteophyte area (OR=6.97) were more likely to have incident radiographic OA ('poor' outcome). Also the K&L-grade (0 or I) at T0 added to clinical and demographic variables as a predictor for radiographic OA incidence at T5y (OR=4.74). This model had (borderline) fair prognostic ability with AUC-ROC of 0.70 (95%CI: 0.66-0.74, p<0.0001), not statistically significantly different from the model with demographics only.

To evaluate whether the quantitative measurement of radiographic features can be applied in clinical practice, to identify individuals that are more likely to develop radiographic OA five years later, a simplified predictive score was calculated. The predictive score was based on the shrunken (shrinkage factor 0.98) and rounded regression coefficients (not the ORs as presented in the table) of the final logistic regression model including the key features of radiographic damage and demographic and clinical variables as: -0.5*minimum JSW + 2*(log[osteophyte area+1]) + 0.5*gender + 0.1*BMI. Based on the ROC-curve three cut-off points were evaluated for predictive ability: predictive score >2.50 (optimal sensitivity), score >3.65 (optimal trade-off between sensitivity and specificity), and score >4.60 (optimal specificity). Table 4 shows sensitivity, specificity, PPV, and NPV for these cut-offs.

Using the predictive score the AUC-ROC was 0.73 (95%CI: 0.69-0.77, p<0.0001) and the mean value was 3.44±1.13 [0.00-7.19]. When the predictive score was larger than 4.60 (e.g. a female with BMI of 30 kg/m^2, minimum JSW of 1.90 mm, and osteophyte area of 5.00 mm^2) the chance of incident radiographic OA at T5y was 54% (PPV), which was evidently larger than the incidence of radiographic OA in 19% of all knees (prior probability; 189 of 985 knees had 'poor' outcome). The chance of not developing radiographic OA (NPV) was 93% instead of 81% (prior probability) when the predictive score was 2.50 or lower (e.g. a female with BMI of 30 kg/m^2, minimum JSW of 3.00 mm and osteophyte area of 0.00 mm^2).

Table 4 Ability to predict radiographic incidence for three cut-off points of predictive score

	Cut-off		
Predictive ability	**2.50**	**3.65**	**4.60**
Sensitivity	93%	66%	38%
Specificity	23%	66%	92%
PPV	23%	32%	54%
NPV	93%	89%	86%

PPV: positive predictive value, NPV: negative predictive value

Clinical OA: persistence

Table 5 summarizes results of regression analyses on persistent knee pain. The difference-values of the radiographic characteristics (knee - contralateral knee) are not depicted since none of these variables were significant predictors of this outcome in univariate and multivariate analyses.

Table 5 Regression analyses with persistent knee pain as dependent variable

Univariate				Multivariate				
Variable	**OR**	**(95%CI)**	**p**	**Variable**	**OR**	**(95%CI)**	**p**	**AUC**
Demographic & clinical				**Demographic & clinical**				0.58
Age	0.97	(0.95-1.00)	0.03					
Gender	0.99	(0.73-1.35)	0.95					
BMI	1.03	(1.00-1.06)	0.03					
ESR	1.00	(0.98-1.01)	0.58					
Pain intensity	1.15	(1.08-1.22)	<0.0001	Pain intensity	1.15	(1.08-1.22)	<0.0001	
Radiographic				**Radiographic (demographic & clinical)**				
Feature				***Feature***				0.60
Minimum JSW	0.88	(0.80-0.98)	0.02	Minimum JSW	0.88	(0.79-0.97)	0.01	
Medial JSW	0.91	(0.80-1.04)	0.16					
Lateral JSW	0.96	(0.88-1.05)	0.37					
Varus angle	1.00	(0.94-1.08)	0.90					
Osteophyte	1.48	(1.09-2.02)	0.01	Osteophyte	1.54	(1.12-2.13)	0.008	
Eminence	1.01	(0.97-1.05)	0.73					
Bone density	1.02	(0.99-1.04)	0.21					
				Demographic & clinical				
				Age	0.97	(0.94-0.99)	0.01	
				Pain intensity	1.13	(1.06-1.20)	0.0002	
K&L grade	1.42	(1.12-1.82)	0.004	***K&L grade***	1.47	(1.14-1.88)	0.002	0.60
				Demographic & clinical				
				Age	0.97	(0.95-1.00)	<0.0001	
				Pain intensity	1.13	(1.07-1.21)	0.04	

OR: odds ratio, (95%CI): 95% confidence interval, p: significance level, AUC: area under the receiver operating characteristic curve, note: pain intensity (and age) depicted for 3 multivariate models: demographic & clinical, and radiographic feature and K&L grade in addition to demographic & clinical

The predictive value (OR) of radiographic characteristics was smaller for persistent knee pain outcome than for radiographic outcome in univariate and multivariate analyses. Of the participants with radiographic OA at T5y, 53% also had persistent pain (95 of 179). And of the participants with persistent pain, 22% had radiographic OA at T5y. Hosmer-Lemeshow tests showed adequate fit for the final models with radiographic characteristics (KIDA and K&L: p>0.05), but lack of fit for the model with demographics only (p=0.01).

The multivariate models all implied poor prognostic ability. The model with demographic and clinical characteristics only (n=1035 of 1060 knees) had AUC-ROC 0.58 (95%CI: 0.54-0.61, p<0.001). Adding radiographic variables hardly improved the ability to predict pain persistence: AUC-ROC=0.60 (95%CI: 0.56-0.64, p<0.001) for the separate radiographic features (n=957 knees) and AUC-ROC=0.60 (95%CI: 0.56-0.63, p<0.01) for K&L-grade (n=970 knees).

Clinical OA: progression

Table 6A and B depict results of regression analyses with WOMAC pain and function outcome as dependent variable, respectively.

The multivariate model with demographics and clinical characteristics (n=282 and 267 participants, respectively) had AUC-ROC of 0.59 (95%CI: 0.52-0.66, p<0.01) and 0.63 (95%CI: 0.56-0.70, p<0.001), respectively. Adding radiographic features slightly improved the prediction of WOMAC pain and function development (AUC-ROC=0.62 (95%CI: 0.55-0.68, p<0.001) and 0.65 (95%CI: 0.58-0.71, p<0.001), for pain and function, respectively). Interestingly, in the multivariate model for WOMAC pain the difference-value (between contralateral knees) of eminence height was a significant predictor of 'poor' outcome. When adding K&L-grade (difference) to demographics and clinical variables comparable poor predictive abilities were found.

Table 6A Regression analyses with WOMAC pain outcome as dependent variable

	Univariate			**Multivariate**				
Variable	**OR**	**(95%CI)**	**p**	**Variable**	**OR**	**(95%CI)**	**p**	**AUC**
Demographic & clinical			**Demographic & clinical**			0.59		
Age	0.96	(0.92-1.01)	0.13					
Female gender	1.40	(0.80-2.46)	0.24					
BMI	1.09	(1.02-1.17)	0.01	BMI	1.09	(1.02-1.17)	0.01	
ESR	1.05	(1.01-1.09)	0.02					
WOMAC pain	1.02	(1.00-1.03)	0.02					
Radiographic				**Radiographic (demographic & clinical)**				
Feature				***Feature***				0.62
Minimum JSW	0.98	(0.88-1.09)	0.77					
Medial JSW	1.09	(0.97-1.23)	0.16					
Lateral JSW	1.01	(0.92-1.11)	0.82					
Varus angle	0.97	(0.91-1.04)	0.46					
Osteophyte	1.37	(0.76-2.46)	0.29					
Eminence	1.01	(0.97-1.05)	0.55					
Bone density	0.99	(0.97-1.01)	0.24					
Feature abs diff				***Feature abs diff***				
Minimum JSW	1.13	(0.82-1.57)	0.45					
Medial JSW	1.05	(0.69-1.60)	0.81					
Lateral JSW	0.94	(0.73-1.21)	0.64					
Varus angle	0.94	(0.80-1.11)	0.49					
Osteophyte	1.44	(0.75-2.74)	0.28					
Eminence	0.86	(0.73-1.03)	0.10	Eminence	0.81	(0.67-0.98)	0.03	
Bone density	0.99	(0.86-1.13)	0.89					
				Demographic & clinical				
				ESR	1.05	(1.01-1.09)	0.02	
				WOMAC pain	1.02	(1.00-1.03)	0.05	
K&L grade	0.97	(0.76-1.25)	0.83					
K&L abs diff	0.53	(0.29-0.98)	0.04	***K&L abs diff***	0.46	(0.24-0.87)	0.02	0.63
				Demographic & clinical				
				ESR	1.05	(1.01-1.09)	0.03	
				WOMAC pain	1.02	(1.00-1.04)	0.03	

OR: odds ratio, (95%CI): 95% confidence interval, p: significance level, AUC: area under the receiver operating characteristic curve, abs diff: absolute difference, note: ESR and WOMAC pain depicted for 2 multivariate models: radiographic feature and K&L grade in addition to demographic & clinical

Table 6B Regression analyses with WOMAC function outcome as dependent variable

Univariate				Multivariate				
Variable	**OR**	**(95%CI)**	**p**	**Variable**	**OR**	**(95%CI)**	**p**	**AUC**
Demographic & clinical				**Demographic & clinical**				0.63
Age	1.01	(0.96-1.06)	0.74					
Female gender	1.79	(1.01-3.18)	0.05					
BMI	1.09	(1.02-1.16)	0.02					
ESR	1.06	(1.02-1.11)	0.002	ESR	1.05	(1.01-1.10)	0.01	
WOMAC function	1.02	(1.00-1.04)	0.01	WOMAC function	1.02	(1.00-1.04)	0.02	
Radiographic				**Radiographic (demographic & clinical)**				
Feature				***Feature***				0.65
Minimum JSW	0.89	(0.80-0.99)	0.03	Minimum JSW	0.89	(0.80-0.99)	0.04	
Medial JSW	0.89	(0.79-1.01)	0.07					
Lateral JSW	0.99	(0.90-1.09)	0.79					
Varus angle	1.04	(0.97-1.11)	0.31					
Osteophyte	1.31	(0.72-2.38)	0.38					
Eminence	1.00	(0.96-1.04)	0.83					
Bone density	0.99	(0.97-1.00)	0.10					
Feature abs diff				***Feature abs diff***				
Minimum JSW	1.04	(0.75-1.45)	0.82					
Medial JSW	1.18	(0.75-1.86)	0.47					
Lateral JSW	1.00	(0.77-1.29)	0.97					
Varus angle	1.02	(0.87-1.23)	0.75					
Osteophyte	1.95	(1.00-3.82)	0.05					
Eminence	0.91	(0.76-1.08)	0.27					
Bone density	0.96	(0.83-1.10)	0.54					
				Demographic & clinical				
				ESR	1.05	(1.01-1.09)	0.02	
				WOMAC function	1.02	(1.00-1.04)	0.02	
K&L grade	1.10	(0.85-1.42)	0.46					
K&L abs diff	0.71	(0.39-1.30)	0.71					

OR: odds ratio, (95%CI): 95% confidence interval, p: significance level, AUC: area under the receiver operating characteristic curve, abs diff: absolute difference, note: ESR and WOMAC pain depicted for the multivariate model with radiographic features in addition to demographic & clinical

Hosmer-Lemeshow tests revealed adequate fit for all multivariate models with WOMAC outcome. For the clinical outcomes, the AUC-ROC was not statistically significantly different between the multivariate models (e.g. model with demographics and clinical variables only *vs.* model where radiographic variables (features or K&L-grade) were added).

For the analyses on participant level a portion of the individuals contributed with both knees to the regression analyses. To account for this dependency generalized estimating equations (GEE) were performed, which resulted in (nearly) the same OR and p-values as the regression analyses.

Discussion

In individuals that presented themselves with very early complaints related to knee osteoarthritis (OA), evaluation of radiographic characteristics added to the prediction of incident radiographic OA five years later. The evaluation of separate quantitative features

performed better in this respect than a simple K&L-grade. Radiographic characteristics hardly added to the prediction of persistence and/or progression of clinical OA, and total predictive ability of these models was too low for use in practice.

The additional value of radiographic characteristics in the prediction of radiographic progression has been described for more advanced OA (e.g. K&L-grade ≥II)[14,29-31]. And in a recent review also radiographic features, varus alignment, age, and BMI[32] were identified as predictors for OA progression later in disease. This is the first study however to demonstrate that quantitative radiographic features, identified in individuals that present themselves with knee pain but without radiographic damage (K&L-grade <II), can add to the prediction of incident radiographic OA within five years.

The finding that radiographic characteristics were of additional value in the prediction of radiographic outcome, but hardly in the prediction of clinical outcome[18,32] is in accordance with the commonly reported inconsistent association between radiographic and clinical characteristics of OA[8,33]. Even when structural damage was evaluated with magnetic resonance imaging, only a weak correlation between change in WOMAC score and in cartilage thickness was found[34].

In the present study basic demographics together with clinical characteristics poorly predicted clinical outcome (persistence and/or development). This limited predictive ability for clinical outcome might be explained by the subjective nature of these outcomes. A limited set of variables that was commonly used in patient care was chosen for analyses, based on undemanding application in clinical practice.

Despite the limited severity and development of complaints in this very early OA cohort, the quintile approach[24] discriminated participants with 'good' and 'poor' outcome with significantly different scores at T5y. The WOMAC function score was 6 [1-12] for participants with 'good' and 29 [19-41] for participants with 'poor' outcome, and the WOMAC pain score was 8 [5-15] for those with 'good' outcome and 28 [20-40] in those with 'poor' outcome. Irrespectively, specifically these individuals (at risk of developing (radiographic) OA) present themselves with complaints for the first time. And although at that time radiography (and demographic and clinical characteristics) can hardly predict clinical outcome, specific key features obtained with KIDA measurement significantly add to the prediction of radiographic outcome.

The fact that WOMAC outcomes were measured on participant level, contrary to the radiographic outcome on joint level, might also in part explain the limited predictive ability of radiographic characteristics for this clinical outcome. The sum of the radiographic characteristics might be more appropriate to detect an association with clinical outcome in case of (more severe) bilateral OA, when the radiographic characteristics are expected to be more pronounced. In case of (milder) unilateral OA, as will be the case when individuals present themselves for the first time with complaints, the sum of one unaffected and one (slightly) affected joint might underestimate radiographic severity and the difference between the knees might be more appropriate[25]. In general however, in our study difference-values between both knee joints in the radiographic characteristics did not appear better in

predicting the incidence of radiographic OA or the persistence or progression of clinical OA. Surprisingly, the difference-value between both knees in eminence height and in K&L-grade were found to be predictors of WOMAC pain outcome, while the sum of the measured values were not significantly associated with this outcome. The odds ratios were low however, and moreover both OR's were smaller than 1 which implies that a larger difference between joints is protective for 'poor' outcome in contrast to our hypothesis. However, the predictive ability of the models (AUC-ROC) was poor indicating that predicting WOMAC outcome is difficult either with or without radiographic features.

The separate key features that were identified as additional predictors for incidence of radiographic OA and for clinical persistence were minimum JSW and osteophyte area. These separate features measured quantitatively using KIDA are also the most important characteristics in K&L-grading. This explains why also K&L-grading added to demographics and clinical variables in predicting radiographic incidence of OA. When all radiographic variables (minimum JSW, osteophyte area and K&L-grading) were added in logistic regression analysis, these were all significant predictors and predictive ability was even slightly higher than for KIDA variables and K&L-grading separately (AUC-ROC=0.76). Since OR was strongest for osteophyte area (5.02), measuring separate features is of value in addition to demographic and clinical characteristics (and K&L-grading). Also, since the KIDA predictors performed better than K&L-grading in this cohort, measuring separate radiographic features might improve the detection of radiographic OA earlier in the disease process. This was supported by the detection of joint space narrowing and osteophyte formation when qualitatively evaluating separate parameters by the Altman atlas[35], with larger AUC-ROC than K&L-grade (0.76, data not shown).

Next to the advantage of evaluation of separate key features of joint damage, KIDA measurement uses a mathematical approach and is performed without any knowledge of the knee and the individual. Intra-observer variation of KIDA was low[22,25], however variation in the measurements might occur during image acquisition. Despite optimal standardization[36] the position of the tibial plateau is subject to variation[37], which decreases comparability between and within individuals. Due to such variations, the additional value of KIDA might be underestimated in the present study, and might be improved when reproducibility of radiographic acquisition is further optimized in clinical trials. Further, the predictive ability of K&L-grading might be overestimated since in the present study the definition of K&L-grade at T0 was determined with knowledge of the K&L-grade at T2y (although not at T5y, this might be regarded a 'proxy' and as such the T0 measurement can not be regarded fully blinded like the KIDA measurements). This is of course also contrary to clinical practice when individuals present themselves with joint pain related to OA, and a choice on treatment (or on inclusion in a trial) is preferably made within a short time span, without waiting for a second radiograph one or two years later.

By use of the predictive scores (based on separate features) a subgroup of individuals was distinguished with a higher chance of onset of radiographic OA (54% compared to incidence of 19%). Although this chance is too low for decision making at the individual level,

identifying this group might advance the design of OA trials[38] for the development of more specific (disease modifying) treatment strategies[39]. The predictive score and cut-off values as determined in this study need internal and external validation before used in (clinical) practice. Also the fit of the models might be improved by investigating for instance different transformations of the predictors.
In conclusion, the prediction of incidence of radiographic OA improves from poor to fair when quantitative radiographic features are evaluated, in addition to basic demographics and clinical assessment, in individuals that visit a physician with early complaints possibly related to OA. Therefore the measurement of separate features might be valuable in identifying individuals at high risk of developing radiographic osteoarthritis.

Acknowledgements

We acknowledge Marja Lafeber (ML) for the KIDA measurements.

References

1. Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: an update with relevance for clinical practice. Lancet 2011;377(9783):2115-26.
2. Peat G, Thomas E, Duncan R, Wood L, Hay E, Croft P. Clinical classification criteria for knee osteoarthritis: performance in the general population and primary care. Ann Rheum Dis 2006;65(10):1363-7.
3. Nevitt MC. Definition of hip osteoarthritis for epidemiological studies. Ann Rheum Dis 1996;55(9):652-5.
4. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16(4):494-502.
5. Boegard TL, Rudling O, Petersson IF, Jonsson K. Joint space width of the tibiofemoral and of the patellofemoral joint in chronic knee pain with or without radiographic osteoarthritis: a 2-year follow-up. Osteoarthritis Cartilage 2003;11(5):370-6.
6. Vignon E, Conrozier T, Hellio Le Graverand MP. Advances in radiographic imaging of progression of hip and knee osteoarthritis. J Rheumatol 2005;32(6):1143-5.
7. Dieppe PA, Cushnaghan J, Shepstone L. The Bristol 'OA500' study: progression of osteoarthritis (OA) over 3 years and the relationship between clinical and radiographic changes at the knee joint. Osteoarthritis Cartilage 1997;5(2):87-97.
8. Kinds MB, Welsing PM, Vignon EP, Bijlsma JW, Viergever MA, Marijnissen AC, et al. A systematic review of the association between radiographic and clinical osteoarthritis of hip and knee. Osteoarthritis Cartilage 2011.
9. Hunter DJ. Lower extremity osteoarthritis management needs a paradigm shift. Br J Sports Med 2011.
10. 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.
11. Grotle M, Hagen KB, Natvig B, Dahl FA, Kvien TK. Obesity and osteoarthritis in knee, hip and/or hand: an epidemiological study in the general population with 10 years follow-up. BMC Musculoskelet Disord 2008;9:132.
12. Heliovaara M, Makela M, Impivaara O, Knekt P, Aromaa A, Sievers K. Association of overweight, trauma and workload with coxarthrosis. A health survey of 7,217 persons. Acta Orthop Scand 1993;64(5):513-8.
13. Ding C, Jones G, Wluka AE, Cicuttini F. What can we learn about osteoarthritis by studying a healthy person against a person with early onset of disease? Curr Opin Rheumatol 2010;22(5):520-7.
14. Eckstein F, Le Graverand MP, Charles HC, Hunter DJ, Kraus VB, Sunyer T, et al. Clinical, radiographic, molecular and MRI-based predictors of cartilage loss in knee osteoarthritis. Ann Rheum Dis 2011;70(7):1223-30.
15. Dougados M, Gueguen A, Nguyen M, Berdah L, Lequesne M, Mazieres B, et al. Requirement for total hip arthroplasty: an outcome measure of hip osteoarthritis? J Rheumatol 1999;26(4):855-61.
16. Vinciguerra C, Gueguen A, Revel M, Heuleu JN, Amor B, Dougados M. Predictors of the need for total hip replacement in patients with osteoarthritis of the hip. Rev Rhum Engl Ed 1995;62(9):563-70.
17. Belo JN, Berger MY, Reijman M, Koes BW, Bierma-Zeinstra SM. Prognostic factors of progression of osteoarthritis of the knee: a systematic review of observational studies. Arthritis Rheum 2007;57(1):13-26.
18. Thomas E, Peat G, Mallen C, Wood L, Lacey R, Duncan R, et al. Predicting the course of functional limitation among older adults with knee pain: do local signs, symptoms and radiographs add anything to general indicators? Ann Rheum Dis 2008;67(10):1390-8.
19. Wesseling J, Dekker J, van den Berg WB, Bierma-Zeinstra SM, Boers M, Cats HA, et al. CHECK (Cohort Hip and Cohort Knee): similarities and differences with the Osteoarthritis Initiative. Ann Rheum Dis 2009;68(9):1413-9.
20. Buckland-Wright C. Protocols for precise radio-anatomical positioning of the tibiofemoral and patellofemoral compartments of the knee. Osteoarthritis Cartilage 1995;3 Suppl A:71-80.
21. Buckland-Wright JC, Wolfe F, Ward RJ, Flowers N, Hayne C. Substantial superiority of semiflexed (MTP) views in knee osteoarthritis: a comparative radiographic study, without fluoroscopy, of standing extended, semiflexed (MTP), and schuss views. J Rheumatol 1999;26(12):2664-74.
22. Marijnissen AC, Vincken KL, Vos PA, Saris DB, Viergever MA, Bijlsma JW, et al. Knee Images Digital Analysis (KIDA): a novel method to quantify individual radiographic features of knee osteoarthritis in detail. Osteoarthritis Cartilage 2008;16(2):234-43.

23. van Dijk GM, Veenhof C, Spreeuwenberg P, Coene N, Burger BJ, van Schaardenburg D, et al. Prognosis of limitations in activities in osteoarthritis of the hip or knee: a 3-year cohort study. Arch Phys Med Rehabil 2010;91(1):58-66.
24. Sharma L, Cahue S, Song J, Hayes K, Pai YC, Dunlop D. Physical functioning over three years in knee osteoarthritis: role of psychosocial, local mechanical, and neuromuscular factors. Arthritis Rheum 2003;48(12):3359-70.
25. Kinds MB, Vincken KL, Vignon E, ten Wolde S, Bijlsma JW, Welsing PM, et al. Radiographic features of knee and hip osteoarthritis represent characteristics of an individual, in addition to severity of osteoarthritis. Scand J Rheumatol 2011:*Accepted for publication*.
26. Muller MP, Tomlinson G, Marrie TJ, Tang P, McGeer A, Low DE, et al. Can routine laboratory tests discriminate between severe acute respiratory syndrome and other causes of community-acquired pneumonia? Clin Infect Dis 2005;40(8):1079-86.
27. Van Houwelingen JC, Le Cessie S. Predictive value of statistical models. Stat Med 1990;9(11):1303-25.
28. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148(3):839-43.
29. Mazzuca SA, Brandt KD, Katz BP, Ding Y, Lane KA, Buckwalter KA. Risk factors for progression of tibiofemoral osteoarthritis: an analysis based on fluoroscopically standardised knee radiography. Ann Rheum Dis 2006;65(4):515-9.
30. Bruyere O, Honore A, Ethgen O, Rovati LC, Giacovelli G, Henrotin YE, et al. Correlation between radiographic severity of knee osteoarthritis and future disease progression. Results from a 3-year prospective, placebo-controlled study evaluating the effect of glucosamine sulfate. Osteoarthritis Cartilage 2003;11(1):1-5.
31. Wolfe F, Lane NE. The longterm outcome of osteoarthritis: rates and predictors of joint space narrowing in symptomatic patients with knee osteoarthritis. J Rheumatol 2002;29(1):139-46.
32. Chapple CM, Nicholson H, Baxter GD, Abbott JH. Patient characteristics that predict progression of knee osteoarthritis: A systematic review of prognostic studies. Arthritis Care Res (Hoboken) 2011;63(8):1115-25.
33. Bedson J, Croft PR. The discordance between clinical and radiographic knee osteoarthritis: a systematic search and summary of the literature. BMC Musculoskelet Disord 2008;9:116.
34. Buck R, Nevitt M, Wirth W, Eckstein F. Correlation of knee cartilage thickness change at 1 year and clinical outcome changes at 3 years: data from the osteoarthritis initiative. Presented at OARSI imaging workshop 2011.
35. Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. Osteoarthritis Cartilage 2007;15 Suppl A:A1-56.
36. Mazzuca SA, Brandt KD, Buckwalter KA, Lane KA, Katz BP. Field test of the reproducibility of the semiflexed metatarsophalangeal view in repeated radiographic examinations of subjects with osteoarthritis of the knee. Arthritis Rheum 2002;46(1):109-13.
37. Botha-Scheepers S, Dougados M, Ravaud P, Hellio Le Graverand MP, Watt I, Breedveld FC, et al. Effect of medial tibial plateau alignment on serial radiographs on the capacity to predict progression of knee osteoarthritis. Osteoarthritis Cartilage 2008;16(2):272-6.
38. Lane NE, Brandt K, Hawker G, Peeva E, Schreyer E, Tsuji W, et al. OARSI-FDA initiative: defining the disease state of osteoarthritis. Osteoarthritis Cartilage 2011;19(5):478-82.
39. Felson DT. Identifying different osteoarthritis phenotypes through epidemiology. Osteoarthritis Cartilage 2010;18(5):601-4.

