Title

Validity and sensitivity to change of three scales for the radiographic assessment of knee osteoarthritis using images from the Multicenter Osteoarthritis Study (MOST)

Journal

Osteoarthritis and Cartilage, 23(9)

ISSN

1063-4584

Authors

Sheehy, L
Culham, E
McLean, L
et al.

Publication Date

2015-09-01

DOI

10.1016/j.joca.2015.05.003

Peer reviewed

HHS Public Access

Author manuscript

Osteoarthritis Cartilage. Author manuscript; available in PMC 2016 April 14.

Published in final edited form as:

Osteoarthritis Cartilage. 2015 September ; 23(9): 1491–1498. doi:10.1016/j.joca.2015.05.003.

Validity and sensitivity to change of three scales for the radiographic assessment of knee osteoarthritis using images from the Multicenter Osteoarthritis Study (MOST)

Lisa Sheehy[1,2], **Elsie Culham**[3], **Linda McLean**[2], **Jingbo Niu**[4], **John Lynch**[5], **Neil A. Segal**[6,7], **Jasvinder A. Singh**[8,9,10], **Michael Nevitt**[5], and **T. Derek V. Cooke**[3]

Lisa Sheehy: lsheehy@bruyere.org; Elsie Culham: elsie.culham@queensu.ca; Linda McLean: mcleanl@queensu.ca; Jingbo Niu: niujp@bu.edu; John Lynch: JLynch@psg.ucsf.edu; Neil A. Segal: neil-segal@uiowa.edu; Jasvinder A. Singh: jasvinder.md@gmail.com; Michael Nevitt: MNevitt@psg.ucsf.edu; T. Derek V. Cooke: derek@cookes.ca

[1]Bruyère Research Institute, Ottawa, ON, Canada

[2]School of Rehabilitation Sciences, University of Ottawa, Ottawa, ON, Canada

[3]School of Rehabilitation Therapy, Queen's University, Kingston, ON, Canada

[4]Boston University School of Medicine, Boston, MA

[5]Department of Epidemiology and Biostatistics, University of California, San Francisco, CA

[6]Department of Rehabilitation Medicine, University of Kansas, Kansas City, KS

[7]Department of Epidemiology, University of Iowa, Iowa City, IA

Corresponding Author: Lisa Sheehy, Bruyère Research Institute, 43 Bruyère St., Rm. 703B, Ottawa, ON, Canada, K1N 5C8, 01-613-562-6262 Ext. 1593 (phone), 01-613-562-4256 (fax).

Author contributions: Dr. Sheehy conceived and designed the study, analyzed and interpreted the data, wrote and edited the manuscript and provided critical revision of the article for important intellectual content. She approves the final version.
Dr. McLean had input into the design, analysis and interpretation of the study, and provided critical revision of the article for important intellectual content. She approves the final version.
Dr. Cooke had input into the conception, design, analysis and interpretation of the study and provided critical revision of the article for important intellectual content. He approves the final version.
Dr. Niu is part of the Multicenter Osteoarthritis Study (MOST). She had input into the design of the study, performed the participant selection, assisted with the statistical analyses and was given the opportunity to provide critical revision of the article for important intellectual content. She approves the final version.
Dr. Lynch is part of the Multicenter Osteoarthritis Study (MOST). He had input into the design of the study, assisted with the interpretation of the data and was given the opportunity to provide critical revision of the article for important intellectual content. He approves the final version.
Dr. Segal is part of the Multicenter Osteoarthritis Study (MOST). He had input into the design of the study and provided critical revision of the article for important intellectual content. He approves the final version.
Dr. Singh is part of the Multicenter Osteoarthritis Study (MOST). He had input into the design of the study and provided critical revision of the article for important intellectual content. He approves the final version.
Dr. Nevitt is part of the Multicenter Osteoarthritis Study (MOST). He had input into the design of the study and was given the opportunity to provide critical revision of the article for important intellectual content. He approves the final version.
Dr. Culham had input into the design, analysis and interpretation of the study and provided critical revision of the article for important intellectual content. She approves the final version. Participants were provided by the MOST study database.

