

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

Margot B. Kinds,^{1,2}

Anne C.A. Marijnissen,¹

Sita M.A. Bierma-Zeinstra,^{3,4}

Johannes W.J. Bijlsma,¹

Koen L. Vincken,² Max A. Viergever,²

Floris P.J.G. Lafeber,¹

Paco M.J. Welsing^{1,5}

¹Rheumatology and Clinical Immunology,

²Image Sciences Institute, University

Medical Center (UMC) Utrecht; ³General

Practice, ⁴Department of Orthopaedics,

University Medical Center Rotterdam;

⁵Julius Center for Health Sciences and

Primary Care, University Medical Center (UMC) Utrecht, the Netherlands

Abstract

For more tailored treatment of osteoarthritis it is worthy 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). Among individuals with early signs related to osteoarthritis (Cohort Hip and 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 measurement and Altman scales. 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. 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 ($\kappa \leq 0.20$). 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

Correspondence: Floris P.J.G. Lafeber, University Medical Center Utrecht, Rheumatology and Clinical Immunology, PO Box 85500, 3508 GA Utrecht, the Netherlands.

Tel. +31.88.7558521 - Fax: +31.30.252374

E-mail: f.lafeber@umcutrecht.nl

Key words: osteoarthritis, knee, radiography, progression, digital analysis.

Acknowledgements: we acknowledge Marja Lafeber (ML) for KIDA measurements.

CHECK is funded by the Dutch Arthritis Association on the lead of a steering committee comprising 16 members with expertise in different fields of OA, chaired by J.W.J. Bijlsma, Prof, MD and coordinated by J. Wesseling, MSc. Involved are: Academic Hospital Maastricht; Erasmus Medical Center Rotterdam; 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.

Contributions: MK, AM, JB, MV, PW, FL, conception and design of this study; MK, AM, SB, KV, FL, PW, analysis and interpretation of data.

Conflict of interests: the authors declare no conflict of interests.

Funding: this work was supported by the Dutch Arthritis Association.

Received for publication: 22 July 2011.

Revision received: 13 February 2012

Accepted for publication: 30 March 2012.

This work is licensed under a Creative Commons Attribution NonCommercial 3.0 License (CC BY-NC 3.0).

©Copyright M.B. Kinds et al., 2012

Licensee PAGEPress, Italy

Rheumatology Reports 2012; 4:e9

doi:10.4081/rr.2012.e9

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) results in a higher sensitivity for radiographic changes (during two-year follow-up) than qualitative grading by the Altman atlas.¹⁵ Participants 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.

Materials and Methods

Study design and participants

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 study was approved by the medical ethics committees of all participating hospitals 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 regular quality visits on compliance were performed by an overall coordinator 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 aluminium step wedge for KIDA).

Knee images digital analysis

Parameters of radiographic OA were quantitatively measured on a continuous scale by KIDA,¹⁵ of which a screenshot is depicted in Figure 1. 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. The minimum joint space width was measured as the smallest distance between femur and tibia. Osteophyte area (in mm^2) was determined at the lateral and medial femur and lateral and medial tibia separately. Bone density (in mm/Al) was determined in the femur and tibia separately, and grey 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 and all baseline and two-year follow-up knee radiographs were analyzed by one experienced observer (M.L.) in random order unaware of the patient characteristics.¹⁵ 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 and Lawrence grade

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

Statistical analysis

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 radi-

ographic 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,

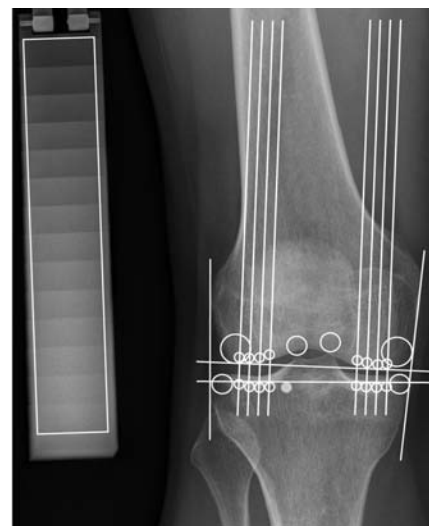


Figure 1. Screenshot of Knee Images Digital Analysis (KIDA). The framework determines the joint dimensions and joint angle. The smaller circles are used for joint space width (JSW) and bone density evaluation in the lateral and medial femur and tibia; the shortest distance between each pair of circles represents JSW. The medium circles are placed on top of the lateral and medial eminence; the shortest distance to the upper horizontal line represents the eminence height. The larger circles are used for osteophyte area measurement in the lateral and medial femur and tibia; the bone area outside these circles is considered osteophyte area. The minimum JSW is determined as the smallest distance between femur and tibia.