Chapter 7

Identifying radiographic phenotypes of early knee osteoarthritis using separate quantitative features might improve patient selection for more targeted treatment

MB Kinds[1,2]
ACA Marijnissen[1]
MA Viergever[2]
PJ Emans[3]
FPJG Lafeber[1]
PMJ Welsing[1,4]

[1]Rheumatology & Clinical Immunology, University Medical Center Utrecht
[2]Image Sciences Institute, University Medical Center Utrecht
[3]Orthopaedic Surgery, Maastricht University Medical Center, Maastricht
[4]Julius Center for Health Sciences & Primary Care, University Medical Center Utrecht

Submitted for publication

Abstract

Objective

The expression of osteoarthritis (OA) varies significantly between individuals, and over time, implying the existence of different phenotypes possibly with specific etiology and targets for treatment. This study aims at identifying phenotypes of progression of radiographic knee OA using separate quantitative features.

Methods

Separate radiographic parameters of OA were measured by Knee Images Digital Analysis (KIDA) in individuals with early knee OA (Cohort Hip & Cohort Knee: CHECK), at baseline, two-year and five-year follow-up. Hierarchical clustering was performed to identify phenotypes of radiographic knee OA progression. The phenotypes identified were compared for development of joint space width (JSW), varus angle, osteophyte area, eminence height, bone density, and on clinical characteristics. Logistic regression analysis evaluated whether baseline radiographic features could predict to which phenotype an individual belonged.

Results

The five identified clusters were interpreted as ‘severe’ or ‘no’, ‘early’ or ‘late’ progression of the radiographic features, or specific involvement of ‘bone density’. Larger medial JSW, varus angle, osteophyte area, eminence height, and bone density at baseline were associated with the ‘severe’ and ‘bone density’ phenotypes. Smaller eminence height and bone density were associated with ‘early’ and ‘late’ progression. Larger varus angle and smaller osteophyte area were associated with ‘no’ progression.

Conclusion

Five phenotypes of radiographic progression of early knee OA were identified using separate quantitative features, which could be predicted by baseline radiographic features. These phenotypes might need specific treatment and represent relevant subgroups for clinical trials.

Introduction

Osteoarthritis (OA) is a degenerative joint disease characterized by pain and functional disability. Particularly knee OA has a high and increasing prevalence and is considered a major health and economic problem[1]. Structural changes affect the whole joint and include cartilage, bone, and soft tissues[2]. The definition of the disease and of diagnostic criteria remains difficult despite all efforts in research on OA over the past years[3,4]. This is mainly due to the (apparent) inconsistent relation between clinical symptoms and radiographic characteristics (representing structural damage) of OA[5-8], and the generally slow progression of the degenerative process early in the disease[9]. Despite this, radiography is, the primary outcome to prove disease-modifying efficacy (tissue structure modification) of treatment[10,11].

In clinical practice expression of disease varies significantly between patients and over time, and therefore it is appreciated that different phenotypes (subpopulations) of OA exist[10-12]. For instance, in patients with prominent inflammation a more destructive type of OA is found[13]. It is hypothesized that also radiographic phenotypes of OA exist. For example, some patients may mainly suffer from bone changes, while others have predominant damage of the cartilage. These radiographic phenotypes may also have their specific clinical character. For example, patients with predominant bone changes may sense more pain[14], and patients with osteophyte growth may have more joint inflammation as these phenomena have been linked[15]. The level of progression and the sequence of occurrence of different radiographic characteristics may vary for different phenotypes of radiographic knee OA. Such subtle differences will be missed when progression of radiographic joint damage is evaluated by the commonly used Kellgren & Lawrence (K&L) grading which is a rough score (0-IV) summarizing multiple characteristics[16]. These limitations hamper selection of subgroups of individuals for which specific treatment strategies might be needed. In case of bone involvement bisphosphonates might be effective, but this benefit will not be detected in the average OA population. Treatment of inflammation might do more harm than good in the overall OA population[17], but might be very helpful for subgroups of patients with evident inflammation. Identifying radiographic phenotypes is expected to improve by quantitative evaluation of separate features on radiographs. The objective of the present study is to identify radiographic phenotypes of early knee OA and to describe their radiological and clinical characteristics.

Methods

Cohort Hip & Cohort Knee (CHECK)

Development of knee OA was evaluated from baseline to five-year follow up in CHECK (Cohort Hip & Cohort Knee). In this cohort 1002 participants with pain and/or stiffness of hip and/or knee, age 45-65 years, and without a previous visit or with a first visit no longer than

six months ago to the general practitioner for these complaints were included[18]. At baseline 82% of the participants had knee complaints, and the radiographic damage of the entire cohort was limited with K&L grade in the knee of 0 in 81%, I in 16%, II in 3%, and III in 0.4%.

Knee Images Digital Analysis (KIDA)

Standardized weight bearing semiflexed views (Buckland-Wright)[19,20] were acquired of both knees at baseline, two-year and five-year follow-up (T0, T2y, and T5y). Radiographs were analyzed for fourteen separate parameters of OA by use of KIDA (Knee Images Digital Analysis)[21]: minimum joint space width (JSW in mm), mean medial and mean lateral JSW, femur-tibia varus angle (in degrees), eminence height (in mm), osteophyte area (in mm^2) in lateral and medial femur and lateral and medial tibia, and bone density (in mmAl equivalents) in these four compartments.[22] The KIDA measurements were performed by one experienced observer (ML) in random order unaware of information on time point, severity and characteristics of an individual. The numbers of analyzed knees are indicated and vary slightly for the different radiographic parameters since poor radiographic quality can hamper KIDA measurement.

Statistical analyses

Using principal component analysis, the measurements of separate KIDA parameters were reduced into five components to represent the following radiographic features: medial JSW, lateral JSW, osteophyte area, eminence height, and bone density (manuscript submitted). By multiplying the factor loadings from the principal component analysis of the KIDA measurements, five component scores were calculated. In the present study these five component scores (standardized using z-scores) were used in a hierarchical cluster analysis (Ward's method) to identify possible phenotypes of progression of radiographic knee OA. Per individual the component scores of the left and right knee at T0, T2y, T5y, and the change scores (T5y-T2y and T2y-T0) were all used in this analysis. The number of selected clusters was based on inspection of dendrograms.

To interpret the clusters (phenotypes), the following radiographic features were evaluated over time and compared between the clusters: minimum JSW, medial JSW, lateral JSW, varus angle, osteophyte area (log transformed sum of four compartments +1), eminence height (sum of both), and bone density (mean of four compartments). Further, the presence of knee and/or hip pain and the WOMAC pain and function scores (Western Ontario and McMaster Universities Osteoarthritis index; 0-100 scale, 100=worst condition) were compared between clusters.

Subsequently, logistic regression analyses were performed to evaluate whether the radiographic features measured at T0, in addition to demographic and clinical characteristics (age, gender, BMI, erythrocyte sedimentation rate (ESR), and pain intensity) at T0, could be used to predict to which specific phenotype an individual belongs. These analyses were performed in participants with knee pain at T0, since these individuals form the population that visits a physician with early complaints suspected for radiographic OA development.

Univariate and multivariate analyses were performed. In the multivariate analyses all variables were initially included and were removed manually using backward stepwise selection of variables that were statistically significantly related to the outcome. Models including only demographics and clinical variables were compared to models where radiographic features were added and to models where conventional K&L grading was added. To represent the total burden of radiographic damage, for each participant the sum of the left and right knee was used in the models. Since radiographic features might be very characteristic of an individual[22], the difference between a knee and the contralateral knee for the radiographic features was also studied as independent variable. These difference scores might detect small changes by using the contralateral knee as a reference in this early OA population with only subtle damage in one joint. Prognostic ability of the final models was summarized and compared using the area under the curve (AUC) of the receiver operating characteristic (ROC) curve. The AUC-ROC provides a measure for the ability to discriminate between a specific phenotype and the other phenotypes; an AUC-ROC <0.70 is regarded as poor, 0.70-0.80 as fair, 0.80-0.90 as good, and ≥0.90 as excellent[23]. Additionally, per phenotype the regression coefficients of the final models were corrected for over-fitting using the van Houwelingen and LeCessie method[24], and were converted into a simple score. Three cut-off points were determined: optimal sensitivity, optimal trade-off between sensitivity and specificity, and optimal specificity. For these cut-off values positive predictive values (PPV) were calculated as estimate of predictive ability.
Analyses were performed using SPSS (Statistical Package for the Social Sciences) version 15.0 and SAS (Statistical Analysis System) version 9.1.3; p-value <0.05 was considered statistically significant.

Results

Identification of radiographic phenotypes

Based on the development over time of the component scores of the knees, five clusters could be identified. Participants were only classified when complete data of KIDA measurements were available on all three time points (417 of 1002 participants). The five clusters were interpreted as:

1. ‘Severe’: severe progression;
2. ‘Bone density’: clear involvement of the bone density feature;
3. ‘Early’: progression mainly in an early phase (T0-T2y);
4. ‘Late’: progression mainly in a later phase (T2y-T5y);
5. ‘No progression’: no progression.

Figure 1 depicts the development of the separate radiographic features over time per cluster; top-row: the clusters representing the level of progression, middle-row: the cluster representing involvement of a specific feature, and bottom-row: the clusters representing different phasing of progression.

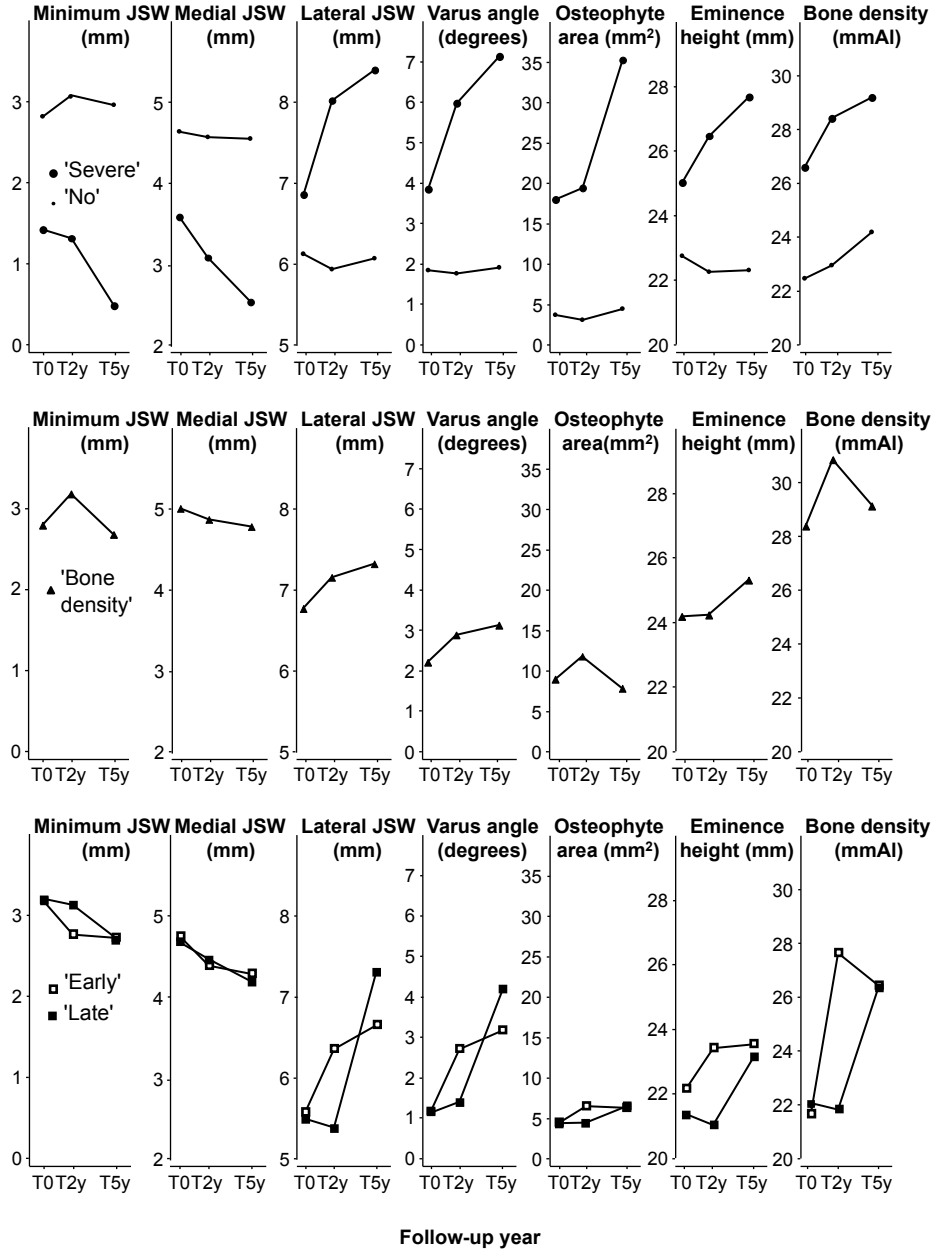

Figure 1: Development of radiographic features (right knee) per cluster of progression of radiographic knee OA. Results were similar for the left knee (data not shown)

In general, the radiographic features showed OA progression during five years follow-up: minimum and medial JSW decreased, and lateral JSW, varus angle, osteophyte area, eminence height, and bone density increased. Participants in the 'severe' cluster (top-row, large black dots: n=17; 4% of 417 available participants) progressed more evidently than participants in the other clusters on all radiographic features. Interestingly, at T0 these participants were already more affected on the features than the participants in the other clusters. The 'bone density' cluster (middle-row, black triangles: n=113; 27% of participants) represented severe involvement of bone density at all three time points compared with the other phenotypes. In this cluster the other features were only mildly affected. Participants in the 'early' cluster (bottom-row, unfilled squares: n=110; 26% of participants) mainly progressed between T0 and T2y, most evidently for lateral JSW, varus angle, and bone density. Participants in the 'late' cluster (lower panel, filled squares: n=69; 17% of participants) mainly progressed between T2y and T5y on lateral JSW, varus angle and eminence height. In the 'no progression' cluster (top-row, small black dots: n=108; 26% of participants) the radiographic features did not progress during follow-up.

Characterization of radiographic phenotypes

Baseline characteristics

Baseline characteristics are depicted per phenotype in table 1.

Table 1 Baseline characteristics per cluster of progression of radiographic knee OA

Characteristic	**'Severe'** (n=17)	**'Bone density'** (n=113)	**'Early'** (n=110)	**'Late'** (n=69)	**'No'** (n=108)	**p-value** overall
Age in years	58 (4)	56 (5)	56 (5)	56 (5)	57 (5)	0.16
Female gender	82%	55%	81%	88%	92%	<0.0001
BMI in kg/m^2	27 [25-31]	27 [24-30]	24 [23-27]	24 [23-27]	24 [22-27]	<0.0001
ESR in mm/hour	9 [5-15]	6 [4-12]	8 [5-15]	8 [5-13]	9 [5-15]	0.07
K&L ≥II knee	26%	4%	3%	2%	2%	<0.0001

Mean (SD) or median [25-75th percentile] depicted, BMI: body mass index, ESR: erythrocyte sedimentation rate, p-values depicted based on ANOVA (age), Chi-square (gender and K&L grade), and Kruskal-Wallis tests (BMI and ESR)

Clinical development

For further interpretation of the phenotypes, table 2 depicts the presence of pain in knee and/or hip at T0, T2y, and T5y. The location of pain was significantly different between the phenotypes (Chi-square test: p=0.002 at T0, p=0.001 at T2y, and p<0.0001 at T5y). Participants with 'severe' radiographic progression specifically presented with knee pain. Participants in the 'late' cluster reported pain in "hip only" more commonly than participants in the other clusters which might suggest (early) hip affection, followed by knee affection. Interestingly, a substantial part of the participants report "knee nor hip" pain at T2y and T5y, specifically in the 'no progression' cluster. This may indicate that this phenotype concerns acute transient joint pain that does not lead to progressive radiographic joint damage.

Table 2 Presence of pain over time per cluster of progression of radiographic knee OA

Pain (%)	**'Severe'** (n=17)			**'Bone density'** (n=113)			**'Early'** (n=110)			**'Late'** (n=69)			**'No'** (n=108)		
Time point	T0	T2y	T5y	T0	T2y	T5y	T0	T2y	T5y	T0	T2y	T5y	T0	T2y	T5y
Knee only	71	70	53	44	32	29	40	32	30	30	31	27	34	28	21
Knee&hip	24	12	35	42	38	40	41	46	39	33	37	41	49	33	29
Hip only	5	0	0	14	13	13	19	13	13	36	23	17	17	14	6
Knee nor hip	-	18	12	-	17	18	-	8	18	- 9	15	-	25	44	

T0: baseline, T2y: two-year follow-up, T5y: five-year follow-up

Figure 2 depicts the development of the average WOMAC pain and function score over time per progression phenotype. The WOMAC scores were moderate at all time points and did not evidently increase during follow-up. Although the average level of WOMAC pain score over time was statistically significantly lower in the 'no progression' phenotype than in the 'severe' (p=0.003), 'bone density' (p=0.02) and 'late' (p=0.02) phenotypes, the development over time was not significantly different between the phenotypes (tested using longitudinal regression analysis including an interaction term for time*phenotype). Also the average level of WOMAC function score was significantly lower in the 'no progression' phenotype than in the 'severe' (p=0.004),'bone density' (p=0.03), 'early' (p=0.04), and 'late' (p=0.04) phenotypes. Furthermore, the development over time was significantly different between the 'no' and 'late' phenotype.

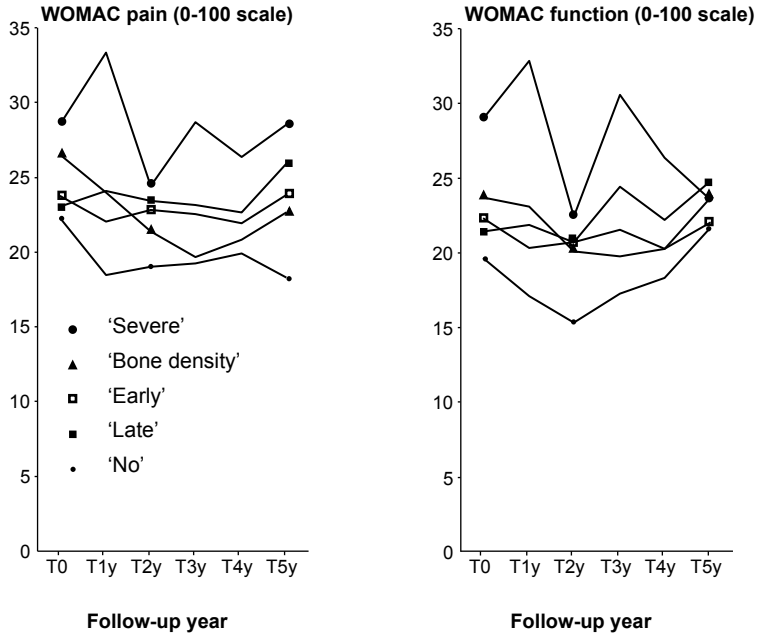

Figure 2 Development of WOMAC scores (0-100 scale, 100=worst condition) per cluster of progression of radiographic knee OA

Prediction of phenotypes

Table 3 depicts results of logistic regression analyses to evaluate which baseline variables predict 'membership' of a specific phenotype, as compared to all other phenotypes. This was evaluated in 336 participants with knee pain at T0, since these are the individuals clinically suspected of knee OA.

Table 3 Regression analyses with phenotype as dependent variable

3A 'Severe' progression				
	Univariate			**Multivariate**
Variable	**OR**	**(95%CI)**	**p**	
Demographic & clinical				**(NA)**
Age	1.08	(0.97-1.19)	0.15	
Female gender	1.28	(0.36-4.62)	0.71	
BMI	1.18	(1.07-1.30)	0.001	
ESR	0.99	(0.93-1.06)	0.84	
Pain intensity	1.20	(0.96-1.51)	0.12	
Radiographic				
Feature sum				
Minimum JSW	0.58	(0.46-0.74)	<0.001	
Medial JSW	0.49	(0.35-0.68)	<0.001	
Lateral JSW	1.15	(0.95-1.40)	0.15	
Varus angle	1.37	(1.13-1.60)	<0.001	
Osteophyte	15.48	(4.53-52.9)	<0.001	
Eminence	1.13	(1.05-1.23)	0.002	
Bone density	1.04	(1.00-1.09)	0.06	
Feature abs d				
Minimum JSW	2.85	(1.69-4.79)	<0.001	
Medial JSW	4.15	(2.18-7.90)	<0.001	
Lateral JSW	1.68	(1.18-2.38)	0.004	
Varus angle	0.57	(0.45-0.74)	<0.001	
Osteophyte	31.0	(6.1-158.1)	<0.001	
Eminence	1.72	(1.29-2.28)	<0.001	
Bone density	0.90	(0.64-1.26)	0.54	
K&L sum	2.62	(1.79-3.84)	<0.001	
K&L abs d	2.17	(0.76-6.20)	0.15	

Since the 'severe' phenotype consisted of only 16 participants with knee pain, for this outcome multivariate analyses were not performed. In the univariate evaluation almost all radiographic features were significant predictors, as were K&L grade and BMI.

In general the multivariate analyses showed that the discriminative ability (AUC-ROC) of the models improved when radiographic features were added to the demographic and clinical variables. The K&L grade was not a significant predictor for any of the phenotypes. The predictors for 'early', 'late', and 'no progression' phenotype generally had an effect opposite to the effect of the predictors for the 'severe' and 'bone density' phenotype.