Conflict of interest: Co-author T.D.V. Cooke is the president of Orthopedic Alignment and Imaging Services, Inc. which provided the Surveyor™ image analysis program 3.1 free-of-charge. He has shares in OAISYS Inc. and has submitted a patent for a Joint Surgery Triage Tool which uses the compartmental grading scale for knee OA described in this paper.
Co-author N. Segal is an OA CME publication editor for Vindico and editor of a book on musculoskeletal care during pregnancy and current PM&R Journal for Springer SBM, LLC. He has an R42 application in for NIBIB.
Co-author J.A. Singh is a member of the executive of OMERACT, an organization that develops outcome measures in rheumatology and receives arms-length funding from 36 companies; a member of the American College of Rheumatology's Guidelines Subcommittee of the Quality of Care Committee; and a member of the Veterans Affairs Rheumatology Field Advisory Committee. He is a consultant for Savient, Takeda, Allergan and Regeneron and has grants from Takeda and Savient.

[8]Department of Medicine at School of Medicine and Division of Epidemiology at School of Public Health, University of Alabama at Birmingham, Birmingham, AL

[9]Medicine Service, VA Medical Center, Birmingham, AL

[10]Department of Orthopedic Surgery, Mayo Clinic College of Medicine, Rochester, MN

Abstract

Objectives—The purpose of this study was to assess the concurrent validity and sensitivity to change of three knee osteoarthritis (OA) grading scales.

The Kellgren-Lawrence (KL) and the Osteoarthritis Research Society International (OARSI) joint space narrowing (JSN) grading scales are well-established. The third scale, the compartmental grading scale for OA (CG) is a novel scale which grades JSN, femoral osteophytes, tibial erosion and subluxation to create a total score.

Methods—One sample of 72 posteroanterior fixed-flexion radiographs displaying mild to moderate knee OA was selected from the Multicenter Osteoarthritis Study to study validity. A second sample of 75 radiograph pairs, which showed an increase in OA severity over 30 months, was selected to study sensitivity to change.

The three radiographic grading scales were applied to each radiograph in both samples. Spearman's rank correlation coefficients were used to correlate the radiographic grades and the change in grades over 30 months with a Whole-organ Magnetic Resonance Imaging Score (WORMS)-based composite score which included five articular features of knee OA.

Results—Correlations between the KL, OARSI JSN and CG grading scales and the MRI-based score were 0.836, 0.840 and 0.773 ($p<0.0001$) respectively while correlations between change in the radiographic grading scales and change in the MRI-based score were 0.501, 0.525 and 0.492 ($p<0.0001$).

Conclusions—All three radiographic grading scales showed high validity and are suitable to assess knee OA severity. They showed moderate sensitivity to change; therefore caution should be taken when using ordinal radiographic grading scales to monitor knee OA over time.

Keywords

Knee osteoarthritis; knee radiographs; grading; validity; sensitivity to change

1.2 Introduction

Knee osteoarthritis (OA) is diagnosed with the presence of symptoms accompanied by radiographic changes [1]. To facilitate objective and consistent assessments, radiographs are generally scored using ordinal grading scales (scales with ordered or ranked categories). The most commonly-used grading scale is the Kellgren-Lawrence (KL) scale, which scores several features of OA in both the medial and lateral tibiofemoral (TF) compartments on an ordinal scale from zero to four [2]. Another commonly-used scale is the Osteoarthritis Research Society International (OARSI) joint space narrowing (JSN) scale [3,4]. This individual grading scale uses an atlas to compare radiographs to representative images and

assign a grade for the severity of JSN from zero to three in either the medial or lateral TF compartment [4]. Osteophytes are the primary feature for grades 0-2 for the KL scale and JSN is the only feature for the OARSI JSN scale. A scale that includes several features of OA at all grading levels might be better for monitoring progression in people with a variety of presentations of OA. To address this issue a composite knee OA grading scale, the compartmental grading scale for OA (CG), was designed to assess several features of knee OA individually but sum the scores for a total score out of 13 [5]. The CG scale is applied to the most severely-damaged TF compartment of the knee.

For grading scales to be recommended to assess knee OA on a radiograph, they must be valid (measure what they purport to measure) and sensitive to change. To assess concurrent validity, grades obtained from each radiograph scale must be compared to grades obtained from a criterion standard such as magnetic resonance images (MRI). MRIs allow the observation of cartilage damage and eliminate issues of magnification, distortion and superimposition [6]. KL grades show moderate associations with cartilage lesions and volume as seen on MRI [7,8]. Comparisons of OARSI JSN and CG grades to MRI findings have not been performed.