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 were used.²¹ 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

Among 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

The standardized response mean (SRM) was not evidently greater for the separate radiographic parameters of KIDA measurement as compared to the corresponding parameters of Altman grading (Table 1). The SRM for KIDA was largest for medial joint space narrowing (0.34) and bone density at the tibia (0.36) and femur (0.23). For Altman grading the SRM was largest for medial joint space narrowing (0.25) and for osteophytes at the lateral tibia (0.27), which were smaller than the largest SRM for KIDA. According to Cohen's approach of effect size, the effect sizes were considered small (trivial: <0.20, small: 0.20-0.50, moderate: 0.50-0.80, large: >0.80).²² Lateral joint space narrowing as measured with KIDA was found to decrease during two-year follow-up, which represents a widening of the lateral joint space.

Smallest detectable difference for knee images digital analysis 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. 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 2A-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. 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 2A-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 high-

Table 1. Standardized response mean for knee images digital analysis measurement and Altman grading.

	No. knees	KIDA change t2-t0			Altman change t2-t0		
		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

KIDA, knee images digital analysis; SRM, standardized response mean; SD, standard deviation; t2, two-year follow-up, t0, baseline.

Table 2. Smallest detectable difference for knee images digital analysis parameters in selection of unchanged knees (Altman parameters grade 0).

	No. knees	KIDA difference (t2-t0)		
		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

KIDA, knee images digital analysis; SDD, smallest detectable difference; SD, standard deviation; t2, two-year follow-up, t0, baseline.

er 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 (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 knee images digital analysis, Altman atlas, and Kellgren & Lawrence grading

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. 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), respec-

tively. 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 (4-88%; range for the distinct parameters) than *vice versa*

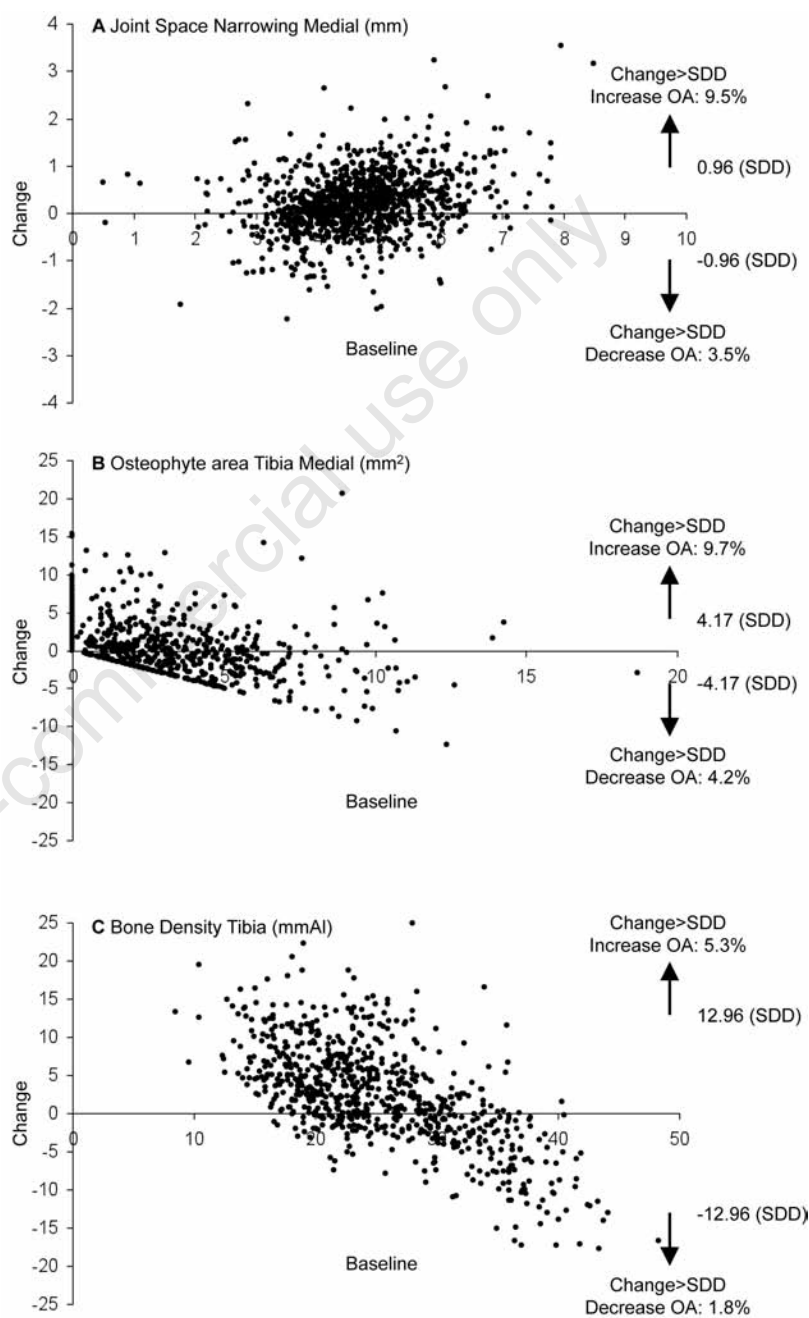


Figure 2. Knee Images Digital Analysis measurement: a knee change (two-year follow-up -baseline) plotted against baseline value (each dot depicts a knee) for A) Joint Space Narrowing Medial (mm) of 1081 knees; B) Osteophyte area Tibia Medial (mm²) of 1035 knees; C) Bone Density Tibia (mmAl) of 733 knees.