3B 'Bone density' involvement								
	Univariate			Multivariate				
Variable	**OR**	**(95%CI)**	**p**	**Variable**	**OR**	**(95%CI)**	**p**	**AUC**
Demographic & clinical					**Demographic & clinical**			0.66
Age	0.97	(0.93-1.02)	0.26					
Female gender	0.19	(0.11-0.32)	<0.001	Female gender	0.17	(0.10-0.31)	<0.001	
BMI	1.13	(1.06-1.20)	<0.001	BMI	1.14	(1.07-1.21)	<0.001	
ESR	0.96	(0.93-0.99)	0.03					
Pain intensity	1.06	(0.95-1.18)	0.33					
Radiographic				**Radiographic (demographic & clinical)**				
Feature sum				***Feature sum***				0.91
Minimum JSW	0.98	(0.88-1.09)	0.67					
Medial JSW	1.25	(1.09-1.41)	0.01	Medial JSW	1.37	(1.13-1.65)	0.001	
Lateral JSW	1.40	(1.26-1.57)	<0.001					
Varus angle	1.15	(1.06-1.24)	<0.001	Varus angle	1.14	(1.02-1.28)	0.02	
Osteophyte	3.38	(2.23-5.14)	<0.001	Osteophyte	2.82	(1.64-4.82)	<0.001	
Eminence	1.11	(1.06-1.16)	<0.001	Eminence	1.09	(1.02-1.16)	0.007	
Bone density	1.15	(1.11-1.19)	<0.001	Bone density	1.16	(1.12-1.21)	<0.001	
Feature abs d								
Minimum JSW	1.20	(0.84-1.71)	0.31					
Medial JSW	0.95	(0.60-1.51)	0.84					
Lateral JSW	1.26	(0.98-1.61)	0.07					
Varus angle	0.90	(0.76-1.07)	0.22					
Osteophyte	2.71	(1.40-5.26)	0.003					
Eminence	0.87	(0.72-1.05)	0.15					
Bone density	1.22	(1.07-1.39)	0.003					
				Demographic & clinical				
				Female gender	0.45	(0.20-0.98)	0.04	
K&L sum	1.12	(0.88-1.43)	0.36					
K&L abs d	1.11	(0.63-1.95)	0.71					

3C 'Early' progression								
	Univariate			Multivariate				
Variable	**OR**	**(95%CI)**	**p**	**Variable**	**OR**	**(95%CI)**	**p**	**AUC**
Demographic & clinical				**Demographic & clinical**				
Age	0.99	(0.94-1.04)	0.68					
Female gender	1.11	(0.61-1.99)	0.74					
BMI	0.96	(0.91-1.03)	0.26					
ESR	1.01	(0.98-1.04)	0.44					
Pain intensity	0.92	(0.82-1.04)	0.17					
Radiographic				**Radiographic (demographic & clinical)**				
Feature sum				***Feature sum***				0.79
Minimum JSW	1.13	(1.01-1.28)	0.04					
Medial JSW	0.98	(0.86-1.12)	0.80					
Lateral JSW	0.76	(0.68-0.85)	<0.001					
Varus angle	0.80	(0.73-0.88)	<0.001	Varus angle	0.85	(0.77-0.95)	0.003	
Osteophyte	0.54	(0.38-0.76)	<0.001					
Eminence	0.92	(0.88-0.97)	0.001	Eminence	0.93	(0.89-0.98)	0.006	
Bone density	0.92	(0.90-0.95)	<0.001	Bone density	0.94	(0.91-0.97)	<0.001	
Feature abs d								
Minimum JSW	0.67	(0.43-1.04)	0.07					
Medial JSW	0.82	(0.50-1.35)	0.43					
Lateral JSW	0.56	(0.37-0.84)	0.006					
Varus angle	0.64	(0.49-0.84)	0.001	Varus angle	0.75	(0.57-0.98)	0.04	
Osteophyte	0.58	(0.29-1.17)	0.13					
Eminence	0.81	(0.66-0.99)	0.05					
Bone density	0.75	(0.62-0.91)	0.004	Bone density	0.78	(0.63-0.98)	0.03	
K&L sum	0.73	(0.54-1.00)	0.05					
K&L abs d	0.54	(0.28-1.05)	0.07					

Female gender and BMI were associated with the 'bone density' phenotype together with multiple radiographic features resulting in a model with 'excellent' discriminative ability (table 3B): AUC-ROC=0.91 (95%CI: 0.88-0.94) decreasing to 'good' with AUC-ROC=0.87 (0.83-0.91) after correction for over-fitting and rounding of coefficients. The PPV, the chance of belonging to the 'bone density' phenotype, was 83%in individuals with a score above the cut-off for optimal sensitivity (table 4).
The 'early' progression phenotype was associated with radiographic features only: table 3C. AUC-ROC of this model was 0.79 (0.74-0.84) and decreased to 0.70 (0.64-0.76) after shrinkage and rounding.

3D 'Late' progression								
	Univariate			**Multivariate**				
Variable	**OR**	**(95%CI)**	**p**	**Variable**	**OR**	**(95%CI)**	**p**	**AUC**
Demographic & clinical				**Demographic & clinical**				
Age	0.98	(0.92-1.04)	0.45					
Female gender	7.13	(1.08-30.17)	0.008	Female gender	7.13	(1.08-30.17)	0.008	0.60
BMI	0.94	(0.86-1.02)	0.15					
ESR	1.00	(0.96-1.04)	0.96					
Pain intensity	1.02	(0.88-1.19)	0.79					
Radiographic				**Radiographic (demographic & clinical)**				
Feature sum				***Feature sum***				0.76
Minimum JSW	1.08	(0.93-1.26)	0.30					
Medial JSW	0.92	(0.77-1.09)	0.33					
Lateral JSW	0.77	(0.67-0.89)	<0.001	Lateral JSW	0.83	(0.71-0.98)	0.02	
Varus angle	0.85	(0.77-0.95)	0.005					
Osteophyte	0.58	(0.38-0.89)	0.01					
Eminence	0.88	(0.83-0.94)	<0.001	Eminence	0.91	(0.85-0.97)	0.004	
Bone density	0.94	(0.91-0.97)	0.001	Bone density	0.95	(0.92-0.99)	0.006	
Feature abs d								
Minimum JSW	0.69	(0.38-1.25)	0.22					
Medial JSW	0.96	(0.51-1.79)	0.89					
Lateral JSW	0.78	(0.50-1.22)	0.27					
Varus angle	0.92	(0.71-1.20)	0.55					
Osteophyte	0.46	(0.18-1.17)	0.10					
Eminence	0.88	(0.68-1.14)	0.34					
Bone density	1.01	(0.85-1.20)	0.89					
K&L sum	0.77	(0.51-1.18)	0.23					
K&L abs d	0.82	(0.36-1.86)	0.63					

Female gender and several radiographic features were associated with the 'late' progression phenotype (table 3D) and AUC-ROC was 0.76 (0.69-0.83) and remained unchanged.
Females and participants with lower BMI were more likely to belong to the 'no progression' phenotype and several radiographic features were also associated with this phenotype. Unexpectedly, individuals with a larger varus angle were more likely to belong to the 'no radiographic' progression phenotype (table 3E). The discriminative ability of the model was fair with AUC-ROC=0.72 (0.66-0.78) decreasing to 0.68 (0.62-0.74).
Table 4 shows PPV for the different cut-offs for the predictive scores per phenotype.

3E 'No progression'								
	Univariate			Multivariate				
Variable	OR	(95%CI)	p	Variable	OR	(95%CI)	p	AUC
Demographic & clinical				**Demographic & clinical** 0.68				
Age	1.04	(0.99-1.09)	0.15					
Female gender	3.92	(1.80-8.52)	0.001	Female gender	3.87	(1.75-8.54)	0.001	
BMI	0.87	(0.81-0.94)	<0.001	BMI	0.87	(0.81-0.94)	<0.001	
ESR	1.03	(1.00-1.06)	0.10					
Pain intensity	0.97	(0.86-1.08)	0.55					
Radiographic				**Radiographic (demographic & clinical)**				
Feature sum				***Feature sum***				0.72
Minimum JSW	1.00	(0.89-1.11)	0.96					
Medial JSW	0.98	(0.86-1.11)	0.74					
Lateral JSW	1.03	(0.93-1.13)	0.59					
Varus angle	1.05	(0.97-1.13)	0.24	Varus angle	1.09	(1.00-1.19)	0.04	
Osteophyte	0.60	(0.43-0.84)	0.003	Osteophyte	0.60	(0.41-0.87)	0.007	
Eminence	0.99	(0.96-1.04)	0.80					
Bone density	0.97	(0.95-0.99)	0.009					
Feature abs d				***Feature abs d***				
Minimum JSW	0.90	(0.61-1.33)	0.59					
Medial JSW	0.63	(0.36-1.10)	0.10					
Lateral JSW	1.03	(0.19-1.35)	0.82					
Varus angle	1.00	(0.83-1.19)	0.97					
Osteophyte	0.41	(0.20-0.83)	0.01					
Eminence	1.22	(1.03-1.45)	0.03	Eminence	1.25	(1.03-1.51)	0.02	
Bone density	0.99	(0.86-1.13)	0.84					
				Demographic & clinical				
				Age	1.06	(1.00-1.11)	0.04	
				Female gender	3.81	(1.69-8.62)	0.001	
				BMI	0.88	(0.82-0.95)	0.001	
K&L sum	0.86	(0.65-1.14)	0.29					
K&L abs d	1.39	(0.78-2.47)	0.27					

NA: not applicable, abs d: absolute difference, osteophyte: osteophyte area, eminence: eminence height, OR: odds ratio, 95%CI, 95% confidence interval, p: significance level: AUC: area under receiver operator characteristic curve

Table 4 Ability to predict progression phenotypes for three cut-off points of predictive score

		Predictive score		
Phenotype	**% (n) present**	**Cut-off**	**PPV**	**% identified (>cut-off)**
'Severe'	5% (16)	NA	NA	NA
'Bone density'	29% (97)	Sn: 9.00	53%	54%
		S&S: 10.30	56%	44%
		Sp: 12.00	83%	14%
'Early'	26% (89)	Sn: -5.60	33%	77%
		S&S: -5.00	40%	50%
		Sp: -4.10	48%	10%
'Late'	13% (44)	Sn: -9.80	20%	63%
		S&S: -9.00	28%	34%
		Sp: -8.00	38%	8%
'No'	27% (90)	Sn: 0.50	34%	77%
		S&S: 1.30	35%	50%
		Sp: 2.40	51%	11%

Predictive scores were calculated as: 'bone density'; (0.3*medial JSW + 0.1*varus angle + 1*(log [osteophyte area+1])+0.05*eminence height + 0.1*bone density -0.5*gender (male is 1 and female is 2)), 'early'; (-0.1*varus angle -0.05*eminence height -0.05*bone density -0.3*absolute difference in varus angle -0.2*absolute difference in bone density), 'late'; (-0.2*lateral JSW -0.1*eminence height -0.05*bone density), 'no'; (-0.1*varus angle -0.5*(log[osteophyte area+1]) + 0.2*absolute difference in eminence height + 0.05*age + 1*gender -0.1*BMI). NA: not applicable, for cut-off points: Sn; optimal sensitivity, S&S; optimal trade-off between sensitivity and specificity, Sp: optimal specificity, PPV: positive predictive value, NPV: negative predictive value

Discussion

This is the first study to identify specific phenotypes of progression of radiographic knee OA in participants with complaints of early OA. Phenotypes were found to represent the level of disease progression ('severe' and 'no progression'), the phase of progression ('early' and 'late'), and the involvement of a specific feature ('bone density'). Although the definition of the phenotypes should be confirmed, these phenotypes might represent a (partly) different etiology. The phenotypes may benefit from different treatment strategies, e.g. an intense regimen that combines pain medication with (cartilage-safe) non-steroidal anti-inflammatory drugs in case of the 'severe' phenotype, and treatment aimed at bone quality (e.g. bisphosphonates) in case of the 'bone density' phenotype. Clinical characteristics were slightly different between the clusters and also the WOMAC scores were only slightly lower in the 'no' cluster than in the other clusters. This is in line with the limited relation between clinical and radiographic OA in earlier studies[5,7,25].

The fact that we were able to identify specific phenotypes of OA progression using detailed KIDA measurement justifies the ongoing developments on more precise evaluation of plain radiographs[26]. The finding that varus alignment is a predictor for progression of OA[27], implies that this radiographic feature should be measured separately. Adding specific separate radiographic features to demographic and clinical characteristics also improved the ability to discriminate between the progression phenotypes substantially, contrary to K&L grading of overall damage. Applying measurement of specific separate radiographic features in clinical trials is therefore recommended.

Female gender[1] and BMI[28,29] are known risk factors for onset and progression of OA, and were also identified as predictors for most (but not all) phenotypes of radiographic progression in this study. Interestingly, being female was protective of belonging to the 'bone density' phenotype, and was a significant predictor (OR=3.87) for belonging to the 'no progression' phenotype. This might be related to the fact that females have lower bone density than males[30]. Osteophyte area was identified as the most important predictor for 'severe' progression and 'bone density' involvement and was protective for the 'no progression' phenotype, which might be in accordance with the fact that osteophyte formation is assumed to occur early in the disease[16]. Unexpectedly however, osteophyte area was not identified as a predictor for the 'early' phenotype. The radiographic features that were identified to be associated with the 'early' and 'late' progression phenotype (e.g. eminence height, bone density, varus angle, and JSW) actually had a protective effect, which needs further evaluation.

Generally the positive predictive values (PPV) based on the predictive scores using demographic and clinical characteristics combined with specific radiographic features were not high enough for prediction at the individual level. However, defining subgroups for inclusion in clinical trials might be significantly improved based on these scores and hence enable the development of a more personalized treatment approach. For instance, 54% of

our population could be classified as belonging to the 'bone density' phenotype with a certainty of 53% (PPV) when the predictive score was >9.0.

Cluster analysis is a technique to group individuals who are 'similar' regarding the variables that are included in the analysis. To come to a set of phenotypes, also 'subjective' choices have to be made. The value of clustering individuals is determined by the relevance and characteristics of the clusters, in our case underlying etiology, disease severity, need for treatment, and (long-term) outcome. Performing a cluster analysis with a different set of variables, for instance including clinical characteristics, might result in different clusters e.g. phenotypes in which radiographic and clinical characteristics are strongly related to each other.

In the present study, cluster analysis aimed at identifying radiographic progression phenotypes by exploring radiographic features at and between different time points. We also explicitly choose to cluster participants and not knees. When performing cluster analysis with radiographic features at T0, T2y, and T5y separately, also a 'severe' cluster with involvement of all feature scores, a cluster with 'bone density' involvement, and a cluster with 'no progression' of all feature scores was identified, which adds to the validity of the defined progression clusters. Interestingly, no clusters were identified with specific progression of e.g. one knee, or of the medial compartment and not of the lateral compartment. This might be explained by the fact that radiographic features within an individual and within a joint are quite similar and small differences are missed due to much larger differences between individuals or knees[22]. Also, this might be a reflection of the systemic character of OA[31], affecting the whole joint and also more joints within an individual[2]. This might also be the reason that the difference scores of the radiographic features were not related to phenotype 'membership'.

Based on separate radiographic features, phenotypes with different levels and phases of progression, and involvement of a specific feature were detected in participants with early complaints related to OA. These phenotypes might represent relevant subgroups for the evaluation of (preventive) treatment strategies in clinical trials and with that drive the discovery of more targeted treatment strategies.

Acknowledgements

We acknowledge Marja Lafeber (ML) for the KIDA measurements.

References

1. Felson DT, Zhang Y. An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. Arthritis Rheum. 1998 Aug;41(8):1343-55.
2. Hunter DJ, Felson DT. Osteoarthritis. Bmj. 2006 Mar 18;332(7542):639-42.
3. McAlindon T, Dieppe P. Osteoarthritis: definitions and criteria. Ann Rheum Dis. 1989 Jul;48(7):531-2.
4. Peat G, Greig J, Wood L, et al. Diagnostic discordance: we cannot agree when to call knee pain 'osteoarthritis'. Fam Pract. 2005 Feb;22(1):96-102.
5. Dieppe PA, Cushnaghan J, Shepstone L. The Bristol 'OA500' study: progression of osteoarthritis (OA) over 3 years and the relationship between clinical and radiographic changes at the knee joint. Osteoarthritis Cartilage. 1997 Mar;5(2):87-97.
6. Hochberg MC, Lawrence RC, Everett DF, et al. Epidemiologic associations of pain in osteoarthritis of the knee: data from the National Health and Nutrition Examination Survey and the National Health and Nutrition Examination-I Epidemiologic Follow-up Survey. Semin Arthritis Rheum. 1989 May;18(4 Suppl 2):4-9.
7. Kinds MB, Welsing PM, Vignon EP, et al. A systematic review of the association between radiographic and clinical osteoarthritis of hip and knee. Osteoarthritis Cartilage. 2011 Jan 31.
8. Lethbridge-Cejku M, Scott WW, Jr., Reichle R, et al. Association of radiographic features of osteoarthritis of the knee with knee pain: data from the Baltimore Longitudinal Study of Aging. Arthritis Care Res. 1995 Sep;8(3):182-8.
9. Boegard TL, Rudling O, Petersson IF, et al. Joint space width of the tibiofemoral and of the patellofemoral joint in chronic knee pain with or without radiographic osteoarthritis: a 2-year follow-up. Osteoarthritis Cartilage. 2003 May;11(5):370-6.
10. Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: an update with relevance for clinical practice. Lancet. 2011 Jun 18;377(9783):2115-26.
11. Duryea J, Neumann G, Niu J, et al. Comparison of radiographic joint space width with magnetic resonance imaging cartilage morphometry: analysis of longitudinal data from the Osteoarthritis Initiative. Arthritis Care Res (Hoboken). 2010 Jul;62(7):932-7.
12. Kerkhof HJ, Meulenbelt I, Akune T, et al. Recommendations for standardization and phenotype definitions in genetic studies of osteoarthritis: the TREAT-OA consortium. Osteoarthritis Cartilage. 2011 Mar;19(3):254-64.
13. Banks SE. Erosive osteoarthritis: a current review of a clinical challenge. Clin Rheumatol. 2011 Jul;29(7):697-706.
14. Hunter DJ, Gerstenfeld L, Bishop G, et al. Bone marrow lesions from osteoarthritis knees are characterized by sclerotic bone that is less well mineralized. Arthritis Res Ther. 2009;11(1):R11.
15. van der Kraan PM, van den Berg WB. Osteophytes: relevance and biology. Osteoarthritis Cartilage. 2007 Mar;15(3):237-44.
16. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. 1957 Dec;16(4):494-502.
17. Ding C, Cicuttini F, Jones G. Do NSAIDs affect longitudinal changes in knee cartilage volume and knee cartilage defects in older adults? Am J Med. 2009 Sep;122(9):836-42.
18. Wesseling J, Dekker J, van den Berg WB, et al. CHECK (Cohort Hip and Cohort Knee): similarities and differences with the Osteoarthritis Initiative. Ann Rheum Dis. 2009 Sep;68(9):1413-9.
19. Buckland-Wright C. Protocols for precise radio-anatomical positioning of the tibiofemoral and patellofemoral compartments of the knee. Osteoarthritis Cartilage. 1995 Sep;3 Suppl A:71-80.
20. Buckland-Wright JC, Wolfe F, Ward RJ, et al. Substantial superiority of semiflexed (MTP) views in knee osteoarthritis: a comparative radiographic study, without fluoroscopy, of standing extended, semiflexed (MTP), and schuss views. J Rheumatol. 1999 Dec;26(12):2664-74.
21. Marijnissen AC, Vincken KL, Vos PA, et al. Knee Images Digital Analysis (KIDA): a novel method to quantify individual radiographic features of knee osteoarthritis in detail. Osteoarthritis Cartilage. 2008 Feb;16(2):234-43.
22. Kinds MB, Vincken KL, Vignon E, et al. Radiographic features of knee and hip osteoarthritis represent characteristics of an individual, in addition to severity of osteoarthritis. Scand J Rheumatol. 2011:*Accepted for publication*.
23. Muller MP, Tomlinson G, Marrie TJ, et al. Can routine laboratory tests discriminate between severe acute respiratory syndrome and other causes of community-acquired pneumonia? Clin Infect Dis. 2005 Apr 15;40(8):1079-86.
24. Van Houwelingen JC, Le Cessie S. Predictive value of statistical models. Stat Med. 1990 Nov;9(11):1303-25.
25. Bedson J, Croft PR. The discordance between clinical and radiographic knee osteoarthritis: a systematic search and summary of the literature. BMC Musculoskelet Disord. 2008;9:116.

26. Oka H, Muraki S, Akune T, et al. Fully automatic quantification of knee osteoarthritis severity on plain radiographs. Osteoarthritis Cartilage. 2008 Apr 17.
27. Brouwer GM, van Tol AW, Bergink AP, et al. Association between valgus and varus alignment and the development and progression of radiographic osteoarthritis of the knee. Arthritis Rheum. 2007 Apr;56(4):1204-11.
28. Grotle M, Hagen KB, Natvig B, et al. Obesity and osteoarthritis in knee, hip and/or hand: an epidemiological study in the general population with 10 years follow-up. BMC Musculoskelet Disord. 2008;9:132.
29. Heliovaara M, Makela M, Impivaara O, et al. Association of overweight, trauma and workload with coxarthrosis. A health survey of 7,217 persons. Acta Orthop Scand. 1993 Oct;64(5):513-8.
30. Breijawi N, Eckardt A, Pitton MB, et al. Bone Mineral Density and Vitamin D Status in Female and Male Patients with Osteoarthritis of the Knee or Hip. Eur Surg Res. 2008 Oct 30;42(1):1-10.
31. Kraus VB, Kepler TB, Stabler T, et al. First qualification study of serum biomarkers as indicators of total body burden of osteoarthritis. PLoS One. 2010;5(3):e9739.

Chapter 8

Influence of variation in knee positioning during image acquisition on separate quantitative radiographic parameters of osteoarthritis

MB Kinds[1,2]
KL Vincken[2]
TN Hoppinga[2]
RLAW Bleys[3]
MA Viergever[2]
ACA Marijnissen[1]
PMJ Welsing[1,4]
FPJG Lafeber[1]

[1]Rheumatology & Clinical Immunology, University Medical Center Utrecht
[2]Image Sciences Institute, University Medical Center Utrecht
[3]Anatomy, University Medical Center Utrecht
[4]Julius Center for Health Sciences & Primary Care, University Medical Center Utrecht

Submitted for publication

Abstract

Objective

The clinical application of quantitative (digital) measurement of separate radiographic parameters of knee osteoarthritis (OA) might be hampered by a lack of reproducible joint positioning during acquisition of the radiographs. The influence of systematic variations in positioning of the knee, on measurement of separate quantitative radiographic parameters, was studied.

Methods

Five components of knee position during radiographic acquisition (beam height, lower and upper leg extension, internal rotation, and lateral shift) were systematically varied within a clinically relevant range, using three cadaver legs. The influence of these variations on the change in measurement of the separate quantitative radiographic parameters was evaluated. Significant changes were validated in vivo. Changes were compared with differences during two-year follow-up in a radiographic progression cohort of early OA.

Results

Systematic variation in upper and lower leg extension induced changes in the measurement of joint space width. Lower leg extension also influenced measurement of osteophyte area and eminence height. Also bone density measurement was influenced by variation in all five position components. Variations were of clinical relevance compared with two-year differences in knees with radiographic progression, and were confirmed in vivo.

Conclusion

Variations in knee positioning, which are considered to occur easily during image acquisition in trials and clinical practice despite standardization, is of significant influence on the quantitative measurement of most separate radiographic parameters of OA. The additional value of quantitative measurement might improve significantly by better standardization during acquisition of radiographs; with radiography still the being the gold standard for structure modification in OA.

Introduction

Osteoarthritis (OA) is a disabling joint disease that commonly affects large weight bearing joints like the knee. Structural changes include articular cartilage damage, osteophyte formation, and subchondral bone changes, and are assumed to (at least partly) underlie symptoms of pain and functional disability[1,2]. Radiography is still the gold standard for demonstrating structural changes in humans[3], since image acquisition is non-invasive, cheap, fast and generally accessible[4,5]. In the past decades developments have been ongoing on more detailed evaluation to improve sensitivity for detection of structural damage on radiographs. The actual measurement (on a continuous scale) of joint space width (JSW) is increasingly applied[6,7], and in recent years digital image analysis tools were developed to increase efficiency and reliability of such measurements[6,8,9]. Compared with the commonly used Kellgren & Lawrence[10] (K&L) grading, the sensitivity to change was improved by actual measurement of JSW[11]. Digital analysis also enables measurement of additional separate radiographic characteristics of knee OA like osteophyte formation[12] and joint angulation[12,13], and even bone density[14]. The measurement of separate OA parameters might improve the detection of structural damage in an early phase of the disease and might improve the association with clinical symptoms.

When the onset and progression of radiographic OA is evaluated, changes caused by variation in knee positioning during acquisition of the subsequent radiographs need to be taken into account. Such variation specifically hampers the detection of differences (over time and between individuals) when the radiographic development is subtle, which is generally the case in a slowly progressive disease like OA and specifically early in the disease. This also accounts for radiographic changes in e.g. hips and hands[15-17].

Particularly when using digital image analysis, reproducible positioning of the knee needs attention, since this objective mathematical method does not take into account subjective evaluation of variation in knee positioning between radiographs. The acquired image is a two-dimensional projection of the knee and is determined by the three-dimensional orientation of the joint towards the X-ray device. It has been reported that variability in JSW measurement is introduced by variations in knee flexion, foot rotation, and beam angle in the extended radiographic view[18], and by variations in beam height in the tunnel view[19]. Therefore, standardization of the radiographic procedure is of great importance. Standardization is commonly aimed at reproducible alignment of the medial tibial plateau, by projection of the anterior rim on the posterior rim. This can be achieved by applying some degree of knee flexion and by inclination of the beam angle. Since verification by fluoroscopy results in increased costs, acquisition time and X-ray exposure, several non-fluoroscopic procedures have been evaluated for the reproducibility of knee joint positioning and the influence on JSW measurement[20-22]. The semiflexed view according to Buckland-Wright is preferred since this procedure repositions the joint best, both at the same day[20] and within a year[23]. Even in the case of unsatisfactory alignment, which was reported in 70% of initial examinations, the position of the baseline image was highly reproducible in the

follow-up image and as such reproducibility of the JSW measurement approached fluoroscopic procedures[24].
Despite this standardization, acquiring reproducible radiographs in clinical studies remains difficult. Interestingly, it has never been reported to what extent specific components of the knee position (e.g. flexion or rotation) influence the reproducibility of radiographic characteristics other than JSW, like osteophyte area, eminence height, and bone density. The present study evaluated which systematic variations in positioning of the knee towards the X-ray detector have an effect on measurements of separate radiographic parameters, which is of relevance in the evaluation of structural differences in the process of OA in clinical studies and in clinical practice.

Methods

Cadaver study

The optimal approach to evaluate the effect of systematic variations on the measurement of OA severity is to take multiple weight bearing radiographs of an individual. However, cumulative X-ray exposure makes this ethically impossible. Another option would be to use computed tomography of a knee under different flexion/extension angles, and to generate projection images under different angles. But next to undesired X-ray exposure, a major drawback of computed tomography is that this is under non weight bearing conditions. Owing to these constraints the use of cadaver knee joints was considered the most feasible and valid method. Three human cadaver legs (two females and one male: age 76, 76, and 65 years) were prepared for analysis. Although incidence of OA is highest at this age (>65 years), the cadaver legs were not reported as suffering from clinical OA symptoms. These legs were considered suitable for the present evaluation since more and more cohorts focus on very early (pre-clinical) OA (e.g. the Cohort Hip & Cohort Knee; CHECK[25] study in the Netherlands and the Osteoarthritis Initiative; OAI[26] in the United States).
To warrant the mechanical condition of the knee joint as good as possible (like weight bearing, and ligament/muscle involvement) while keeping the experimental conditions feasible/acceptable, the whole leg including the hip joint was used with the lumbar vertebrae fixed to a framework. This set-up allowed (semi-) weight bearing during radiography with the possibility to image multiple systematic variations in knee positioning.

Validation in vivo

The significant changes in radiographic parameters by variations in positioning in the cadaver legs were validated in vivo. For each component of knee position two radiographs were acquired of a healthy volunteer, representing the two extremes of the variation in the component of leg positioning. The volunteers (four males and one female) were aged between 50 and 66 years and had no known history of joint disease. The medical ethical

committee of the University Medical Center Utrecht approved this study, and volunteers gave written informed consent.

Reference cohort

To evaluate if the changes observed during variation in knee positioning were of clinical relevance, they were compared with differences observed in a radiographic progression cohort of individuals with early OA. Knees were selected from the CHECK cohort with no radiographic OA (K&L grade <II) at baseline and radiographic OA (K&L grade ≥II) at five-year follow-up (310 knees, mean age at baseline: 56 years [range: 44-66]). In this cohort knee radiographs are taken according to a standardized protocol. The mean differences from baseline to two-year follow-up in the separate radiographic parameters of these knees were used as a clinical reference.

Radiographic parameters

Posteroanterior radiographs were acquired according to the standardized protocol by Buckland-Wright[20,27] using a clinical digital radiography system (*Digital Diagnost*, Philips Healthcare, Best, the Netherlands) at the University Medical Center Utrecht. Acquisition settings were: tube potential of 55 kilo voltage (kV), tube charge of 5 milliampere seconds (mAs), no added tube filtration, and a source image distance of 120 cm. The radiographic protocol and the acquisition settings were according to those used in the CHECK[25] cohort).
Separate radiographic parameters were quantitatively measured by use of Knee Images Digital Analysis (KIDA)[14], an interactive tool to evaluate radiographic characteristics with low inter- and intra-observer variation[14,28]. All radiographs were analyzed by one experienced observer (ML) in random order. Analyses revealed minimum, medial, and lateral JSW (joint space width in mm), the angle between the femur and tibia in the frontal plane (varus angle in degrees), height of the eminences (mm), osteophyte area (in mm^2), in four compartments (lateral and medial femur and tibia), and bone density of the four compartments (expressed in mmAl equivalents)[14].

Knee position

For radiographs in the standard position (Buckland-Wright) the leg was placed in the semiflexed position with the knee leaning against the detector, the first metatarsophalangeal (MTP) joint perpendicular to the detector, and the foot in 7.5 degrees exorotation (by use of a foot plate with a triangular wedge)[20,27]. Compared with the standard position, five separate components of knee positioning were systematically varied (figure 1).
The choice and the range of variation of the position components were based on expert opinion (FL/AM). The knee position was varied by systematically changing one position component while the rest of the leg remained in the standard position. Variations were done in both directions (e.g. more extension (+) and more flexion (-) compared with standard position). Per component the position was varied in fixed steps (one radiograph per step) that were of similar size for all cadaver legs and which were considered to be in a clinically

relevant range. ‘Beam height’, ‘lower leg extension’ (shifting the foot forwards/backwards on the panel), and ‘lateral shift upper leg’ (frontal plane) were changed in steps of 1 cm. ‘Internal rotation’ (transversal plane) and ‘upper leg extension’ was changed in steps of approximately 5 degrees as measured by a goniometer. After radiographic acquisition the measurement of ‘upper leg extension’ was verified by angle measurement on standard photographs, which were taken from a lateral view simultaneously with the radiographs. On these photographs the knee extension angle was measured between bars that were fixed to the bone (by pins through soft tissue) on the lateral side of the femur and of the tibia.

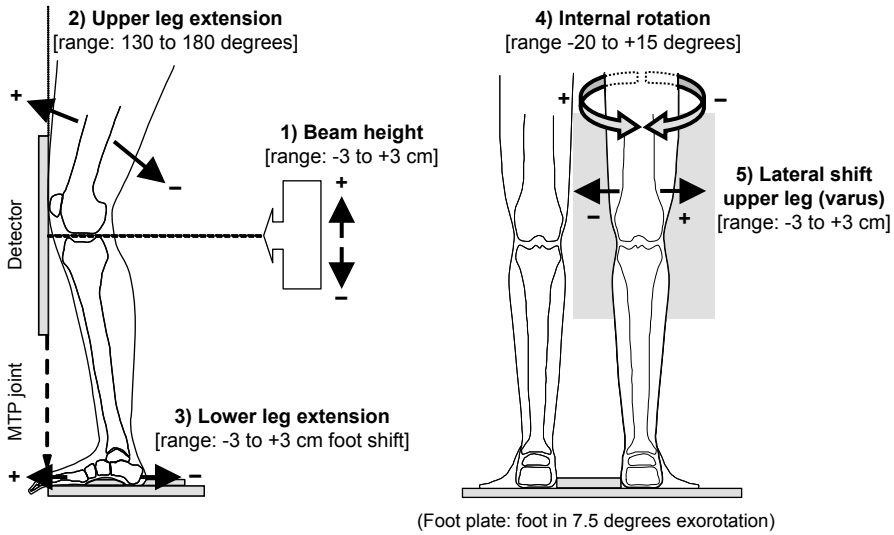

Figure 1 Standard position and systematic variations in knee position components for radiography, MTP: metatarsophalangeal

Statistical analyses

Linear regression analyses were performed to evaluate whether variations in knee positioning (e.g. ‘beam height’) by stepwise radiographs, induced a systematic effect on the measurement of the separate quantitative parameters (KIDA) in the cadaver legs. This analysis accounted for the dependency of repeated observations within the same cadaver leg (mixed model analyses). Results were only reported when the effect was statistically significant, and when the regression coefficient of the three individual cadaver legs (using ordinary linear regression) had the same direction and p-value was <0.10 for at least two of the three cadaver legs. Those KIDA parameters that were influenced by variations in knee positioning in the cadaver joints were verified in vivo.

The clinical relevance of the influence of systematic variation was evaluated by comparing the change in radiographic parameters per unit increase (regression coefficients) with the differences observed over two years in knees with radiographic progression.
Analyses were performed using SPSS (Statistical Package for the Social Sciences) version 15.0 and SAS (Statistical Analysis Software) version 9.1.3.

Table 1 Regression analyses: influence of systematic variation in knee position components on separate quantitative radiographic parameters

	Beam (cm)		Upper ext (5dgr)		Lower ext (cm)		Internal (5dgr)		Lateral shift (cm)	
	ß	p	ß	p	ß	p	ß	p	ß	p
Joint space										
Minimum (mm)	-	-	-	-	+0.07	0.01	-	-	-	-
Medial (mm)	-	-	-	-	+0.07	0.005	-	-	-	-
Lateral (mm)	-	-	-0.20	<0.0001	-0.18	0.001	-	-	-	-
Varus angle (degrees)	-0.12	0.004	-0.25	<0.0001	-0.32	<0.0001	-	-	-	-
Osteophyte area										
Femur lateral (mm^2)	-	-	-	-	-	-	-	-	-	-
Tibia lateral (mm^2)	-	-	-	-	-	-	-	-	-	-
Tibia medial (mm^2)	-	-	-	-	+0.49	0.0001	-	-	-	-
Eminence height										
Lateral (mm)	-	-	-	-	-0.23	0.001	-	-	-	-
Medial (mm)	-	-	-	-	-0.17	0.002	-	-	-	-
Bone density										
Femur lateral (mmAl)	-0.31	0.001	+0.38	<0.0001	-	-	+0.48	0.0004	+0.79	<0.0001
Femur medial (mmAl)	-	-	-	-	-	-	+0.28	<0.0001	+0.47	0.0002
Tibia lateral (mmAl)	-	-	+0.38	<0.0001	+0.29	<0.0001	+0.29	0.005	+0.73	<0.0001
Tibia medial (mmAl)	-	-	+0.29	0.0002	-	-	-	-	+0.61	<0.0001

Beam: beam height, upper: upper leg, ext: extension, lower: lower leg, internal: internal rotation, lateral shift: of upper leg, dgr: degrees, ß: regression coefficient per cm or 5 dgr, - : mixed regression analyses; not significant and individual regression analyses; not the same direction for all three cadaver legs and/or only one cadaver leg p-value <0.10

Results

Influence of variation in knee positioning on radiographic parameters

In table 1 per component of knee position that was varied regression coefficients (ß) and p-values are provided which represent the change in outcome per unit change in positioning, in case of significance according to the above-described criteria.

Joint space

Varying the 'beam height' (-3 to +3 cm) did not influence minimum JSW, medial JSW, and lateral JSW. Increasing the beam height induced a significant decrease in varus angle (-0.12 degrees varus angle per cm beam height; table 1). The effect is not considered clinically relevant, since the change in the cadaver legs was considerably smaller than the mean difference (increase) of 0.77 degrees in knees with radiographic progression (table 2).
Systematically varying the 'upper leg extension' (130 to 180 degrees) significantly influenced the lateral JSW and varus angle (-0.20 mm and -0.25 degrees per 5 degrees more extension; table 1 and figure 2A). The decrease in lateral JSW is in accordance with the

decrease in varus angle, meaning a relative decrease in lateral JSW compared to medial JSW. These effects were verified in vivo, showing a decrease in lateral JSW similarly to the cadaver legs (-0.21 mm per 5 degrees) and a decrease in varus angle of -0.08 degrees per 5 degrees (smaller than cadaver legs). The variation in 'upper leg extension' on lateral JSW is considered clinically relevant, since this change was only slightly smaller than the 0.27 mm difference during two-year follow-up in individuals with early OA that progressed from K&L grade <II to ≥II (table 2).

The variation in 'lower leg extension' by shifting the foot forward (from -3 to +3 cm) on the foot plate induced a slight but systematic increase in minimum JSW and medial JSW (both +0.07 mm per cm). The decrease in lateral JSW (-0.18 mm per cm) and varus angle (-0.32 degrees per cm) is considered in accordance with the increase in medial and minimum JSW (since the medial compartment is commonly smallest). These effects were also found in vivo, and specifically the increase in minimum and medial JSW was substantial (per 5 degrees: +0.17 mm for minimum JSW, +0.24 mm for medial JSW, -0.03 mm for lateral JSW, and -0.31 degrees for varus angle). Particularly the increase in minimum JSW (+0.07 mm per cm; table 1 and figure 2B) is clinically relevant when compared with the mean difference of 0.11 mm in knees with radiographic progression. 'Internal rotation of the upper leg' (-20 to +15 degrees) and 'lateral shift of the upper leg' (-3 to +3 cm) had no clear systematic influence on the measurement of JSW and varus angle on radiographs.

Osteophyte area

Osteophyte formation in these (relatively) healthy knees was minimal, and at the medial femur of all three cadaver legs no osteophyte area was present. Only by systematically increasing 'lower leg extension' (shifting the foot forward) the osteophyte area increased significantly in the medial tibia (0.49 mm^2 per cm; table 1 and figure 2C). An increase in osteophyte area in this compartment was also found in vivo (+0.22 mm per cm). The change is of limited clinical relevance however, since this is clearly smaller than the difference over two years of 1.33 mm^2 in medial tibia osteophyte area in knees with radiographic progression (table 3). Variations in the other position components did not significantly influence the quantitative measurement of osteophyte area. Clearly, the effects may be underestimated due to the minimal osteophyte area in these healthy individuals.

Eminence height

The height of the tibial eminences was significantly influenced by variation in 'lower leg extension', not surprisingly because of their position on the tibia (table 1 and figure 2D). This effect was confirmed in vivo, with a change of similar size on the lateral eminence (per cm change -0.21 mm in vivo compared with -0.23 in cadaver) and a smaller change on the medial eminence (-0.04 mm in vivo compared with -0.17 mm in cadaver per cm). The influence on the lateral eminence measurement was considered clinically relevant since the change caused by systematic repositioning was only slightly smaller than the difference of 0.27 mm in knees with progression during two-year follow-up (table 2).

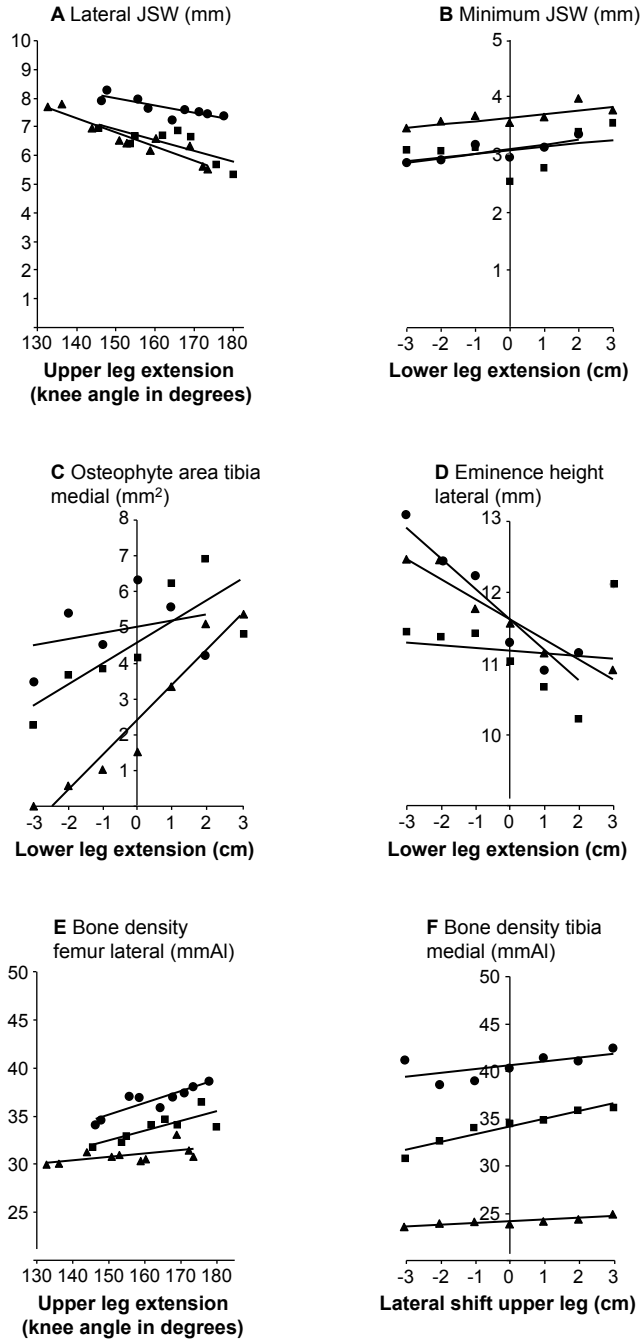

Figure 2 Influence of systematic variation in position components on separate quantitative radiographic parameters

Table 2 Difference (mean & SD) in separate quantitative radiographic parameters in CHECK

	Progression (310 knees)	
	Mean Δ	**SD Δ**
Joint space		
Minimum (mm)	-0.11	1.28
Medial (mm)	-0.29	0.78
Lateral (mm)	0.27	1.69
Varus angle (degrees)	0.77	2.33
Osteophyte area		
Femur lateral (mm^2)	0.47	3.67
Femur medial (mm^2)	0.66	3.31
Tibia lateral (mm^2)	0.78	4.38
Tibia medial (mm^2)	1.33	4.11
Eminence height		
Lateral (mm)	0.26	1.53
Medial (mm)	0.05	1.35
Bone density		
Femur lateral (mmAl)	1.46	5.86
Femur medial (mmAl)	2.34	6.26
Tibia lateral (mmAl)	1.33	6.15
Tibia medial (mmAl)	2.62	6.74

Progression: knees that changed from K&L grade <II to K&L grade ≥II, Δ: difference during two-year follow-up

Bone density

Surprisingly the bone density measurement was influenced by systematic variations in many of the components of knee positioning. By increasing 'upper leg extension', bone density in the lateral femur increased with 0.38 mmAl per 5 degrees (table 1 and figure 2E). The effect of 'upper leg extension' on the lateral femur corroborates the effect of 'beam height' (-0.31 mmAl per cm). Since increasing 'beam height' artificially causes an increase in knee flexion angle, this resulted in decreased bone density. Varying 'beam height' did not significantly influence bone density in the other compartments. Surprisingly, although the position of the tibia was not changed by 'upper leg extension', also tibial bone density was influenced to a similar extent as in the lateral femur (lateral tibia: 0.38 mmAl and medial tibia: 0.29 mmAl per 5 degrees).

Varying 'lower leg extension' resulted in a significant increase in lateral tibia bone density (0.29 mmAl per cm). No influence on the femur was observed, which is in line with the fact that the femur is not actually varied in position by changing the 'lower leg extension'.

Varying 'internal rotation' resulted in a limited increase in bone density measurement in the lateral and medial femur and lateral tibia. This effect corroborates the bone density increase due to 'upper leg extension'. During radiographic procedures (closed chain movement due to fixed foot position) the femur rotates internally and medially (varus) at the last 30 degrees of extension[29]. Similarly, this effect fits the significant increase in bone density in the femur by variation in 'lateral shift upper leg'. Moreover, it supports relatively normal joint kinematics in the cadaver legs. Unexpectedly, but in accordance with 'upper leg extension' and 'internal

rotation', 'lateral shift upper leg' induced a bone density increase in the lateral tibial compartment and also in de medial tibial compartment (figure 2F).
In general, the influence of systematic variation in positioning on bone density measurements was considered of minor clinical relevance since the changes due to systematic repositioning were all smaller than the differences in knees with radiographic progression (table 2). For lateral femur and tibia however, changes in bone density over two years in the knees with radiographic progression (1.46 and 1.33 mmAl in two years) was only twice that of the systematic variation induced by 1 cm lateral shift (0.79 and 0.73 mmAl, respectively). Surprisingly, when evaluating variation in lateral shift in vivo, the change was in the opposite direction but substantially smaller (-0.07 and -0.13 mmAl, respectively).

Discussion

Systematic variation in knee joint positioning during image acquisition, and particularly in the extension angle, influenced the quantitative measurement of different radiographic parameters in this study. These clinically relevant changes were confirmed by in vivo evaluation. Several of these changes were relevant compared with the detected differences during two-year follow-up in knees with radiographic OA progression early in the disease.
The clinical relevance of the influence of knee positioning on JSW measurement is confirmed by the commonly reported annual progression rate of joint space narrowing due to OA of around 0.2 mm[30,31]. Specifically very early in the disease narrowing is probably even less, as shown by the difference of 0.11 mm for minimum JSW and 0.29 mm for medial JSW during two years of follow-up, in knees with K&L grade progression in the CHECK cohort. Even subtle variations in 'upper leg extension' (5 degrees) and 'lower leg extension' (1 cm shift) influenced the medial JSW measurement with 0.07 mm in the cadaver legs and even 0.24 mm in vivo. Although the Buckland-Wright protocol aims at medial tibia plateau alignment[20,23], these changes were only slightly smaller than the two-year differences in progressive radiographic OA. Less well known is the influence of positioning on lateral JSW and varus angle measurement. Clearly also these parameters are influenced by variations in knee extension (by 'lower' and 'upper leg extension') both in cadaver legs and in vivo. To improve the additional value of digital analyses this needs further attention, as JSW is a commonly applied outcome to evaluate radiographic knee OA[6,32].
As expected, the measurement of eminence height is only significantly influenced by 'lower leg extension'. From a mathematical point of view, only 'beam height' might have been of influence additionally. By varying the height of the X-ray beam the eminence height was expected to decrease when shifting up (+) but also when shifting down (-) compared with the standard position, which indeed occurred (data not shown). Although the role of the eminences in the OA process is argued, recent studies (in CHECK) have demonstrated clear progression in eminence height during follow-up (manuscript submitted) and a predictive value of this parameter for progression of disease (manuscript submitted).

In the cadaver legs and healthy volunteers osteophyte area was hardly present as a characteristic of OA. Irrespectively, small changes in osteophyte area are considered to be of relevance since osteophyte formation is assumed to occur first when OA develops, as defined in the commonly used K&L grading[10]. When separate features were measured, the formation of osteophytes was found to be important, e.g. in predicting incidence of radiographic OA (manuscript submitted) and in predicting phenotypes of radiographic knee OA progression (manuscript submitted). As for eminence height and JSW, specifically 'lower leg extension' influenced osteophyte area measurements. Although the two-year difference in osteophyte area in knees with progressive radiographic OA exceeds the influence of 1 cm variation in positioning, this difference will be smaller than the change when a shift in positioning of 2 cm is applied.

The influence of variations in all components of knee position on bone density measurement is of interest. Although the effect was smaller than differences during two-year follow-up in case of radiographic progression, this effect should not be underestimated. Slight variations in positioning may alter projection of compact bone areas, which results in significant changes in bone density measurement. On the other hand, the observed changes may also be due to the use of digital image acquisition (in contrast to conventional film-screen acquisition) in which the appearance of the image is influenced by variable automated adjustments of contrast and noise. When the leg is positioned differently, this can influence the projected gray values of the bone and with that the post-processing[33].

It can be argued that the use of cadaver legs is not representative of clinical practice. Although the set-up was optimized by use of a whole leg and a frame, which aimed at fixation and similar weight bearing in all radiographs, knee positioning might be hampered by e.g. the lack of muscle tension. The validation in vivo however, confirmed the observed changes. It should be noted though that only three cadaver joints were imaged and only one validation for each of the extreme variations in positioning was used. It is therefore of relevance to validate the data from the present study, to demonstrate that accurate positioning improves reliability of quantitative analysis of radiographs.

Since strict criteria were applied to distinguish between actual effects owing to systematic variations in knee positioning and random effects (ß's of three cadaver legs in same direction and two $p<0.10$), the number of clinically relevant effects was limited. Additionally, the comparison with two-year differences in an early OA cohort with clear radiographic progression in the knees might have underestimated the clinical relevance of the observed changes. In clinical studies radiographic progression is preferably evaluated already after one year, and not all individuals will progress in radiographic OA severity. As such, it is concluded that the presently identified position variations that influence radiographic analyses are the most relevant but probably not the only one. Moreover, in clinical practice the influence might actually be larger since, when a patient is positioned for a radiograph, a combination of small variations in different position components likely occurs instead of a variation in only one component.

Besides technical limitations, variations in knee positioning might be introduced by actual progression of pain or structural damage. Pain may result in limitations in knee movement[34], forcing different knee positioning during image acquisition. Also weight bearing on the affected leg might be limited due to pain, which can influence the width of the projected joint space.

Clearly the influence of knee positioning exceeds the variation in digital image analysis techniques such as KIDA, since intra-observer variation was low with these measurements[14,28]. As such, to benefit from very robust analyses methods like KIDA, optimal attention of the technicians involved in image acquisition is needed. But presumably better, new techniques, like the use of molds might need to be developed to improve this standardization.

The present study demonstrates that variations in knee positioning, which can easily occur during acquisition in trials and clinical practice despite standardization, significantly influence the quantitative measurement of most separate radiographic characteristics of osteoarthritis. Although the parameters measured by digital analysis are sufficiently robust, the surplus value of these quantitative measurements over qualitative grading will pay off only when standardization during image acquisition is improved. Since radiography remains cheap and easily accessible, it is considered of value to further improve standardization of acquisition.

Acknowledgements

We acknowledge Willem van Wolferen from the department of Anatomy of the University Medical Center Utrecht for his assistance with the preparation of the cadaver legs and the experimental set-up. We thank the healthy volunteers. We acknowledge the Radiology department of the University Medical Center Utrecht for their cooperation with the acquisition of the radiographs. And we acknowledge Marja Lafeber (ML) for measurements of the radiographs using KIDA.