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 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

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

	No. knees	Increase OA		Decrease OA	
		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%*

OA, osteoarthritis; KIDA, knee images digital analysis. * Number of knees with increase or decrease significantly different (McNemar $P < 0.05$) between KIDA measurement and Altman grading.

Table 4. Agreement between knee images digital analysis and Altman grading in radiographic change (two-year follow-up baseline).

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

%: percentage of available knees per parameter. *Agreement between knee images digital analysis and Altman grading.

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 and differences in knee joint position are likely to introduce variation in the objective measurement of KIDA parameters.^{24,25} 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 3A-D. When the change in a specific radiographic parameter larger than the SDD is measured, this is considered an actual change. However, for smaller changes in radiographic measurements it is uncertain whether this is a real change or whether this is due to the random measurement error (partly) due to variations in knee positioning. Furthermore, as seen in Figure 2A-C regression to the mean occurred, with a smaller value at t0 being related to a larger value at t2 (and larger change) and *vice versa*, which emphasizes the difficulty with reproducible positioning. For example, when the t0 radiographs are acquired with more flexion than according to the protocol, it is likely that the knee will be more extended (in the position according to protocol) during acquisition at t2; which relates to changes in apparent JSW due to acquisition flaws and not to cartilage thickness changes. 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 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. In fact, 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



Figure 3. 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) osteoarthritis increase according to Altman and no change according to KIDA; C-D) no change according to Altman and osteoarthritis increase according to KIDA.

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 might be identified slightly earlier in the disease.²⁷ Irrespective of the method used evaluation of distinct parameters might enable the identification of (independent) progression of specific radiographic features.

References

- McAlindon T, Dieppe P. Osteoarthritis: definitions and criteria. *Ann Rheum Dis* 1989; 48:531-2.
- 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;65: 1363-7.
- Hunter DJ, Felson DT. Osteoarthritis. *BMJ* 2006;332:639-42.
- 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;5:87-97.
- Buckland-Wright JC. Quantitative radiography of osteoarthritis. *Ann Rheum Dis* 1994;53:268-75.
- Nevitt MC. Definition of hip osteoarthritis for epidemiological studies. *Ann Rheum Dis* 1996;55:652-5.
- Kellgren JH, Lawrence JS. Radiological assessment of osteoarthrosis. *Ann Rheum Dis* 1957;16:494-502.
- Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. *Osteoarthritis Cartilage* 2007; 15supplA:A1-56.
- 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: 370-6.
- Vignon E, Conrozier T, Hedio Le Graverand MP. Advances in radiographic imaging of progression of hip and knee osteoarthritis. *J Rheumatol* 2005;32:1143-5.
- 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;31:1588-97.
- 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;9:281-6.
- Gossec L, Jordan JM, Mazucca 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;16:742-8.
- Mazucca SA, Brandt KD, Katz BP, et al. Comparison of quantitative and semi-quantitative indicators of joint space narrowing in subjects with knee osteoarthritis. *Ann Rheum Dis* 2006;65:64-8.
- 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;16: 234-43.
- 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; 68:1413-9.
- Buckland-Wright C. Protocols for precise radio-anatomical positioning of the tibiofemoral and patellofemoral compartments of the knee. *Osteoarthritis Cartilage* 1995;3supplA:71-80.
- 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;26:2664-74.
- 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;62:924-31.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307-10.
- Dwyer AJ. Matchmaking and McNemar in the comparison of diagnostic modalities. *Radiology* 1991;178:328-30.
- Cohen J. Statistical power analysis for the behavioral sciences; 2nd edition. Hillsdale: Lawrence Erlbaum Associates Inc; 1988.
- Botha-Scheepers S, Watt I, Breedveld FC, Kloppenburg M. Reading radiographs in pairs or in chronological order influences radiological progression in osteoarthritis. *Rheumatology (Oxford)* 2005;44:1452-5.
- 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;52:1225-30.
- Mazucca 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;44:1786-94.
- Mazucca 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;46:109-13.
- Bierma-Zeinstra SM, Verhagen AP. Osteoarthritis subpopulations and implications for clinical trial design. *Arthritis Res Ther* 2011;13:213.