References

1. McAlindon T, Dieppe P. Osteoarthritis: definitions and criteria. Ann Rheum Dis. 1989 Jul;48(7):531-2.
2. Peat G, Thomas E, Duncan R, et al. Clinical classification criteria for knee osteoarthritis: performance in the general population and primary care. Ann Rheum Dis. 2006 Oct;65(10):1363-7.
3. Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: an update with relevance for clinical practice. Lancet. 2011 Jun 18;377(9783):2115-26.
4. Buckland-Wright JC. Quantitative radiography of osteoarthritis. Ann Rheum Dis. 1994 Apr;53(4):268-75.
5. Nevitt MC. Definition of hip osteoarthritis for epidemiological studies. Ann Rheum Dis. 1996 Sep;55(9):652-5.
6. Buckland-Wright JC, Macfarlane DG, Williams SA, et al. Accuracy and precision of joint space width measurements in standard and macroradiographs of osteoarthritic knees. Ann Rheum Dis. 1995 Nov;54(11):872-80.
7. Ravaud P, Chastang C, Auleley GR, et al. Assessment of joint space width in patients with osteoarthritis of the knee: a comparison of 4 measuring instruments. J Rheumatol. 1996 Oct;23(10):1749-55.
8. Schmidt JE, Amrami KK, Manduca A, et al. Semi-automated digital image analysis of joint space width in knee radiographs. Skeletal Radiol. 2005 Oct;34(10):639-43.
9. Neumann G, Hunter D, Nevitt M, et al. Location specific radiographic joint space width for osteoarthritis progression. Osteoarthritis Cartilage. 2008 Nov 12.
10. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. 1957 Dec;16(4):494-502.
11. Ravaud P, Giraudeau B, Auleley GR, et al. Radiographic assessment of knee osteoarthritis: reproducibility and sensitivity to change. J Rheumatol. 1996 Oct;23(10):1756-64.
12. Oka H, Muraki S, Akune T, et al. Fully automatic quantification of knee osteoarthritis severity on plain radiographs. Osteoarthritis Cartilage. 2008 Nov;16(11):1300-6.
13. Wong AK, Inglis D, Beattie KA, et al. Reproducibility of computer-assisted joint alignment measurement in OA knee radiographs. Osteoarthritis Cartilage. 2009 May;17(5):579-85.
14. Marijnissen AC, Vincken KL, Vos PA, et al. Knee Images Digital Analysis (KIDA): a novel method to quantify individual radiographic features of knee osteoarthritis in detail. Osteoarthritis Cartilage. 2008 Feb;16(2):234-43.
15. Auleley GR, Duche A, Drape JL, et al. Measurement of joint space width in hip osteoarthritis: influence of joint positioning and radiographic procedure. Rheumatology (Oxford). 2001 Apr;40(4):414-9.
16. Vignon E, Conrozier T, Piperno M, et al. Radiographic assessment of hip and knee osteoarthritis. Recommendations: recommended guidelines. Osteoarthritis Cartilage. 1999 Jul;7(4):434-6.
17. Neumann G, dePablo P, Finckh A, et al. Patient repositioning reproducibility of joint space width measurements on hand radiographs. Arthritis Care Res (Hoboken). 2011 Feb;63(2):203-7.
18. Ravaud P, Auleley GR, Chastang C, et al. Knee joint space width measurement: an experimental study of the influence of radiographic procedure and joint positioning. Br J Rheumatol. 1996 Aug;35(8):761-6.
19. Lynch JA, Buckland-Wright JC, Macfarlane DG. Precision of joint space width measurement in knee osteoarthritis from digital image analysis of high definition macroradiographs. Osteoarthritis Cartilage. 1993 Oct;1(4):209-18.
20. Buckland-Wright JC, Wolfe F, Ward RJ, et al. Substantial superiority of semiflexed (MTP) views in knee osteoarthritis: a comparative radiographic study, without fluoroscopy, of standing extended, semiflexed (MTP), and schuss views. J Rheumatol. 1999 Dec;26(12):2664-74.
21. Buckland-Wright C. Review of the anatomical and radiological differences between fluoroscopic and non-fluoroscopic positioning of osteoarthritic knees. Osteoarthritis Cartilage. 2006;14 Suppl A:A19-31.
22. Ravaud P, Giraudeau B, Auleley GR, et al. Variability in knee radiographing: implication for definition of radiological progression in medial knee osteoarthritis. Ann Rheum Dis. 1998 Oct;57(10):624-9.
23. Buckland-Wright JC, Ward RJ, Peterfy C, et al. Reproducibility of the semiflexed (metatarsophalangeal) radiographic knee position and automated measurements of medial tibiofemoral joint space width in a multicenter clinical trial of knee osteoarthritis. J Rheumatol. 2004 Aug;31(8):1588-97.
24. Mazzuca SA, Brandt KD, Buckwalter KA, et al. Field test of the reproducibility of the semiflexed metatarsophalangeal view in repeated radiographic examinations of subjects with osteoarthritis of the knee. Arthritis Rheum. 2002 Jan;46(1):109-13.
25. Wesseling J, Dekker J, van den Berg WB, et al. CHECK (Cohort Hip and Cohort Knee): similarities and differences with the Osteoarthritis Initiative. Ann Rheum Dis. 2009 Sep;68(9):1413-9.
26. Eckstein F, Maschek S, Wirth W, et al. One year change of knee cartilage morphology in the first release of participants from the Osteoarthritis Initiative progression subcohort: association with sex, body mass index, symptoms and radiographic osteoarthritis status. Ann Rheum Dis. 2009 May;68(5):674-9.
27. Buckland-Wright C. Protocols for precise radio-anatomical positioning of the tibiofemoral and patellofemoral compartments of the knee. Osteoarthritis Cartilage. 1995 Sep;3 Suppl A:71-80.

28. Kinds MB, Vincken KL, Vignon E, et al. Radiographic features of knee and hip osteoarthritis represent characteristics of an individual, in addition to severity of osteoarthritis. Scand J Rheumatol. 2011;*In press*.
29. Mansfield PJ, Neumann DA. Essentials of Kinesiology for the Physical Therapist Assistant. 2008.
30. Benichou OD, Hunter DJ, Nelson DR, et al. One-year change in radiographic joint space width in patients with unilateral joint space narrowing: data from the Osteoarthritis Initiative. Arthritis Care Res (Hoboken). 2010 Jul;62(7):924-31.
31. Brandt KD, Mazzuca SA, Katz BP, et al. Effects of doxycycline on progression of osteoarthritis: results of a randomized, placebo-controlled, double-blind trial. Arthritis Rheum. 2005 Jul;52(7):2015-25.
32. Reichmann WM, Maillefert JF, Hunter DJ, et al. Responsiveness to change and reliability of measurement of radiographic joint space width in osteoarthritis of the knee: a systematic review. Osteoarthritis Cartilage. 2011 May;19(5):550-6.
33. Kinds MB, Bartels LW, Marijnissen AC, et al. Feasibility of bone density evaluation using plain digital radiography. Osteoarthritis Cartilage. 2011 Nov;19(11):1343-8.
34. McAlindon T. Glucosamine and chondroitin for osteoarthritis? Bull Rheum Dis. 2001 Jul;50(7):1-4.

Chapter 9

Feasibility of bone density evaluation using plain digital radiography

MB Kinds[1,2]
LW Bartels[2]
ACA Marijnissen[1]
KL Vincken[2]
MA Viergever[2]
FPJG Lafeber[1]
HWAM de Jong[3]

[1]Rheumatology & Clinical Immunology, University Medical Center Utrecht
[2]Image Sciences Institute, University Medical Center Utrecht
[3]Radiology, University Medical Center Utrecht

Osteoarthritis & Cartilage 2011 Nov;19(11):1343-1348

Abstract

Objective

For the radiographic evaluation of subchondral bone changes (sclerosis) in osteoarthritis (OA), bone density is commonly subjectively assessed. Bone density evaluation using plain digital radiography might be influenced by acquisition and post-processing (PP) settings. Objective of this study was to evaluate the effects of these settings on the measurement of bone density using digital radiographs.

Methods

A bone density standard (BDS) of hydroxyapatite (HA) mimicked a bone density range of 1.0-5.75 g/cm^2. Digital radiographs were acquired with variation in acquisition settings, and with clinical and minimal PP. An aluminum step wedge served as an internal reference to express the gray values of the BDS in mm aluminum equivalents (mmAl). The relation (R^2) between actual bone density and bone density normalized to the reference wedge was evaluated with linear regression analyses for radiographs with variations in PP and acquisition settings. Precision of bone density measurement of the BDS was evaluated for application in clinical practice.

Results

The correlation between actual bone density and bone density normalized to the reference was improved by changing PP from clinical (R^2=0.96) to minimal (R^2=0.98). Higher tube voltage (kV) improved the correlation further. Even for clinical PP, average SD was 0.97 mmAl, much smaller than the change of 2.51 mmAl clinically observed in early OA, which implies the feasibility of bone density measurements on digital radiographs.

Conclusion

Changing PP and acquisition settings in clinical practice can have profound effect on outcome. If done with care, accurate bone density measurement is feasible using plain digital radiography.

Introduction

Changes in subchondral bone density are an important feature of osteoarthritis (OA) and comprise a complex sequence of changes in the subchondral bone plate and underlying trabecular bone[1,2]. For measurement of bone density several methods have been reported on, including dual energy digital radiography (DEDR)[3], quantitative computed tomography (QCT)[4], radiographic absorptiometry[5], but most importantly dual energy X-ray absorptiometry (DEXA)[6] which is the most validated and commonly used method. For evaluation of structural changes due to OA, plain radiographs are commonly acquired. If the quantitative evaluation of clinically relevant bone density changes on these radiographs is proved feasible, it obviates the need to acquire additional DEXA scans. Thus far, radiographs are used by rheumatologists and orthopedic surgeons for diagnosis of sclerosis (presence of increased bone density)[7-9] and for grading of bone density on categorical scales[10,11]. Combining diagnosis and quantification in one examination reduces time, costs, and patient radiation exposure.

Although the use of film-screen radiography has been described in the evaluation of bone density[9,12], this technique has been almost completely replaced by digital radiography. The accuracy of digital radiography in bone density measurement has received no attention however. One important feature of digital radiography is that image post-processing (PP) is incorporated in the scan protocol. This PP generally includes adjustment of contrast curves and application of non-linear image filters to optimize image quality parameters such as contrast and noise. PP aims at improving diagnostic readability, rather than allowing quantitative analyses to assess bone density changes for longitudinal evaluation Furthermore, the acquisition settings including tube voltage (in kilovolt: kV), exposure (in milliampere seconds: mAs), and filtration can vary between technologists, exam rooms, and institutes, which may influence cross-sectional or longitudinal bone density evaluation[13,14].

To enable quantification of bone density, independently of PP and acquisition settings, the inclusion of an aluminum (step) wedge in the radiographic field-of-view has been suggested[15,16]. In this way the gray values of the bone can be expressed in mm aluminum equivalents (mmAl).

The objective of this study is to determine the feasibility of bone density evaluation using digital radiography. The influence of variations in acquisition and PP settings on bone density measurements was evaluated by means of phantom experiments. To evaluate the applicability of bone density measurements in clinical practice, the precision was calculated and compared to clinically observed bone density values.

Methods

Evaluation of bone density standard

A perspex bone density standard (BDS) was constructed using hydroxyapatite (HA: $Ca_5(OH)(PO_4)_3$; Sigma Aldrich) to simulate bone densities of 1.00, 2.00, 2.75, 3.50, 4.25, 5.00, and 5.75 g/cm^2. The BDS consisted of eight columns of 15x15 mm^2 (34 mm deep) and was closed by a perspex lid resulting in a total of 6 mm of perspex on the bottom and top of the columns. The bone density range of the BDS was based on bone density of the (subchondral) bone of a healthy human knee joint (2.21 g/cm^2, determined by DEXA).

Radiographs of the BDS were acquired using a clinical digital radiography system (*Digital Diagnost*, Philips Healthcare, Best, The Netherlands), equipped with *Unique* PP software. An aluminum step wedge (40x200 mm; the same in all digital radiographs) with increasing thickness of 2 mm at each of 20 steps was added in the field-of-view of the BDS as an internal reference for bone density measurement. Bone density was measured on the digital radiographs by placing a circular region of interest (ROI) with a diameter of 1 cm in each column of the BDS.

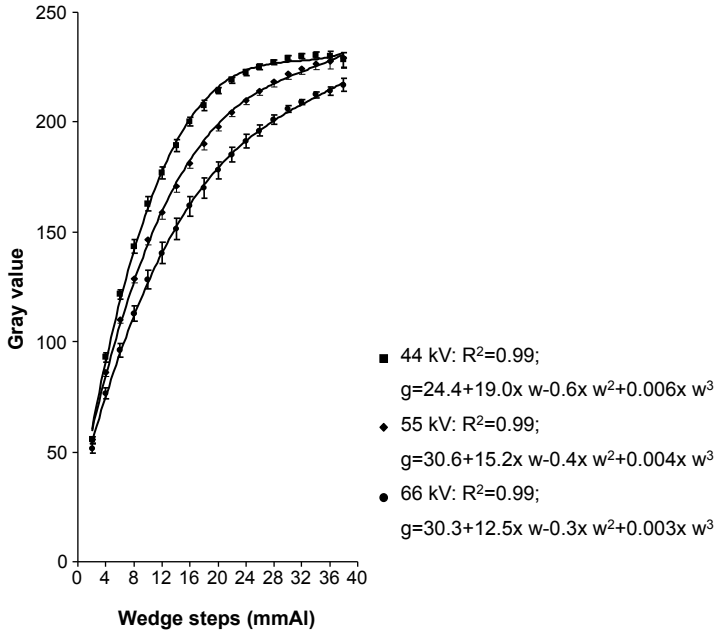

Figure 1 Gray values of the reference wedge for radiographs with Clinical PP. Mean and standard deviation (SD: error bars) of three repeated radiographs at 44 kV (squares), 55 kV (default: diamonds), and 66 kV (dots). Correlation (R^2) and equation between wedge steps (w) and gray value (g)

For this, the in-house developed Knee Images Digital Analysis (KIDA) software was used[16]. This method enabled semi-automatic extraction of the reference step wedge from the image to calibrate pixel size in the image and to provide a look-up-table for conversion of ROI values from the radiograph to mmAl. To this end, a third order polynomial was fitted to the wedge for each radiograph (variation and repetition) separately (figure 1). The effect of X-rays hitting the reference wedge, BDS, and detector under oblique angles was taken into account in the KIDA method. The ROIs were placed in the center of each step of the reference wedge and of each column of the BDS where no partial attenuation was observed. To mimic the clinical situation a human cadaver knee (male adult) was added to the radiographic field-of-view (see figure 3) in plane with the aluminum step wedge.

The default protocol for acquisition parameters and PP algorithm was: a source image distance (SID) of 120 cm, 55 kV, 5 mAs, no added tube filtration, the BDS in the center of the field-of-view (from left to right: step wedge, BDS, and cadaver knee), and clinical PP as determined by the manufacturer. This protocol was also used for radiographs in the Dutch Cohort Hip & Cohort Knee (CHECK) study in which 1002 participants with early signs of hip and/or knee OA are monitored[17]. Based on this protocol, the following settings were systematically varied:

i) *Acquisition parameters*: 44, 55, or 66 kV; 5 mAs or automatic exposure control (AEC); tube filtration without added filter or with 2 mm of aluminum; position of the BDS in the field-of-view at the center (with step wedge left and cadaver knee right), at the outer right (with step wedge left and cadaver knee at center), or at the top (with step wedge left and cadaver knee at center).

ii) *PP algorithm*: either clinical (*Unique*) or minimal (with PP at minimal strength).

To assess reproducibility (precision), three repeated series of radiographs were made in random order for each of the variations in the acquisition parameters and in PP algorithm.

Application in clinical practice

It was evaluated whether the results from the BDS experiments have implications for bone density evaluation in clinical practice. Therefore, variation in bone density measurement on the BDS was compared with changes in bone density measurement in knee radiographs from a large clinical study in early OA (CHECK).

In this study knee radiographs of 1002 participants (n=1095) were acquired with the default settings with clinical PP, according to the posteroanterior semiflexed protocol aimed at correct alignment of the medial tibia. In this study also an aluminum step wedge was added in the plane with the joint in the field-of-view. Circular ROIs were placed at the joint margins of the lateral and medial femur, and lateral and medial tibia (for details see: [16]), to determine bone density in subchondral bone areas where fat marrow is expected not to be of influence.

Statistical analyses

Evaluation of bone density standard

DEXA scanning (Hologic Discovery) was used to validate the use of the BDS for evaluation of bone density measurement using digital radiography. For digital radiographs linear regression analysis was used to evaluate the relation between actual bone density (HA in g/cm^2: independent variable) and measured bone density by normalization of gray value to the reference wedge (mmAl: dependent variable) for the eight columns of the BDS. Regression coefficients (ß) with 95% confidence interval (95%CI) and explained variance (R^2) were determined from the mean bone density values of three repeated radiographs. For R^2 values 95%CI were determined by calculation of mean ± 1.96 x standard error of the mean (SEM). Furthermore, impact of variations in kV, mAs, tube filtration, PP algorithm, and cadaver knee location was investigated.

Application in clinical practice

To assess application in clinical practice, the precision (reproducibility) of the bone density measurement, which is dependent on acquisition of the digital radiographs and on KIDA measurement, was evaluated. From the three repeated radiographs the mean and SD were calculated for each column separately. Average SD was calculated as: square root (average variance). The coefficient of variation (CV) was determined as SD divided by mean (x 100%) for each column and the average CV of eight columns. The SD and CV were evaluated in three situations: default protocol with clinical PP (CHECK), default protocol with minimal PP, and 66 kV with minimal PP (optimal settings). The SD was compared to the changes in bone density found during two-year follow-up in the clinical study (CHECK). Statistical analyses were performed using Statistical Package for the Social Sciences (SPSS) 15.0 software.

Results

Evaluation of bone density standard

As expected an excellent correlation between actual amounts of HA (g/cm^2) and the bone density values measured with DEXA (g/cm^2) was found: R^2=0.9947±0.005 (measured bone density=0.14 + 0.96 x actual bone density). The DEXA value of the medial tibia of the cadaver knee was 1.70 g/cm^2 (ROI placement according to KIDA).

In general for the digital radiographs a strong correlation between actual and normalized bone density was found for all investigated acquisition settings (variation in kV settings: figure 2A for clinical PP and figure 2B for minimal PP).

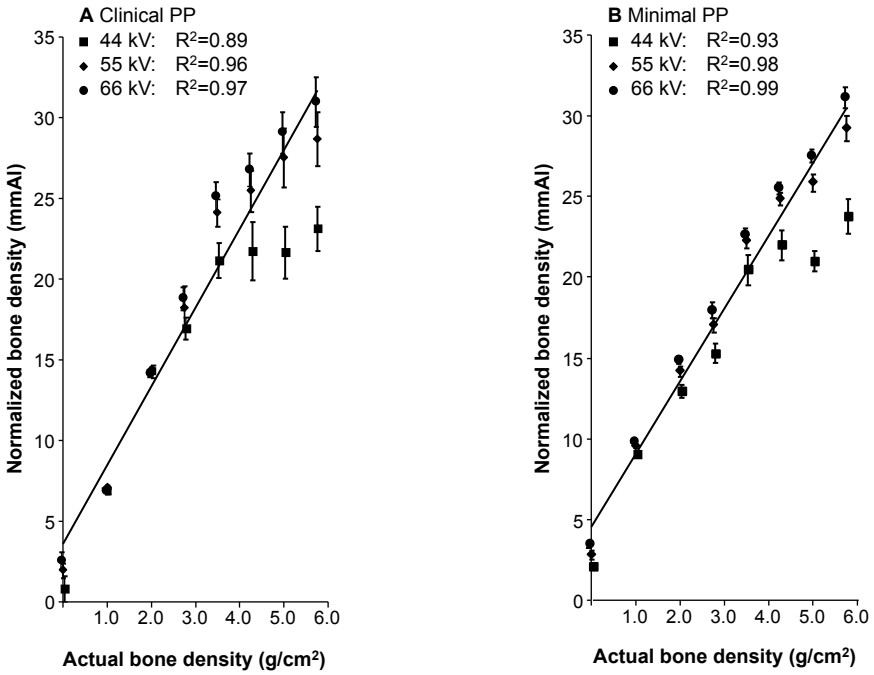

Figure 2 Correlation (R^2) between actual bone density and bone density normalized to the reference wedge for **A** Clinical PP; **B** Minimal PP. Mean and standard deviation (SD: error bars) of three repeated radiographs at 44 kV (squares), 55 kV (default: diamonds with regression line), and 66 kV (dots)

Table 1 presents for all radiographs R^2 (correlation) with 95%CI, ß (regression coefficients) with 95%CI, and constants (intercept) as obtained with linear regression analyses. For all variations in acquisition parameters the correlation was better for minimal PP (table 1B) than for clinical PP (table 1A). For example, for default radiographs with clinical or minimal PP, R^2 equaled 0.96 (95%CI: 0.94-0.97) or 0.98 (95%CI: 0.98-0.98) respectively. The regression coefficients (ß) and the 95%CI around ß were smaller with minimal PP than with clinical PP for all variations in acquisition parameters. Correlation improved with higher kV, independent of PP settings. Settings of mAs, position of the BDS in the field-of-view, and filtration were of no influence on the correlation.

Figure 3 illustrates the appearance of radiographs with variations in acquisition and PP settings.

Table 1 Linear regression analyses: relation between actual BD and BD normalized to the reference for default radiographs and variations in acquisition settings for **A** Clinical PP; **B** Minimal PP

	kV	mAs	Filter	Position	PP	R^2	(95%CI)	ß	(95%CI)	Cons
A Clinical PP										
Default	55	5	no	center	clinical	0.96	(0.94-0.97)	4.89	(3.86-5.93)	3.57
	Variations in acquisition settings									
Regular	***66***	***0.45-0.52***	no	center	clinical	0.97	(0.96-0.98)	5.19	(4.22-6.16)	3.31
kV	***66***	5	no	center	clinical	0.97	(0.96-0.98)	5.26	(4.31-6.20)	3.34
	44	5	no	center	clinical	0.89	(0.87-0.91)	3.87	(2.52-5.23)	4.06
mAs	55	***0.62-0.77***	no	center	clinical	0.95	(0.94-0.96)	4.73	(3.66-5.81)	3.74
Filter	55	5	***2mmAl***	center	clinical	0.96	(0.93-0.98)	4.98	(3.94-6.02)	3.72
Position	55	5	no	***right***	clinical	0.97	(0.95-0.98)	5.92	(4.78-7.05)	3.29
	55	5	no	***top***	clinical	0.94	(0.93-0.95)	5.34	(4.03-6.65)	5.67
B Minimal PP										
Default	55	5	no	center	minimal	0.98	(0.98-0.98)	4.52	(3.86-5.17)	4.51
	Variations in acquisition settings									
Regular	***66***	***0.42-0.51***	no	center	minimal	0.99	(0.98-0.99)	4.74	(4.21-5.27)	4.51
kV	***66***	5	no	center	minimal	0.99	(0.99-0.99)	4.73	(4.27-5.20)	4.76
	44	5	no	center	minimal	0.93	(0.92-0.94)	3.65	(2.64-4.66)	4.74
mAs	55	***0.60-0.80***	no	center	minimal	0.98	(0.97-0.98)	4.48	(3.76-5.20)	4.48
Filter	55	5	***2mmAl***	center	minimal	0.98	(0.98-0.99)	4.51	(3.90-5.13)	4.49
Position	55	5	no	***right***	minimal	0.99	(0.99-0.99)	4.84	(4.30-5.38)	3.48
	55	5	no	***top***	minimal	0.99	(0.99-0.99)	5.02	(4.50-5.55)	3.92

Variations in acquisition settings marked in ***bold italic***, mAs: fixed value of 5 or range for three repeated radiographs provided in case AEC was used, Cons: constant in regression equation

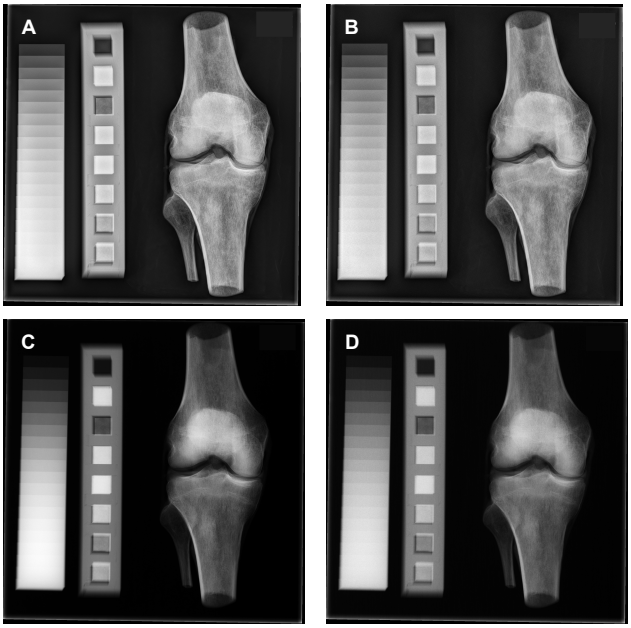

Figure 3 Screenshot of digital radiographs:
A Default and Clinical PP; **B** 66 kV, AEC and Clinical PP;
C Default and Minimal PP; **D** 66 kV, AEC and Minimal PP

Also for DEXA scanning of the cadaver knee a clear linear correlation was found (R^2=0.91) between the actual DEXA bone density values and the mmAl values measured on the plain radiographs, with ROI on the medial tibia on the DEXA and digital radiographs (as in the KIDA measurement).

Application in clinical practice

Table 2 presents the mean, SD, and CV of measured bone density per column of the BDS, respectively for default protocol with clinical PP (A), default with minimal PP (B), and for 66 kV and minimal PP (C). As can be expected because of attenuation of the X-rays, the SD increased with increasing bone density values of the BDS. The CV did not systematically increase with increasing bone density. The CV was highest in the column with bone density 0.00 (without HA), where mean bone density values in mmAl were relatively low and SDs relatively large, for all radiographic protocols. The SD was largest for default radiographs with clinical PP (compared to radiographs with minimal PP) with average SD of 1.15 mmAl (95%CI: 0.66-1.49) and largest SD of 1.81 mmAl. The average SD was found to be slightly larger (1.22 mmAl; 95%CI: 0.65-1.60) when only the clinically relevant range was used (5 columns ranging from 2.0 to 5.0 g/cm^2).

When bone density measurement with KIDA was applied on the cadaver knee, bone density normalized to the reference was 21.46 mmAl in the medial tibia. Bone density measurement in clinical practice (CHECK) resulted in mean bone density of the medial tibia of 25.34 mmAl, with an SD of 6.67 mmAl at baseline (n=1095 knee joints). The mean increase from baseline to two-year follow-up was 2.51 mmAl (SD: 6.76 mmAl), which was larger than the average (and largest) SD for measurement on the BDS on radiographs according to clinical practice (default with minimal PP). An even larger increase in bone density of 4.80 mmAl (SD: 4.38mmAl) was found in a subgroup of 17 knee joints with actual radiographic progression during two-year follow-up (based on change of Altman grade of medial tibia sclerosis (increased bone density) from 0: absent to 1: present).

Table 2 Bone density normalized to reference: mean, SD, and CV for three repeated radiographs for **A** default and clinical PP; **B** default and minimal PP; **C** 66 kV and minimal PP

	A default and clinical PP			**B** default and minimal PP			**C** 66 kV and minimal PP		
Bone density (g/cm^2)	**Mean** (mmAl)	**SD** (mmAl)	**CV** (%)	**Mean** (mmAl)	**SD** (mmAl)	**CV** (%)	**Mean** (mmAl)	**SD** (mmAl)	**CV** (%)
0.00	1.93	0.47	24.1	2.78	0.27	9.9	3.43	0.20	6.0
1.00	7.06	0.16	2.3	9.51	0.12	1.3	9.83	0.06	0.6
2.00	14.39	0.18	1.3	14.15	0.32	2.3	14.85	0.22	1.5
2.75	18.17	1.36	7.5	17.04	0.45	2.6	17.93	0.48	2.7
3.50	24.10	0.84	3.5	22.21	0.44	2.0	22.65	0.36	1.6
4.25	25.45	1.28	5.0	24.85	0.42	1.7	25.51	0.38	1.5
5.00	27.49	1.81	6.6	25.84	0.52	2.0	27.54	0.40	1.4
5.75	28.66	1.69	5.9	29.21	0.81	2.8	31.13	0.65	2.1
Average		1.15	7.0		0.46	3.1		0.38	2.2

Bone density: actual bone density in hydroxyapatite (HA)

Discussion

This is the first study to demonstrate that bone density measurement on plain digital radiographs is feasible. However, since variations in digital radiography settings at acquisition influence the outcome, bone density measurements should be interpreted with caution. In clinical practice variations in acquisition settings within institutes and specifically between institutes (in multicenter trials such as the CHECK study) have been found to vary in the ranges as described in the present study. Furthermore, differences in performance between different brands and types of digital radiography systems will exist.

Digital radiographs without a reference need to be evaluated with caution since in optimization of images, PP algorithm plays an important role. For example, in the longitudinal evaluation of a knee joint, the clinical PP algorithm might yield radiographs with a similar appearance while bone density actually changed based on disease (e.g. OA) or while acquisition settings are different (figure 3A compared to 3B). On the other hand, radiographs might appear different as a result of variations in PP settings (or different radiography systems) within and between centers rather than as a result of bone density changes (figure 3A compared to 3C). Although in clinical practice variations in digital radiography and PP settings will occur that influence bone density measurement, the addition of a reference enables an adequate assessment of the gray values.

One limitation of the present study might be that the used BDS is a simplified representation of tissue composition of a human knee without anatomical resemblance. However, the mean bone density values determined with DEXA at the medial tibia (similar to KIDA) were 2.21 g/cm^2 for the healthy human knee joint, with a linear correlation with bone density values on digital radiographs (R^2=0.91), and 1.70 g/cm^2 for the cadaver knee joint, which showed that the BDS represented a clinically relevant range.

The BDS experiments indicated that the precision of bone density measurement can be increased by using minimal PP rather than clinical PP and by applying relatively high kV. The relation between actual bone density and bone density normalized to the reference is weak when low tube voltage (44 kV) is used especially at larger bone density values, which might be due to relatively more absorption of the beam by the knee joint. Although the application of higher kV improves linearity of the relation between actual and normalized bone density, patient exposure needs to be taken into account. Improved accuracy without additional patient exposure can be reached by using higher kV in combination with lower mAs. Applying minimal PP to improve accuracy is not easily applicable in regular clinical practice since clinical PP is required to provide optimal diagnostic image quality, and in general can not easily be bypassed in clinical practice.

In conclusion, the BDS experiments and the comparison to clinical data indicate that bone density measurement using digital radiography is feasible in a clinically relevant range. Variations in acquisition and post-processing settings within and between clinics can have profound effect on bone density evaluation and should therefore be considered with caution.

As compared to the default clinical protocol, the accuracy of bone density measurements can be improved by applying only minimal image post-processing and a relatively high kV. Provided properly performed, plain digital radiography may yield, in addition to OA characteristics, reliable data on bone density which reduces the need for additional imaging techniques.

Acknowledgements

We acknowledge the department of Radiology of the University Medical Center Utrecht for their cooperation on the acquisition of digital radiographs and DEXA scans. We acknowledge Paco Welsing for methodological support.

References

1. Burr DB. The importance of subchondral bone in osteoarthrosis. Curr Opin Rheumatol. 1998 May;10(3):256-62.
2. Dedrick DK, Goldstein SA, Brandt KD, et al. A longitudinal study of subchondral plate and trabecular bone in cruciate-deficient dogs with osteoarthritis followed up for 54 months. Arthritis Rheum. 1993 Oct;36(10):1460-7.
3. Tahvanainen PS, Lammentausta E, Pulkkinen P, et al. Dual-energy digital radiography for the assessment of bone mineral density. Acta Radiol. 2010 Jun;51(5):543-8.
4. Kleerekoper M, Nelson DA, Flynn MJ, et al. Comparison of radiographic absorptiometry with dual-energy x-ray absorptiometry and quantitative computed tomography in normal older white and black women. J Bone Miner Res. 1994 Nov;9(11):1745-9.
5. Kleerekoper M, Nelson DA, Peterson EL, et al. Reference data for bone mass, calciotropic hormones, and biochemical markers of bone remodeling in older (55-75) postmenopausal white and black women. J Bone Miner Res. 1994 Aug;9(8):1267-76.
6. Blake GM, Fogelman I. Technical principles of dual energy x-ray absorptiometry. Semin Nucl Med. 1997 Jul;27(3):210-28.
7. Dieppe PA, Cushnaghan J, Shepstone L. The Bristol 'OA500' study: progression of osteoarthritis (OA) over 3 years and the relationship between clinical and radiographic changes at the knee joint. Osteoarthritis Cartilage. 1997 Mar;5(2):87-97.
8. Szebenyi B, Hollander AP, Dieppe P, et al. Associations between pain, function, and radiographic features in osteoarthritis of the knee. Arthritis Rheum. 2006 Jan;54(1):230-5.
9. Nevitt MC, Lane NE, Scott JC, et al. Radiographic osteoarthritis of the hip and bone mineral density. The Study of Osteoporotic Fractures Research Group. Arthritis Rheum. 1995 Jul;38(7):907-16.
10. Ledingham J, Regan M, Jones A, et al. Factors affecting radiographic progression of knee osteoarthritis. Ann Rheum Dis. 1995 Jan;54(1):53-8.
11. Miller ME, Rejeski WJ, Messier SP, et al. Modifiers of change in physical functioning in older adults with knee pain: the Observational Arthritis Study in Seniors (OASIS). Arthritis Rheum. 2001 Aug;45(4):331-9.
12. Michel BA, Lane NE, Jones HH, et al. Plain radiographs can be useful in estimating lumbar bone density. J Rheumatol. 1990 Apr;17(4):528-31.
13. Buckland-Wright JC, Wolfe F, Ward RJ, et al. Substantial superiority of semiflexed (MTP) views in knee osteoarthritis: a comparative radiographic study, without fluoroscopy, of standing extended, semiflexed (MTP), and schuss views. J Rheumatol. 1999 Dec;26(12):2664-74.
14. Vignon E, Piperno M, Le Graverand MP, 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 Feb;48(2):378-84.
15. Meakim DW, Ott EA, Asquith RL, et al. Estimation of mineral content of the equine third metacarpal by radiographic photometry. J Anim Sci. 1981 Oct;53(4):1019-26.
16. Marijnissen AC, Vincken KL, Vos PA, et al. Knee Images Digital Analysis (KIDA): a novel method to quantify individual radiographic features of knee osteoarthritis in detail. Osteoarthritis Cartilage. 2008 Feb;16(2):234-43.
17. Wesseling J, Dekker J, van den Berg WB, et al. CHECK (Cohort Hip and Cohort Knee): similarities and differences with the Osteoarthritis Initiative. Ann Rheum Dis. 2009 Sep;68(9):1413-9.

Chapter 10

Summary & Discussion

Summary

Identifying radiographic joint damage in an early phase of osteoarthritis (OA) will advance the understanding of this disabling disease. By detecting specific radiographic characteristics that are important in the onset and early progression of OA and in the relation with symptoms, clinical trials leading to (preventive) treatment strategies can be optimized.
Therefore, in this thesis the use of quantitative measurement (on a continuous scale) of separate radiographic features (by use of digital image analysis) is described. The challenges in evaluating OA from an early phase of the disease, to enable application in clinical research practice, are discussed. In the first section of this thesis the importance of methodological quality in the evaluation of radiographic OA is described. The second section aims at describing the application of measuring separate features of radiographic OA by digital analysis in CHECK (Cohort Hip & Cohort Knee), a cohort of individuals with (very) early signs related to knee and/or hip OA. The last section reports on the role of acquisition quality of standardized radiographs for digital image analysis.

The first step was to answer the question: *Which methodological criteria are important to detect an association between radiographic and clinical OA of hip and knee?* To evaluate the role of such criteria, studies describing a comparison between radiographic and clinical OA characteristics were systematically reviewed (**chapter 2**). In the (only) 39 studies that could be included for this evaluation, associations were more frequently detected when the studies fulfilled the defined methodological quality criteria. These criteria included the definition of OA for study inclusion, the use of standardized radiographic protocols, and the use of standardized radiographic and clinical outcome measures for the evaluation of OA. The lower frequency of an association in studies of lower quality emphasizes the importance of good methodological quality. The standardization of outcome measures appears most important. In the present literature many studies are reported that do not fulfill these methodological quality criteria. This introduces an unintended - but generally appreciated - inconsistent association between radiographic and clinical OA features. Therefore, the importance of methodological quality needs specific attention in future studies to prevent further reporting bias in the literature.

In the following chapters of this thesis participants from CHECK were evaluated for radiographic joint damage in early OA since this study fulfilled most methodological quality criteria. In CHECK inclusion was based on OA related symptoms, and knee and hip radiographs were acquired according to standardized protocols. Multiple outcome measures for radiographic and clinical OA were collected. This enabled evaluation of whether the newly introduced quantitative measurement of separate features by KIDA (Knee Images Digital Analysis) was a valid outcome measure for radiographic OA, and for detecting an association between such features and clinical OA.

In chapter 3 and 4 the measurement of distinct radiographic features by digital analysis was validated, by comparison with commonly applied and already validated outcome measures that are considered the gold standard for evaluating radiographic OA status and progression: K&L (Kellgren & Lawrence) grading and grading according to the Altman atlas.
Chapter 3 addressed the question: *Does the quantitative measurement by digital image analysis result in higher sensitivity to detect progression of radiographic knee damage than qualitative grading according to the Altman atlas?* Studying radiographic progression in symptomatic knees from baseline to two-year follow-up in CHECK showed that the sensitivity to detect early progression of radiographic knee damage was similar for quantitative KIDA measurement and qualitative Altman grading. Importantly, only in a small percentage of knees a radiographic change was identified with any of the methods. The sensitivity to change was small using KIDA measurement and Altman grading, and surprisingly only a limited level of agreement exists between the two methods in the identification of individuals with radiographic progression. The inter- and intra-observer variability of quantitative measurement (KIDA) was significantly smaller than that of qualitative grading. Finally, it was considered that the surplus value of KIDA might improve when either more radiographic progression occurs, or more importantly when reproducibility of image acquisition can be improved. This is supported by the results of a study on quality of image acquisition as described in chapter 8.

In **chapter 4** it was studied how the measurement of separate features represents characteristics of the joint at the baseline assessment in CHECK. The question was: *To what extent do radiographic features of knees and hips, which are normally related to osteoarthritis, represent characteristics of an individual in addition to osteoarthritis severity?* For this purpose, in addition to KIDA for digital analysis of the knee, Holy's software was used for digital image analysis of joint space width (JSW) of the hip. The finding that a large range of values existed for the distinct quantitative features of knee and hip within each K&L grade suggests that the features also represent characteristics of individuals. It appeared that, within participants with hardly any radiographic joint damage according to the gold standard (K&L grade 0-I), radiographic features were correlated between joints. This applied for the contralateral joints, but also ipsilateral and diagonal hips and knees were correlated. As such, radiographic features of knee and hip OA (when analyzed in more detail by use of quantitative measurement) are considered to represent characteristics of an individual, in addition to severity of osteoarthritis. This implies that separate radiographic features of joints within individuals need to be accounted for when joint damage due to OA is studied in a population. The difference in the specific radiographic features between both contralateral knees was taken into account in further analyses of these characteristics in the process of OA (chapters 5 to 7).

The conclusion from the first part of this thesis is that the use of standardized outcome measures of OA is important in evaluating radiographic damage, and in finding an association between radiographic and clinical characteristics. Radiographic OA can be described by quantitative measurement of separate features. Proper image acquisition is of major importance to utilize the surplus value of digital analysis, specifically in an early phase of OA. Digital image analysis of separate features enables identification of individual characteristics that dominate differences within individuals early in the disease, whereas later in the disease radiographic characteristics of joint degeneration prevail.

The second part of this thesis (chapter 5 to 7) deals with the application of quantitative measurements of distinct radiographic features in clinical research. In **chapter 5** the objective was: *To evaluate radiographic OA development over time from an early phase of the disease using separate features (measured by knee images digital analysis), and to evaluate how these features relate to each other and to clinical characteristics of OA.* Using digital analysis radiographs from baseline, two-year and five-year follow-up were evaluated, and specific radiographic features were found to develop between different time points in CHECK. These features defined as minimum JSW, medial JSW, lateral JSW, varus angle, osteophyte area, eminence height, and bone density all progressed over time, were related to each other, and were related to clinical outcome. Importantly, the mutual relation between these specific features and their relation with clinical outcome changed over time, for distinct features in a different way. These findings imply that quantitative measurement of separate features indeed represents characteristics of radiographic joint damage already in an early phase of the disease. Measurement of these features might be of use to identify specific phenotypes of OA. This may advance the design of clinical trials and the development of more targeted treatment strategies. In chapter 7 it is shown that the identified features of radiographic OA indeed enable identification of relevant subpopulations (phenotypes) of CHECK participants.

Next, it was analyzed whether measurement of separate features in an early phase of the disease improved the prediction of OA outcome at five-year follow-up. In **chapter 6** the question was addressed: *Whether and which separate features, measured on knee radiographs of individuals with recent onset knee pain, are associated with incidence of radiographic OA and persistence and/or progression of clinical OA during five-year follow-up?* The additional value of acquiring radiographs and of measuring distinct radiographic features in detail, next to the standard clinical assessment in individuals who visit a physician with their first complaints related to OA, was analyzed. For these individuals, using separate quantitative measurements to predict persistence and/or progression of clinical OA appeared challenging. The predictive ability was too low for use in practice at the individual level. In this early phase of OA however, evaluation of radiographic characteristics clearly added to the prediction of incident radiographic OA five years later. The prediction of incidence of radiographic OA improved from poor to fair when quantitative radiographic

features were evaluated, in addition to demographics and clinical variables. Evaluation of separate quantitative features performed slightly better in this respect than K&L grading. As such, it was concluded that measuring distinct radiographic features is of value in identifying individuals at high risk of developing radiographic OA. When the 'predictive score' (based on multivariate regression coefficients) was larger than the cut-off for optimal specificity, the chance of incident radiographic OA was 54% instead of the prior probability of 19% in the population of individuals presenting themselves with early complaints related to OA. This knowledge is of use when selecting individuals very early in the disease for treatment strategy trials (e.g. using disease modifying osteoarthritic drugs: DMOADs) aimed at preventing joint degeneration.

In addition to identifying predictors for radiographic progression in individuals with very early signs of OA, identification of phenotypes of radiographic OA progression is considered of relevance. Whether measuring specific features in detail (as continuous variables) is of use in the identification of such phenotypes, was evaluated in **chapter 7**. The question addressed was: *Can phenotypes of progression of radiographic knee OA be identified by quantitative measurement of separate radiographic features?* Based on separate radiographic features as described in chapter 5, five phenotypes of radiographic progression were detected: 'severe' and 'no' progression phenotype, 'early' and 'late' progression phenotype, and 'bone density' phenotype. Demographic and clinical characteristics were shown to differ between phenotypes, and also specific features were identified that could predict phenotypes of progression. Predictive ability was improved by defining a cut-off for radiographic features, although this appeared insufficient for use at an individual level. Nonetheless, the identification of these phenotypes based on radiographic features gives the opportunity to improve clinical trials design and the development of more specific treatment strategies.

In summary, this part of the thesis shows that the measurement of separate quantitative features of radiographic OA by KIDA provides tools for the evaluation of radiographic OA, even in a very early phase of disease. Importantly, such specific radiographic features are found to be related to clinical outcome, and to be predictors of onset of radiographic OA. Furthermore, based on the patterns of development of these separate features, phenotypes of radiographic progression can be identified. This finding is of value in selection of individuals early in the disease to evaluate more targeted treatment strategies.

In the third and final part of this thesis the importance of acquisition quality for reliable measurement of radiographic parameters of knee OA is discussed. In **chapter 8** the following question was addressed: *What is the influence of changes in knee position during acquisition of radiographs on the measurement of radiographic characteristics?* It appeared that systematic variations in knee joint position during image acquisition, and particularly in the extension angle, influenced the quantitative KIDA measurement of radiographic OA

parameters. Changes in the KIDA parameters were considered relevant when they were similar to or larger than the detected differences in knees with OA progression from CHECK during two-year follow-up. Thus variations in knee positioning - which can easily occur during acquisition in trials and clinical practice despite standardization - have a significant influence on the quantitative measurement of most separate radiographic parameters of OA. Although the parameters measured by digital analysis are sufficiently robust, the surplus value of these quantitative measurements over qualitative grading will only pay off when standardization during radiographic acquisition is improved. It is still of value to further improve standardization of acquisition since radiography remains cheap and easily accessible, in contrast with other imaging techniques like magnetic resonance imaging.

In the previous chapter it appeared that variations in all components of the knee position for image acquisition influenced the bone density (sclerosis) measurement. For the evaluation of changes in bone density the use of conventional film-screen radiography has been described previously. In recent years however this technique has been replaced almost completely by digital radiography. With digital radiography, image quality is considered to be improved for clinical purposes since contrast is optimized and noise is reduced by the implementation of post-processing after acquisition. The projected bone density is influenced by this post-processing however, and also by changes in acquisition settings that are likely to occur in (long-term) clinical trials. Therefore the question in **chapter 9** was: *What are the effects of acquisition and post-processing settings on the measurement of bone density using plain digital radiographs?* For the first time the feasibility of bone density measurement on plain digital radiographs was demonstrated. Bone density measurements should be interpreted with caution however, since variations in digital radiography settings at acquisition influence the outcome. There is a trade-off between image quality and accuracy of bone density measurement. The accuracy was improved when post-processing settings were minimized compared to standard clinical protocols, and when the energy (in kilovolts) was increased. Even with clinically applied acquisition settings and standard post-processing, the slight variation in repeated radiographs was much smaller than the difference measured during the first two years of follow-up in CHECK. Consequently, when acquisition and post-processing settings are monitored throughout clinical trials, accurate bone density measurement seems feasible using plain digital radiographs. This reduces the need for additional imaging techniques and the accompanied X-ray exposure when using Dual Energy X-ray Absorptiometry.

The results from the last part of this thesis emphasize the need for optimal standardization and careful monitoring of radiographic procedures during acquisition of images that are used for digital image analysis. Better standardization will clearly improve the ability of separate features to describe, specifically in an early phase of OA, the onset and progression of specific radiographic characteristics, and the relation with clinical characteristics.

Discussion

Evaluation of radiographic joint damage in early osteoarthritis

Radiographic evaluation in early osteoarthritis

Structural joint changes are commonly evaluated by use of radiography, although not all structures that are involved in the OA process are imaged on radiographs. Direct evaluation of different joint tissues is enabled by magnetic resonance imaging (MRI)[1]. This technique may allow for the detection of specific structural changes before they become evident on radiographs[2]. Interestingly however, these developments still do not outweigh the advantages of plain radiography; the technique is generally available, fast, cheap, and non-invasive[3]. For these reasons, radiography remains the gold standard to evaluate structural damage in clinical practice[3-5].

For further understanding of OA, evaluating individuals in an early phase of the disease or at high risk of development of the disease is essential[5]. Individuals in an early OA population are the most interesting to study, since these individuals with symptoms can be evaluated for the presence of subtle radiographic characteristics that will not be graded as K&L ≥II (evident radiographic OA) yet. Detecting subtle and specific radiographic features of OA may identify susceptibility for onset and progression of structural joint damage. This can be a valuable tool, e.g. to select individuals in a very early phase of the disease for clinical trials that evaluate more targeted treatment strategies. Furthermore, detecting specific radiographic features enables identification of individuals who have painful joints suspected for OA, but who will not develop joint degeneration in the upcoming years. Such individuals would potentially benefit from an entirely different type of treatment than those who will develop radiographic progression.

Early osteoarthritis in CHECK

The CHECK study aimed at evaluating individuals at risk of developing OA, by including individuals of 45-65 years, with pain and/or stiffness of the knee and/or hip, and without a visit or with a first visit no longer than six months ago to the general practitioner for these complaints[6]. Although the ACR (American College of Rheumatology) criteria for clinical OA were fulfilled in 76% of knees and 24% of hips[6] at inclusion, CHECK participants suffered from only mild symptoms. Structural damage was not present yet or was only limited, since at inclusion only 3% of knees and 7% of hips had K&L grade ≥II. Thus the far majority of participants in CHECK had no radiographic OA or was just suspected to have radiographic OA. Since standardized radiographs are made frequently during at least ten years of follow-up in 1002 individuals, this cohort is explicitly useful for detecting early subtle structural changes. During longitudinal evaluation individuals can be identified who develop radiographic OA, and in retrospect radiographic and clinical features can be identified that predict progression.

On average, the participants in CHECK developed structural damage during the first five years of follow-up, as measured on radiographs using digital image analysis (see chapter 5).

Only in a (small) portion of participants the complaints were progressive to such an extent that joint replacement was indicated (9 knees and 37 hips at five-year follow-up). Interestingly, an evident increase in OA symptoms during follow-up was not clearly present on average for the entire cohort[7]. A substantial part of participants actually had a decrease in symptoms after study inclusion. This might be because some participants benefit from adaptations to their lifestyle (e.g. exercise and/or weight loss), the development of coping strategies, and/or the use of adequate pain medication. On the other hand, the symptoms of these participants might have had another origin than OA, despite the inclusion and exclusion criteria. There is expectedly a subgroup of individuals with transient pain, which is supported by the observations in chapter 7. In this chapter participants were identified who did not show any progression in any of the separate radiographic features during five-year follow-up, and interestingly this subgroup reported evidently less joint pain during follow-up.

Digital image analysis to evaluate radiographic joint damage

Digital image analysis versus conventional grading

To enable the evaluation of subtle individual radiographic characteristics, dedicated digital image analysis was developed. More detailed evaluation is considered to be provided by quantitative digital analysis instead of conventional qualitative grading. In addition to several other digital analysis techniques[8-11], KIDA (Knee Images Digital Analysis)[12] and Holy's hip evaluation software[8-11,13-15] have been developed for this purpose. Most of these software packages have been designed to measure separate characteristics such as JSW[8-11,13-16] and joint alignment[10,17]. KIDA is actually the first method that provides detailed analysis of these and additional radiographic parameters (i.e. osteophyte area, eminence height, and bone density). The inter- and intra-observer variation of KIDA is very low[12], because the method uses a mathematical approach without the need for any personal clinical interpretation. This is in contrast with qualitative grading methods like the Altman atlas with which multiple separate characteristics of OA are scored on an ordinal scale[18]. Chapter 3 describes that only a limited level of agreement exists between KIDA measurements and Altman grading in identifying individuals with radiographic progression. This underscores the intrinsic differences between both techniques; the one with and the other (KIDA) without subjective interpretation.

Radiographic quality for digital image analysis using KIDA

An objective method like KIDA has practical advantages. A learning curve hardly exists, the method can be applied by everyone, and no specific educational background is required. On the other hand there are drawbacks as well, e.g. some changes visualized on radiographs are not related to joint degeneration. Such changes may be the result of individual characteristics, quality of images, standardization of joint position for imaging, and settings during acquisition of radiographs. This may in part explain the observed lack of agreement between both methods described in chapter 3. To prove of additional value, specifically digital analysis requires reproducible positioning of the joint[19], since with this objective

method variations in joint position between subsequent radiographs are commonly not taken into account during the measurement (which can be done with subjective grading).
Chapters 8 and 9 demonstrate that even subtle changes in joint position and acquisition settings are of influence on the projected image, and thereby on the measurement of KIDA parameters. In chapter 8 it appeared that relatively small changes in knee positioning that easily occur in clinical trial setup are of major influence on the KIDA measurement. In particular the bone density measurement was influenced by variation in almost all components of knee joint position.
In chapter 9 it was described that acquisition settings influence the bone density measurement in KIDA. Next to acquisition settings, post-processing settings, which were introduced with the transition from film-screen to digital radiography, also appeared of influence on reliable bone density measurement.
As demonstrated in this thesis, in clinical trials or observational cohorts like CHECK, where radiographs are acquired during long follow-up and in multiple centers, variations in joint positioning occur. This is despite the standardization of radiographic protocols that significantly improved reproducibility of joint position during image acquisition in longitudinal studies[20]. As such, variations in the measurement of radiographic OA features are mainly introduced during image acquisition in CHECK, and to a much smaller extent during KIDA evaluation. The quantitative measurement by KIDA is of limited value when the changes in the separate features are overruled by changes caused by lack of reproducible image acquisition. The studies described in chapters 2 to 7 actually all emphasize the urge for optimal image acquisition. In future studies improvement of image acquisition will increase the surplus value of KIDA significantly, since currently evaluation by the objective KIDA procedure is far more robust than the standardization of image acquisition procedures.

Clinical application of digital image analysis

Surplus value of KIDA in early osteoarthritis

In accordance with common clinical practice, participants in CHECK presented themselves with symptoms that were possibly related to OA. In individuals with symptoms, the question is whether they will actually develop the disease. So far, acquiring radiographs in this early phase of OA was not considered relevant since signs of structural damage were not detected with conventional K&L grading[21]. In clinical practice such a radiograph is actually only used as a reference to detect structural damage on a radiograph acquired later in time (e.g. based on K&L grading). By that time structural damage is already established, and that limits the treatment options to pain medication and eventually joint replacement. To advance prevention of OA it is of major interest to use the radiograph with limited joint damage from an early phase of the disease, and not to wait on a follow-up radiograph. Since digital analysis enables more detailed evaluation of separate radiographic features, this may aid in identification of subtle radiographic characteristics in an early phase. As such, digital analysis can identify characteristics that are important in onset and development of the disease. Subsequently these features can be evaluated for a relation with symptoms, to

predict radiographic progression, and/or to characterize phenotypes with specific course of disease, as reported in chapter 5 to 7.

Implementation of radiographic features of KIDA in CHECK

When separate quantitative features are used to assess the severity of structural damage, individual characteristics need to be taken into account since they influence the measurement. The size of the quantitative features is dependent on e.g. gender, height and weight (or body mass index; BMI), and characteristics of the joint (size of the bone and alignment). The similarities in radiographic characteristics were evaluated, and in chapter 4 a correlation in the digital measurements between joints was demonstrated. Within participants with hardly any joint damage according to K&L grading (0-I), radiographic features were shown to be correlated between joints. This was the case between contralateral joints, but also between hips and knees mutually (ipsilateral and diagonal). When a radiograph from one time point is analyzed, the interpretation of the radiographic features can be improved by taking into account the characteristics of the individual. In chapters 5 to 7 the radiographic characteristics of the contralateral joint were evaluated, next to actual measurement of radiographic features of the affected joint. By calculating the difference of the features between the contralateral joints, it can be assessed whether e.g. the joint space is narrowed or whether this is just an individual characteristic based on BMI and joint structure. By implementing a combination of the features of a joint and the difference between joints, the radiographic status of an individual as a whole can be assessed. E.g. next to a large feature value that indicates affection of a joint, a large difference between the contralateral joints implies unilateral affection while a small difference implies bilateral affection. Furthermore, radiographic features of other joints might be used as a reference for the structural damage of an individual.

Digital image analysis is also of value in the longitudinal evaluation of onset and progression of radiographic OA, since sensitivity to change is improved compared to qualitative grading. By measuring OA features on the radiograph at the time of complaints, the development of established radiographic OA could better be predicted by quantitative KIDA measurement than by qualitative K&L grading (chapter 6). Furthermore, separate features measured by KIDA were all found to progress in OA severity at different time points during five-year follow-up, which is difficult to assess using 0-IV grading according to K&L (chapter 5).

Evaluation of clinical characteristics and the relation with radiographic features

Although symptoms of OA are thought to originate from structural damage, detection of an association between radiographic and clinical OA is currently hampered by lack of methodological quality. Next to standardizing radiographic protocols and defining a population at risk for OA, the outcome measures of radiographic and clinical OA are of major importance (see chapter 2). To detect an association, the advances in detailed evaluation of radiographs need to be accompanied by more thorough assessment of symptoms.

Ideally, by using digital analysis the association can be improved since specific radiographic features can be related to specific symptoms, e.g. bone changes and pain, or osteophytes and functional limitations. The common approach however, is to relate radiographic characteristics of the joint to complaints of the individual as a whole without specifying the complaints of the affected joint[22]. Even when the complaints are assessed for a specific joint, detailed assessment is hampered when the pain is not localized but has a more regional or diffuse pattern[23]. Furthermore, the measures of symptoms are clearly subjective and are subject to considerable variation within individuals even during the day[24]. Moreover, during longitudinal evaluation the pain pattern may develop from constantly dull and aching in an early phase to more intense and unpredictable in a later phase[25]. Such an inconsistent course in symptoms[26] was also found in CHECK (chapter 5).
Despite the use of more precise evaluation by e.g. visual analogue scales (VAS) and the development of specific OA outcomes like the WOMAC (Western Ontario & McMaster Universities Osteoarthritis) index[27], symptoms remain difficult to assess[28]. Improvement of assessing pain and functional limitations are clearly needed and remain to be subject of study.

Heterogeneity of osteoarthritis populations versus identification of phenotypes
When individuals with symptoms are included in a clinical study or observational cohort like CHECK, it is acknowledged that they will differ in (the development of) symptoms and/or structural damage of OA. E.g. in some patients pain will be accompanied by bone changes[29], while others have symptoms that might be due to a combination of (local or systemic) inflammation and osteophyte formation[30]. When individuals with different characteristics and course of the disease are studied in one cohort, the development of radiographic and clinical characteristics will not be evident on average. The heterogeneity between individuals hampers the identification of characteristics that are important in onset and development of disease. This emphasizes that focus in future studies should shift to the identification of phenotypes within OA populations[5,31]. Identification of different phenotypes can advance the development and application of more targeted treatment strategies. E.g. an intense regimen that combines pain medication with anti-inflammatory drugs might be used in case of rapidly progressive OA, and treatment with bisphosponates might be effective in case of bone density involvement.
Recently the existence of phenotypes is more and more acknowledged and developments on how to identify phenotypes are ongoing. Preferably, phenotypes are identified early in the disease to enable prevention of irreversible joint damage. Currently phenotypes are not easily identified, particularly early in the disease. This was improved by using digital analysis, since this enabled detection of separate features that represent high (or low) risk of OA.
Chapter 7 describes that KIDA enabled the identification of five specific phenotypes of progression of radiographic knee OA in CHECK, which appeared of clinical relevance. The approach for phenotype identification is dependent on the purpose, and expectedly different

approaches result in the identification of different (additional) phenotypes. In chapter 7, cluster analysis was aimed at identifying radiographic progression phenotypes by exploring radiographic features at different time points. When performing cluster analysis with the use of cross-sectional data, different or similar phenotypes may be identified. Performing cluster analysis on joint level might enable the evaluation of whether progression of features in one knee is related to progression in another joint (contralateral knee or hip) of the same individual. Furthermore, clinical characteristics can be incorporated and the phase of disease can be studied. Clearly several approaches have to be used to identify the most distinct phenotypes.
The presently defined radiographic progression phenotypes in CHECK could be predicted on a population level by radiographic features analyzed at baseline. As such these phenotypes are already of value for clinical trial design, and future studies might even advance the identification of phenotypes that may be predicted in an early phase of the disease.

Steps forward and new challenges ahead

- The limited radiographic joint damage at baseline and the progression of separate quantitative radiographic features in the following years in CHECK enable evaluation of onset and development of early OA.
- Optimizing image acquisition will improve the value of radiography early in the development of OA, and will maintain radiography as the gold standard for evaluating structural tissue damage in OA.
- Quantitative (digital) image analysis with KIDA enables accurate bone density measurement on radiographs, which provides an additional detailed outcome measure for OA development.
- Radiographic features represent characteristics of an individual, next to osteoarthritis severity, which need to be taken into account during longitudinal evaluation.
- In early OA separate quantitative radiographic features enable the evaluation of progression of joint damage, of an association with clinical OA, and of the prediction of unfavorable outcome.
- Development and relations of radiographic features and clinical outcome differ over time, which needs to be considered when radiographic joint damage is evaluated in longitudinal studies.
- Phenotypes of very early progression of radiographic knee OA are identified, by evaluating separate quantitative features over time. This approach needs further development and validation in CHECK and other early OA, enabling specific patient selection to advance more targeted treatment.

References

1. Hunter DJ, Zhang W, Conaghan PG, et al. Systematic review of the concurrent and predictive validity of MRI biomarkers in OA. Osteoarthritis Cartilage. 2011 May;19(5):557-88.
2. Javaid MK, Lynch JA, Tolstykh I, et al. Pre-radiographic MRI findings are associated with onset of knee symptoms: the most study. Osteoarthritis Cartilage. 2010 Nov 11.
3. Nevitt MC. Definition of hip osteoarthritis for epidemiological studies. Ann Rheum Dis. 1996 Sep;55(9):652-5.
4. Buckland-Wright JC. Quantitative radiography of osteoarthritis. Ann Rheum Dis. 1994 Apr;53(4):268-75.
5. Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: an update with relevance for clinical practice. Lancet. 2011 Jun 18;377(9783):2115-26.
6. Wesseling J, Dekker J, van den Berg WB, et al. CHECK (Cohort Hip and Cohort Knee): similarities and differences with the Osteoarthritis Initiative. Ann Rheum Dis. 2009 Sep;68(9):1413-9.
7. Holla JF, Steultjens MP, Roorda LD, et al. Prognostic factors for the two-year course of activity limitations in early osteoarthritis of the hip and/or knee. Arthritis Care Res (Hoboken). 2010 Oct;62(10):1415-25.
8. Dacre JE, Coppock JS, Herbert KE, et al. Development of a new radiographic scoring system using digital image analysis. Ann Rheum Dis. 1989 Mar;48(3):194-200.
9. Neumann G, Hunter D, Nevitt M, et al. Location specific radiographic joint space width for osteoarthritis progression. Osteoarthritis Cartilage. 2008 Nov 12.
10. Oka H, Muraki S, Akune T, et al. Fully automatic quantification of knee osteoarthritis severity on plain radiographs. Osteoarthritis Cartilage. 2008 Apr 17.
11. Schmidt JE, Amrami KK, Manduca A, et al. Semi-automated digital image analysis of joint space width in knee radiographs. Skeletal Radiol. 2005 Oct;34(10):639-43.
12. Marijnissen AC, Vincken KL, Vos PA, et al. Knee Images Digital Analysis (KIDA): a novel method to quantify individual radiographic features of knee osteoarthritis in detail. Osteoarthritis Cartilage. 2008 Feb;16(2):234-43.
13. Conrozier T, Jousseaume CA, Mathieu P, 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 Sep;37(9):961-8.
14. Conrozier T, Tron AM, Mathieu P, et al. Quantitative assessment of radiographic normal and osteoarthritic hip joint space. Osteoarthritis Cartilage. 1995 Sep;3 Suppl A:81-7.
15. Conrozier T, Vignon E. Quantitative radiography in osteoarthritis: computerized measurement of radiographic knee and hip joint space. Baillieres Clin Rheumatol. 1996 Aug;10(3):429-33.
16. Maillefert JF, Sharp JT, Aho LS, et al. Comparison of a computer based method and the classical manual method for radiographic joint space width assessment in hip osteoarthritis. J Rheumatol. 2002 Dec;29(12):2592-6.
17. Wong AK, Inglis D, Beattie KA, et al. Reproducibility of computer-assisted joint alignment measurement in OA knee radiographs. Osteoarthritis Cartilage. 2008 Oct 17.
18. Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. Osteoarthritis Cartilage. 2007;15 Suppl A:A1-56.
19. Auleley GR, Duche A, Drape JL, et al. Measurement of joint space width in hip osteoarthritis: influence of joint positioning and radiographic procedure. Rheumatology (Oxford). 2001 Apr;40(4):414-9.
20. Buckland-Wright C. Protocols for precise radio-anatomical positioning of the tibiofemoral and patellofemoral compartments of the knee. Osteoarthritis Cartilage. 1995 Sep;3 Suppl A:71-80.
21. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. 1957 Dec;16(4):494-502.
22. Neogi T, Felson D, Niu J, et al. Association between radiographic features of knee osteoarthritis and pain: results from two cohort studies. Bmj. 2009;339:b2844.
23. Thompson LR, Boudreau R, Newman AB, et al. The association of osteoarthritis risk factors with localized, regional and diffuse knee pain. Osteoarthritis Cartilage. 2010 Jul 13.
24. Allen KD, Coffman CJ, Golightly YM, et al. Daily pain variations among patients with hand, hip, and knee osteoarthritis. Osteoarthritis Cartilage. 2009 Apr 17.
25. Hawker GA, Stewart L, French MR, et al. Understanding the pain experience in hip and knee osteoarthritis--an OARSI/OMERACT initiative. Osteoarthritis Cartilage. 2008 Apr;16(4):415-22.
26. Neogi T, Nevitt MC, Yang M, et al. Consistency of knee pain: correlates and association with function. Osteoarthritis Cartilage. 2010 Oct;18(10):1250-5.
27. Bellamy N. Pain assessment in osteoarthritis: experience with the WOMAC osteoarthritis index. Semin Arthritis Rheum. 1989 May;18(4 Suppl 2):14-7.
28. Barker K, Lamb SE, Toye F, et al. Association between radiographic joint space narrowing, function, pain and muscle power in severe osteoarthritis of the knee. Clin Rehabil. 2004 Nov;18(7):793-800.

29. Hunter DJ, Gerstenfeld L, Bishop G, et al. Bone marrow lesions from osteoarthritis knees are characterized by sclerotic bone that is less well mineralized. Arthritis Res Ther. 2009;11(1):R11.
30. van der Kraan PM, van den Berg WB. Osteophytes: relevance and biology. Osteoarthritis Cartilage. 2007 Mar;15(3):237-44.
31. Bierma-Zeinstra SM, Verhagen AP. Osteoarthritis subpopulations and implications for clinical trial design. Arthritis Res Ther. 2011 Apr 5;13(2):213.

Nederlandse samenvatting

Artrose

Artrose is een veel voorkomende gewrichtsaandoening die wordt gekenmerkt door pijn en stijfheid. Door deze klachten is er verminderde functie van het betreffende gewricht en dat zorgt voor beperkingen in het dagelijks leven. Artrose komt vooral voor in de grotere gewrichten; in de volwassen populatie heeft 6% klachten aan de knie en heeft 3% klachten aan de heup. Waarschijnlijk zijn er verschillende factoren nodig om artrose te ontwikkelen, waarvan hogere leeftijd en overgewicht de belangrijkste risicofactoren zijn. Genezing van artrose is (nog) niet mogelijk. De huidige behandeling is meestal gericht op het verminderen van pijn, het versterken van spieren rondom het gewricht en het bewerkstelligen van gewichtsverlies. Het is belangrijk om de diagnose artrose in een zo vroeg mogelijk stadium van de ziekte te stellen. In deze fase is de schade aan het gewricht namelijk nog beperkt en dan kan behandeling het best mogelijke resultaat opleveren.

Veranderingen op röntgenfoto's

De klachten van artrose worden veroorzaakt door veranderingen in verschillende weefsels van het gewricht. Er treedt schade op aan het kraakbeen dat de beide botuiteinden bedekt. Hierdoor neemt de afstand tussen de beide botuiteinden, de gewrichtsspleet, af. Ook treden er veranderingen op in de structuur (dichtheid) van het bot direct onder het kraakbeen en er ontstaan botuitgroeisels aan de randen van het gewricht. Al deze veranderingen zijn op röntgenfoto's zichtbaar. De meest gangbare methode om de ernst van artrose op röntgenfoto's te beoordelen is momenteel de Kellgren & Lawrence score. Hierbij wordt de ernst van artrose van het gehele gewricht beoordeeld met een stapsgewijze score op een schaal van 0 tot en met 4. Deze methode beoordeelt de verschillende weefsels van het gewricht als een geheel. Met deze 'grove' methode is het moeilijk om kleine veranderingen in het gewricht vast te stellen. Dit belemmert het stellen van de diagnose artrose in een vroege fase van de ziekte, waarin de veranderingen in de weefsels van het gewricht nog minimaal zijn.

Digitale analyse

In de laatste decennia zijn digitale analyse methoden ontwikkeld om veranderingen als gevolg van artrose op röntgenfoto's in meer detail te kunnen meten. Bij digitale analyse worden veranderingen in het gewricht gemeten op een continue schaal, in plaats van gescoord op een stapsgewijze schaal van 0 tot en met 4. Bovendien worden de verschillende weefsels van het gewricht apart beoordeeld, dit in tegenstelling tot de Kellgren & Lawrence score waarbij het gewricht in zijn geheel beoordeeld wordt. Digitale analyse is gericht op het eerder diagnosticeren van artrose, door het meten van subtiele veranderingen in de verschillende weefsels van het gewricht. Zo kan bijvoorbeeld versmalling van de gewrichtsspleet gemeten worden in (tienden van) millimeters, wat veel gevoeliger is dan het scoren op een schaal van 0 tot 4. Bovendien wordt met deze grove methode niet alleen de gewrichtsspleet beoordeeld, maar worden ook andere weefsels in de score meegenomen.

Cohort Heup en Cohort Knie (CHECK)

Het Cohort Heup En Cohort Knie (CHECK) is in 2001 door het Reumafonds gestart om een beter beeld te krijgen van de vroege fase van de ziekte artrose. Meer dan duizend deelnemers met pijn en/of stijfheid van de heup en/of knie worden onderzocht. Gedurende 10 jaar worden deze CHECK deelnemers gevolgd voor onder anderen het beloop van hun klachten, veranderingen in hun bloed en urine (biochemische markers voor artrose) en de ontwikkeling van veranderingen in het gewricht op röntgenfoto's. Tot op heden zijn de gegevens tot en met 5 jaar verzameld en momenteel zijn de eerste resultaten van CHECK beschreven door verschillende onderzoeksgroepen.

Aan het begin van de studie hadden de deelnemers nog geen (of nauwelijks) veranderingen in het gewricht, volgens de Kellgren & Lawrence score. De doelstelling van dit proefschrift is om te bestuderen of met behulp van digitale analyse van röntgenfoto's inderdaad kleine veranderingen in de verschillende weefsels van het gewricht eerder opgespoord kunnen worden. Als een relatie tussen deze subtiele veranderingen in het gewricht en de klachten gevonden kan worden, kan dit het inzicht in artrose verbeteren. Bovendien bevestigt dat het belang van het maken van röntgenfoto's in een vroege fase van artrose. De centrale vraagstelling in dit proefschrift is: kunnen subtiele veranderingen in weefsels van het gewricht vroeg in de ontwikkeling van artrose gemeten worden met behulp van digitale analyse?

Dit proefschrift

Dit proefschrift kan worden opgedeeld in drie delen. Het eerste deel betreft de methoden om veranderingen in de weefsels van het gewricht op röntgenfoto's te beoordelen. In het tweede deel wordt digitale analyse toegepast om de ontwikkeling van knie artrose in de CHECK deelnemers te onderzoeken. En in het derde deel is het belang van gestandaardiseerde procedures bij het maken van röntgenfoto's beschreven.

Meten van artrose veranderingen

In **hoofdstuk 2** is bestudeerd of er een eenduidige relatie bestaat tussen klachten en weefselveranderingen bij artrose. Hiervoor zijn gepubliceerde studies die deze relatie bestuderen, beoordeeld op criteria voor de kwaliteit van de gebruikte onderzoeksmethoden. Deze criteria bestaan uit de definitie van artrose om patiënten te selecteren voor een studie, de uitkomstmaten om klachten en weefselveranderingen te beoordelen en het gebruik van standaard protocollen tijdens het maken van röntgenfoto's. De conclusie van dit hoofdstuk is dat een relatie tussen klachten en schade vaker beschreven wordt in studies die voldoen aan meer criteria voor methodologische kwaliteit. Het belangrijkste criterium waaraan een onderzoek moet voldoen om de relatie te vinden is het gebruik van gestandaardiseerde uitkomstmaten om de klachten en weefselveranderingen te beoordelen.

CHECK voldoet aan de methodologische kwaliteitscriteria zoals beschreven in hoofdstuk 2. In de volgende hoofdstukken van dit proefschrift zijn daarom weefselveranderingen in de knie- en heupgewrichten van de CHECK deelnemers bestudeerd.

In de hoofdstukken 3 en 4 is onderzocht of digitale analyse gebruikt kan worden om kleine veranderingen in de weefsels van het gewricht te meten. Daarvoor is een methode gebruikt om knie röntgenfoto's te beoordelen, die is ontwikkeld in het Universitair Medisch Centrum Utrecht; KIDA (Knee Images Digital Analysis). Deze methode meet op continue schalen de veranderingen in het gewricht, in de verschillende weefsels en op verschillende plekken. KIDA is vergeleken met gangbare methoden die gebruik maken van stapsgewijze scores, van het gewricht in zijn geheel of van een beperkt aantal specifieke weefsels.

In **hoofdstuk 3** is de gevoeligheid voor het meten van veranderingen in de verschillende weefsels van het kniegewricht onderzocht. Om dit te bestuderen zijn de röntgenfoto's gebruikt die zijn gemaakt bij aanvang van CHECK en na twee jaar. De gevoeligheid om weefselveranderingen na twee jaar te vinden is vergeleken tussen meten met KIDA en scoren aan de hand van de Altman atlas. Met behulp van deze atlas worden de verschillende weefsels van het gewricht (kenmerken) gescoord op een grove schaal van 0 tot 3, dan wel 0 tot 1. In dit hoofdstuk wordt geconcludeerd dat de gevoeligheid voor verandering vergelijkbaar is tussen de digitale analyse methode KIDA en de Altman score. Overigens is slechts in een klein deel van de knieën een verandering gevonden, met beide methoden. Dit komt doordat weefselveranderingen zijn onderzocht in een vroege fase van de CHECK studie, waarin de deelnemers voornamelijk klachten hadden maar nog weinig veranderingen in het gewricht. Door het verbeteren van het maken van röntgenfoto's zal digitale analyse naar verwachting van grotere meerwaarde zijn dan tot nu is gebleken.

Bij het meten van veranderingen door artrose op twee opeenvolgende röntgenfoto's moet men er rekening mee houden dat er verschillen (variaties) kunnen optreden tijdens de meting. Bijvoorbeeld doordat de beoordelaar bij de digitale analyse niet altijd op precies dezelfde manier zal meten. Ook bij het maken van de röntgenfoto's kan er variatie optreden omdat een deelnemer niet op precies dezelfde manier voor het röntgenapparaat gaat staan (dit is onderzocht in hoofdstuk 8). Deze variatie is kleiner wanneer veranderingen in een gewricht binnen dezelfde persoon worden gemeten, dan wanneer gewrichten worden vergeleken tussen verschillende personen. Hoe een gewricht eruit ziet op een röntgenfoto is namelijk ook afhankelijk van karakteristieken van de persoon. Lange personen kunnen bijvoorbeeld een grotere gewrichtsspleet hebben en vrouwen hebben over het algemeen een lagere dichtheid van het bot. In **hoofdstuk 4** is onderzocht of kenmerken van het gewricht op röntgenfoto's samenhangen met kenmerken van een ander gewricht van een persoon. Er is bestudeerd of deze relatie verandert bij het ontstaan van artrose. In dit hoofdstuk is beschreven dat er een duidelijke relatie bestaat tussen de digitaal gemeten kenmerken op röntgenfoto's van de knie- en heupgewrichten van dezelfde persoon. De

grootte van de gewrichtsspleet van de ene knie hangt bijvoorbeeld samen met de grootte van de gewrichtsspleet van de andere knie, maar ook met de grootte van de gewrichtsspleet van de heupen van een persoon. Omdat deze relatie het sterkst is in gewrichten zonder weefselveranderingen, spelen karakteristieken van een persoon een belangrijke rol in de gemeten kenmerken. Een interessant resultaat is dat de relatie verdwijnt bij het ontstaan van artrose in één gewricht. Deze kennis is belangrijk bij het verder bestuderen van het ontstaan van veranderingen in gewrichten door artrose. Bij het bestuderen van artrose kenmerken op de röntgenfoto van een knie is het daarom zinvol om ook rekening te houden met de kenmerken op de röntgenfoto van de andere knie.

In het eerste deel van dit proefschrift is beschreven dat het belangrijk is om gestandaardiseerde methoden te gebruiken om weefselveranderingen door artrose te bestuderen. Met behulp van digitale analyse van röntgenfoto's kunnen veranderingen in de verschillende weefsels door artrose gemeten worden, waarbij rekening gehouden moet worden met karakteristieken van een persoon. Dit is in de hoofdstukken 5 tot en met 7 gedaan.

Digitale analyse in CHECK

In het tweede deel van dit proefschrift wordt de digitale analyse van röntgenfoto's van de knie (KIDA) toegepast in een studie naar vroege artrose; CHECK. In **hoofdstuk 5** wordt beschreven of metingen met digitale analyse kenmerken van knie artrose representeren. De ontwikkeling van deze kenmerken van artrose zijn geëvalueerd door de kenmerken te meten bij aanvang van het cohort, na twee jaar en na vijf jaar. Over het algemeen werden bij deze personen met klachten in een vroege fase van artrose duidelijke veranderingen in de verschillende weefsels van het gewricht gevonden. Deze veranderingen duiden op een ontwikkeling van artrose.

In **hoofdstuk 6** zijn de aparte kenmerken op de knie röntgenfoto bij aanvang van CHECK gebruikt om te onderzoeken of deze kenmerken bijdragen, naast de klachten en de karakteristieken van de persoon, aan het voorspellen van het beloop van de ziekte. Dit kan van meerwaarde zijn in de klinische praktijk, omdat patiënten met een grote kans op een slechter beloop anders behandeld zouden kunnen worden dan patiënten met een hoge kans op een gunstiger beloop van de ziekte. De resultaten tonen aan dat de aanwezigheid van een kleine gewrichtsspleet en het ontstaan van botuitgroeisels (osteofyten) op de röntgenfoto zeer vroeg in het ziekteproces (bij aanvang van CHECK) het ontstaan van duidelijke weefselveranderingen na vijf jaar kunnen voorspellen.

Het wordt verondersteld dat het beloop van artrose verschillend is voor verschillende personen. In **hoofdstuk 7** is bestudeerd of er subtypes van artrose bestaan, die verschillend zijn wat betreft de veranderingen in de weefsels van het gewricht. In dit hoofdstuk zijn de metingen met digitale analyse gebruikt op de knie röntgenfoto's bij aanvang van CHECK en

na twee jaar en na vijf jaar. Er is geconcludeerd dat verschillende subtypes van knie artrose bestaan. Er zijn personen gevonden met ernstige ontwikkeling van alle aparte kenmerken van artrose. Daarnaast zijn er personen waarbij veranderingen in het gewricht vroeg plaats vinden (in de eerste twee jaar van de studie), en andere personen die deze verandering pas later laten zien (vanaf twee tot vijf jaar). Een specifiek subtype van artrose bestaat uit personen die vooral veranderingen in de dichtheid van het bot hebben. Maar er zijn ook personen in CHECK die geen weefselveranderingen in het gewricht hebben bij aanvang en deze ook niet ontwikkelen gedurende de eerste vijf jaar van de studie.

Dit deel van het proefschrift beschrijft dat digitale analyse van röntgenfoto's gebruikt kan worden in de CHECK studie. In een vroege fase van artrose worden aparte kenmerken gemeten die zich gedurende de tijd verder ontwikkelen. Ook kunnen deze kenmerken gebruikt worden om het beloop van de ziekte te voorspellen en om subtypes van artrose te omschrijven.

Röntgenfoto's

In het laatste deel van dit proefschrift is het optreden van variaties tijdens het maken van röntgenfoto's onderzocht. Voor de CHECK studie wordt een protocol gebruikt met voorgeschreven instellingen van het röntgenapparaat en de positie van het gewricht tijdens het maken van de röntgenfoto. Bij het onderzoeken van artrose veranderingen tussen aanvang van CHECK en na bijvoorbeeld twee jaar, kunnen er kleine variaties optreden bij het maken van de röntgenfoto's. Als deze variaties groter zijn dan de weefselveranderingen, dan belemmert dit het vinden van veranderingen door artrose. Wanneer deze variaties beperkt kunnen worden, kan digitale analyse gebruikt worden om artrose veranderingen te meten. In **hoofdstuk 8** is de invloed van kleine variaties in de positie van de knie tijdens het maken van de röntgenfoto op de meting met digitale analyse bestudeerd. De resultaten laten zien dat variaties in de knie positie de uitkomst van de digitale analyse inderdaad kunnen beïnvloeden. Het is daarom belangrijk het protocol met voorgeschreven instellingen voor het maken van röntgenfoto's zo goed mogelijk te volgen.

Tot op heden kunnen veranderingen in de botdichtheid door artrose niet goed gemeten worden op röntgenfoto's en daarom zijn speciale röntgenfoto's nodig om dit aan te tonen. In **hoofdstuk 9** is onderzocht of veranderingen in de dichtheid van het bot direct onder het kraakbeen gemeten kan worden met behulp van digitale analyse op gewone röntgenfoto's. Ook bij deze meting kunnen variaties tijdens het maken van de röntgenfoto de uitkomst van de digitale analyse beïnvloeden. In dit hoofdstuk is de invloed onderzocht van variaties in de instellingen van de röntgenapparatuur. Naast de instellingen tijdens het maken van de röntgenfoto, zijn ook computer instellingen onderzocht die de röntgenfoto's optimaliseren, maar tegelijkertijd de meting van botdichtheid kunnen beïnvloeden. Er is geconcludeerd dat instellingen van de röntgenapparatuur de meting van botdichtheid inderdaad beïnvloeden. Het is dus belangrijk het protocol voor het maken van röntgenfoto's zo goed mogelijk te

volgen en zo standaard mogelijk te houden. De invloed van de röntgen- en computer instellingen is echter kleiner dan de veranderingen die in CHECK optreden door artrose. Daarom is geconcludeerd dat digitale analyse gebruikt kan worden om veranderingen in botdichtheid te meten, waardoor aanvullende speciale röntgenfoto's niet meer nodig zijn.

Het laatste deel van dit proefschrift beschrijft dat het volgen van een standaard protocol voor het maken van röntgenfoto's belangrijk is. Als dit in acht wordt genomen, kan digitale analyse gebruikt worden om kenmerken vanaf een vroege fase van artrose op röntgenfoto's te meten.

Samengevat is digitale analyse van röntgenfoto's, zoals met KIDA, van meerwaarde in het vinden van weefselveranderingen in vroege artrose. Daarmee is digitale analyse een belangrijk instrument om weefselveranderingen in CHECK te bestuderen, waarbij optimale standaardisatie bij het maken van de röntgenfoto's van groot belang is.

Dankwoord

Mijn proefschrift is klaar! Het was een bijzondere en leerzame tijd waarin ik veel mooie herinneringen heb verzameld. Op deze plek wil ik iedereen die daaraan heeft bijgedragen, bedanken voor de inhoudelijke hulp, persoonlijke begeleiding en/of tijd voor ontspanning.

Een flink aantal promotoren en co-promotoren heeft zich ingezet voor dit proefschrift.
Floris, het begon met een dynamisch sollicitatie gesprek en zo is de energie tussen ons eigenlijk altijd gebleven. We wisten elkaar altijd te prikkelen, met een knallend proefschrift als resultaat! Ik ben onder de indruk van jouw onuitputtelijke bron van verfrissende ideeën en ik wil je bedanken voor je betrokkenheid. Max, het was zinvol om tijdens de CHECK meetings te sparren over de verschillende methodes om de (röntgen) beelden van alle deelnemers te evalueren. Jij hield altijd de toepasbaarheid en de relevantie in de gaten. Hans, op onze veelzijde afdeling waar zoveel verschillende disciplines elkaar kruisen, bewonder ik hoe jij zowel de details als het grote geheel in het oog kan houden.
Koen, wat moest ik aan het begin wennen aan je programmeertaal! Maar het is uiteindelijk gelukt om het technische met het klinische te combineren. Dank voor je inwijding in deze binaire wereld en dank voor je hilarische verhalen en gezelligheid! Anne Karien, bij jou kon ik altijd terecht, bedankt dat ik je hand mocht vasthouden tijdens deze achtbaan van emoties! Je relativeringsvermogen en rust waren subliem, ik ga je persoonlijkheid zeker missen.
Paco, ook jij verdient een prominente plek. Hoewel je geen co-promotor was, heb je een geweldige bijdrage geleverd aan een groot deel van dit proefschrift. Jij bent wat mij betreft een typische onderzoeker; lang van stof en altijd nieuwsgierig. Lastig is alleen dat je soms verdwaalt in allerlei interessante bevindingen en dan de oorspronkelijke vraag vergeet. Je wordt nog wel eens een warrige professor, dank voor al je enthousiasme en geweldige hulp!

Zonder CHECK was dit proefschrift er niet geweest! CHECK is er dankzij de financiering van het Reumafonds en dit grote cohort wordt gecoördineerd door Janet, daarbij geweldig gesteund door Joline, Mariska en Jeffrey.
Zonder data geen CHECK! Bedankt aan alle deelnemers voor hun geweldige inzet, door het jaarlijks invullen van vragenlijsten en het frequent bezoeken van het ziekenhuis gedurende de afgelopen (en de komende) jaren! De enorme hoeveelheid data kan verzameld worden dankzij de toewijding van vele röntgenlaboranten, artsen en andere medewerkers in de tien deelnemende ziekenhuizen verspreid door Nederland: Academisch Ziekenhuis Maastricht, Erasmus Medisch Centrum Rotterdam, Reade (Jan van Breemen Instituut)/ Vrije Universiteit Medisch Centrum Amsterdam, Kennemer Gasthuis Haarlem, Martini Ziekenhuis/ Paramedisch Centrum voor Reumatologie en Revalidatie Groningen, Medisch Spectrum Twente Enschede/ Twenteborg Ziekenhuis Almelo, St. Maartenskliniek Nijmegen, Leids Universitair Medisch Centrum, Universitair Medisch Centrum Utrecht, Wilhelmina Ziekenhuis Assen. Gerard, bedankt dat ik altijd snel bij je terecht kon als ik weer een specifiek deel van die enorme hoeveelheid röntgenfoto's wilde hebben. Verder de medewerkers van het Julius Centrum dank voor het beheren en aanleveren van de data.

Zonder de metingen van de röntgenfoto's met digitale analyse was er bijna geen hoofdstuk van dit proefschrift overgebleven! Marja, bedankt voor het meten van de knie foto's met KIDA, het waren er duizenden! Prof. Vignon, thank you for welcoming us in Lyon and for your help with the Holy measurements. Petra, bedankt voor je hulp met de heup metingen en succes met je eigen promotie. De CHECK symposia en onderzoeksmeetings waren altijd inspirerend en ik wil in het bijzonder de leden van projectgroep II bedanken voor hun input.

Zonder de geweldige en diverse groep AIO's, had ik een flink aantal mooie herinneringen moeten missen! Als het even tegenzat was er altijd wel iemand met een luisterend oor, een goede tip en het gevoel van herkenning, of er werd gewoon geborreld om dingen te relativeren. Een (verre) buur was regelmatig beter dan een goede vriend! De AIO-etentjes en –activiteiten waren altijd ontspannen en gezellig.
Om te beginnen met mijn fantastische voorgangers. Nathalie, kleine nauwkeurige wervelwind, wat ben je toch altijd attent! Sarita, wat heerlijk om mijn liefde voor sneakers met jou te delen. En dan ben je ook nog eens geweldig, grappig en goed gezelschap tijdens congres, teambuilding, Tivoli feestjes enzovoorts! Femke, het was zoveel fijner toen jij er nog was om me bij te vallen als ik weer eens met de botte bijl hakte. Maar ik mis ook je gezellige borrel talent; ik kijk nu al uit naar een mooi feestje, want goed voorbeeld doet volgen! Marije, jij was de stabiele factor tijdens mijn hele promotie onderzoek! Dank voor het zijn van mijn nuchtere kamergenoot en voor het aanhoren van al (!) mijn verhalen, heerlijk om mijn liefde voor sport met jou te kunnen delen. Jos, jouw kijk op het leven en jouw lach zijn aanstekelijk, hopelijk heb ik daar een klein beetje van kunnen overnemen.
Dan nog de vele AIO's die nog volop aan het analyseren en schrijven zijn; zet hem op! Tineke, over twee dagen is het jouw moment, geniet ervan! Angela, gekke Taiwanese schoenen koningin, respect voor je harde werken en doorzettingsvermogen. Ik ga je hilarische verhalen en singstar kwaliteiten zeker missen! Monique, bewonderenswaardig dat jij altijd een plan hebt en de komende jaren al hebt uitgestippeld voor jezelf. Gelukkig zorg je ook voor genoeg ontspanning buiten het werk, al kunnen wij je daar niet altijd op betrappen! En dan onze vier dokters; Marlies, wat was jij top als kamergenoot tijdens de laatste stressvolle maanden! We hebben hard gewerkt, maar ik heb ook veel met je gelachen tijdens een potje frustratie-voetbal of als je weer eens op je handen in de kamer stond! Erwin, ook al zo'n type met een missie, het gaat je lukken om de biomarkers verder op de kaart te zetten en zeker ook met het vervolg. Laurens, zet hem nog even op met je experimenten voordat je weer de kliniek in gaat. Ik ga je langdradige verhalen en vooral de bijbehorende motoriek missen! En Karen, fanatiekeling, het was altijd leuk om allerlei sportevenementen met je te bespreken. Jouw studie gaat zeker weer mooie resultaten opleveren voor de artrose wereld. Fréderique, bij jou schiet het ook op met alle IL's, terwijl ik jou nog steeds als onze slimme en schattige studente zie! Sandhya, het gaat je lukken om je modellen begrijpelijk te maken binnen onze afdeling en daarbuiten. Succes met alle veranderingen die op komst zijn. Maud, ook van jou weet ik dat je de epidemiologie binnen de afdeling op de kaart blijft zetten. Al het uitzoekwerk gaat weer veel opleveren voor het RA

onderzoek. Maarten, in mijn promotie-stress heb ik weinig mee gekregen van jouw werk, maar dankzij je enthousiaste sollicitatie-presentatie weet ik dat het goed zit met het immunologie onderzoek!

Alle anderen binnen de onderzoeksgroep reumatologie en klinische immunologie wil ik ook bedanken voor hun betrokkenheid bij mijn onderzoek. Joël, toen we bureau-buren waren, merkte ik dat je een stereotype onderzoeker bent; hardop denken, artikelen kwijt raken, lang(zaam) van stof en droge humor. Hoewel we totaal verschillend onderzoek deden, kwam je regelmatig met verrassend verfrissende en goede input. Simon, jij bent altijd goed gezelschap en lekker recht voor zijn raap. Dank voor je feedback op presentaties en voor de mogelijkheid om altijd binnen te lopen om te sparren over promotie-gerelateerde (en computer) problemen!
Hoewel ik weinig op het laboratorium te zoeken had, op wat hydroxapatite avonturen na, voelde ik me hier altijd welkom. Dit dankzij onze veelzijdige analisten: Marion, Kim, Angelique, Arno, Karin, Dorien en Katja. Marieke, jouw hulp en aanwezigheid bij de uitstapjes via de anatomie naar de röntgenafdeling waren niet alleen erg gezellig, maar ook erg zinvol! Ook de andere medewerkers van de afdeling wil ik bedanken voor hun input tijdens research besprekingen en voor de leuke afdelingsuitjes, nieuwjaarsborrels etc.

Ook buiten de afdeling reumatologie en klinische immunologie zijn er veel verschillende mensen in het UMC geweest die mij geholpen hebben. Omdat bij een proefschrift over radiologische artrose veel röntgenfoto's horen, was ik regelmatig op de röntgenafdeling te vinden om de meest bijzondere series foto's te maken. Hartelijk dank aan Martine en Henk en aan alle laboranten voor het beschikbaar stellen van ruimte en voor de hulp met de apparatuur en het branden van CD's!
Tineke, het was leuk en leerzaam om jou als studente te begeleiden. We hebben, mede dankzij jouw expertise op de röntgenafdeling, een bijzonder en leuk project uitgevoerd dat hopelijk resulteert in een mooie publicatie! Willem, ik hoefde maar heel eventjes te wennen aan jouw werkplek, want jouw enthousiasme was aanstekelijk. Dankzij je geweldige ideeën en aanpakkers-mentaliteit hebben we een serie mooie experimenten gedaan!
Hugo, hoewel ik me aan het begin wel eens afvroeg wat onze uitstapjes naar de DEXA-scanner en de röntgenafdeling op zouden leveren, is het dankzij jouw geduld en doortastendheid gelukt om een lastig fysisch vraagstuk begrijpelijk te maken! Wilbert, soms maken je hilarische woordgrappen het lastig om de concentratie vast te houden, maar jij hebt het talent om zelfs de ingewikkeldste dingen eenvoudig te laten lijken!

En dan verdienen nog al mijn lieve, grappige en waardevolle vrienden en familieleden een bijzondere vermelding voor de ontspannen en fijne momenten in de afgelopen (en komende!) jaren.

Allereerst de DVO meisjes; alle dingen die we samen beleefd hebben zijn de basis voor een geweldige vriendschap die al jarenlang teruggaat. Gelukkig houden we van tradities, maar ook van vernieuwing, waardoor (onder anderen) ons jaarlijkse weekendje weg garant staat voor gezelligheid, goede gesprekken en veel nieuwe herinneringen!

Dan mijn jaarclubgenootjes, onze dubbele studietijd zit er inmiddels wel echt op, maar gelukkig lukt het nog regelmatig om samen (pasta) te eten en bij te kletsen!
Verder mijn studiegenootjes Sieke, Stephanie en Karen, we hebben uren lang samen praktische vaardigheden geoefend, biomechanica sommen gemaakt en met matlab practica geworsteld. Dank voor de support in tentamen tijd en de gezelligheid daar buiten!
James and all the others from the Biomechanics Research Laboratories in Ann Arbor, thanks for the hospitality and help during my research project. Also thanks to everybody in the co-op who made it such a great time!

En wat kwam mijn besluit om te gaan voetballen op een goed moment, het was heerlijk om mijn werk te kunnen vergeten door tegen een bal aan te trappen! Bovendien heeft het ook nog eens veel gezelligheid met de dames (en heren) met zich mee gebracht.
Maaike, attente lieverd, we kwamen elkaar overal tegen; op het veld, bij de activiteiten commissie en in het UMC voor koffie met lekkers. Dat is straks toch niet allemaal voorbij?! Debbie, ontzettende fladderaar, ik ben wel eens jaloers op jouw super positieve kijk op de wereld, al schiet ik wel in de stress als ik aan jouw drukke leven denk! Bienvenida, topper, dank voor het zijn van mijn exponent!

Lieve Bijleveldjes, wat vijftien jaar geleden begon als een oppas-adres, is een tweede thuis geworden. Hoewel niet meer zo vaak in Bennekom, is een goede buur vaak beter dan een verre vriend (en dat geldt niet alleen voor mij). De (achter)deur staat bij jullie altijd open en één biertje (uit de fles!) worden er altijd twee. Bedankt dat ik mijn verhaal altijd bij jullie kwijt kan. Tineke bedankt voor het zijn van mijn surrogaat-moeder en Harry bedankt dat je me altijd op scherp weet te zetten! En leuk hoor om er zo twee knappe broertjes bij te krijgen!

Joop en Henny, jullie Drentse humor en nuchterheid is geweldig en relativerend! Ik wil jullie bedanken voor de betrokkenheid en lieve zorgen voor ons allemaal in de afgelopen periode.

Tenslotte, lieve pap en mam, jullie lieten me vrij, maar waren er op de achtergrond altijd voor me. Daar heb ik mijn zelfstandigheid en ‘aanpakken’ aan te danken! Dank voor jullie onvoorwaardelijke vertrouwen in mijn keuzes en de goede afloop. Mam, ik moet je missen, maar ik draag een groot deel van jou bij me en ik weet dat je trots bent. Pap, het praatgrage

heb ik zeker niet van een vreemde. Maar ik wil je juist ook bedanken voor je luisterend oor, je kalmte, je advies en je (veronderstelde) kennis van allerlei zaken!
Lieve zussen, we zijn zo verschillend, maar ondertussen ook zo hetzelfde. Ik vind het erg bijzonder om jullie straks als paranimfen bij me te hebben! Dorien, attente diva, hoewel je wat verder weg woont, ben je zeker niet minder betrokken. Ik bewonder je warmte en doorzettingsvermogen! Irene, stoere chick, toen ik in Utrecht ging studeren kon ik altijd bij je terecht om te logeren, mee uit te gaan of lekker een vorkje mee te prikken. Gelukkig kan dat nog steeds en het brengt inmiddels andere voordelen met zich mee dat ik dichtbij woon!
Lieve zwagers, dank voor het gewoon er zijn al die afgelopen jaren. Marc, het is een knallende omslag geworden en Paul we maken er een knallend feest van!

Lief, wat een jaar! Het was druk, maar ook zo fijn en gezellig. Dank voor je rust, je geduld en je geweldige support, nu is het tijd voor ons!

Curriculum vitae

Margot Bernadette Kinds was born on June 12th 1981 in Ede, the Netherlands. In 1999 she graduated secondary school at the Marnix College in Ede.
In the same year she started her study Physical Therapy at the University of Applied Sciences Utrecht (Hogeschool Utrecht), and she obtained her Bachelor of Health (BH) degree in January 2004. For this study, next to practical internships, a literature study was performed on the characteristics of bi-articular muscles.
In September 2004 she started the study Human Movement Sciences at the VU University Amsterdam with 'rehabilitation' as the field of study, and she obtained a Master of Science (MSc) degree in January 2007. A bachelor research project was performed in collaboration with K van Stein Callenfels at 'Revalidatiecentrum de Hoogstraat' in Utrecht, under supervision of prof.dr. LHV van der Woude. For this project technical and practical tests were developed and performed to select a wheelchair that was most suitable for the variety of rehabilitants.
For her master thesis, she did a research internship from June until December 2006 at the Biomechanics Research Laboratories in Ann Arbor, Michigan, USA. Under supervision of prof. JA Ashton-Miller, it was studied whether lengthening contraction training of leg extensor muscles was feasible to prevent falling in elder women.

In May 2007 she started as a PhD student at the department of Rheumatology & Clinical Immunology in collaboration with the Image Sciences Institute at the University Medical Center Utrecht under direct supervision of dr. ACA Marijnissen, dr.ir. KL Vincken, and prof.dr. FPJG Lafeber. For digital image analysis of radiographic features of the hip joint she collaborated with prof.dr. EP Vignon from the department of Rheumatology at Lyon-Sud University Hospital in France, and she visited his research group twice; in February 2008 and February 2009.
During her PhD trajectory she followed the post initial Master of Epidemiology at EpidM (EMGO) of the VU University Medical Center Amsterdam, and in July 2011 she obtained her MSc degree.