Sensitivity to change for radiographic grading scales is assessed using pairs of images taken from the same individual, at two time-points. Change in severity of knee OA observed using the radiographic grading scales is compared to change in severity observed using a criterion standard such as MRI. Sensitivity to change has not been assessed for any of the three radiographic grading scales.

Therefore the first goal of this study was to determine the validity of the KL, OARSI JSN and CG ordinal grading scales to measure the severity of TF OA on a radiograph and to establish if one of these scales was superior to the others. The second goal was to determine the sensitivity to change in the severity of TF OA over a 30-month period of the KL, OARSI JSN and CG grading scales and to ascertain if one of these scales was more sensitive than the others for detecting change over time.

1.3 Participants and Methods

1.3.1 Radiograph Selection

Knee radiographs for this cross-sectional ancillary study were obtained from the Multicenter Osteoarthritis Study (MOST) database. Potential participants were recruited from Iowa City, Iowa and Birmingham, Alabama, from April 2003 to April 2005; follow-up is ongoing [9]. The MOST study was approved by institutional review boards of the participating institutions; participants provided written informed consent. There are data on 3026 persons between the ages of 50 and 79 with, or at risk of developing knee OA, including individuals who are overweight or obese, those with knee pain and those with a history of knee injury or surgery [10,11]. Exclusion criteria include: a diagnosis of rheumatoid arthritis, ankylosing spondylitis, psoriatic arthritis, Reiter's syndrome, significant kidney disease or cancer; bilateral knee replacement; inability to walk without assistance; plans to move out of the study area within three years [10]. Further detail is available at http://most.ucsf.edu/default.asp.

Selected knees must have had KL and OARSI JSN grades assessed from bilateral fixed-flexion posteroanterior (PA) radiographs and whole-organ magnetic resonance imaging scores (WORMS) assessed from 1.0 Tesla MRIs, all performed at baseline and 30 months later. Consequently 1694 knees were available for selection [10]. The baseline hip-knee-ankle (HKA) angle, measured on anteroposterior full-length radiographs was also available for each participant. Samples were selected using an automated computer process (SAS®, version 9.2, SAS Institute Inc., Cary, NC). See supplementary material for participant flow diagram.

1.3.1.1 Concurrent Validity—One sample of 72 PA fixed-flexion knee radiographs (left or right), taken at baseline, was selected. Sample size was calculated based on a Pearson's correlation with two independent variables, a medium effect size, $\alpha = 0.05$ and statistical power $(1 - \beta) = 0.80$; it was estimated to be 67 [12].

To ensure that a wide range of knee OA severity was represented, potential participants were stratified according to a custom summed WORMS score [6]. This score was made up of the individual scores for the medial and lateral tibial (anterior, central, posterior) and femoral (central, posterior) sub-regions for the following features of knee OA: cartilage morphology (each subregion scored out of six), osteophytes (seven), bone attrition (three) and meniscal extrusion (each meniscus scored out of two), for a maximum total of 164. Potential participants were divided into four groups using the following divisions of the custom summed WORMS scores: 0-19 (976 knees), 20-39 (442 knees), 40-59 (159 knees) and 60-164 (117 knees).

To ensure that the same number of knees with, or at risk of medial and lateral TF compartment OA were included within each stratum, the most-affected TF compartment was determined. MOST defined this as the one with the greater OARSI JSN grade [4,10]. If OARSI JSN grades were equal, lower-limb alignment, measured using the HKA angle, was used, with the HKA > 1° varus for medial involvement (964 knees, 57%) and the HKA < 1° varus for lateral involvement (730, 43%) [13,15]. Eighteen participants were randomly selected from each group in this proportion of medial and lateral involvement.

1.3.1.2 Sensitivity to Change—A second sample, of 75 PA fixed-flexion radiograph pairs, taken at baseline and 30 months later, was selected. The sample size estimation was the same as for participant sample one.

Only radiograph pairs that showed change were selected. A minimal increase of at least 15% on the custom summed WORMS score from baseline to 30 months was required because a small increase in severity would not be expected to be detected on a radiograph. The CG grading scale was estimated to have a minimal detectable change of 2 out of 13, which is approximately a 15% change. An absolute minimum level of change was also required because in a knee with limited radiographic evidence of OA at baseline a 15% increase would be a small absolute number, which would not be detectable on a radiograph. To determine this minimum, we used the following procedure. For MRIs with a custom summed WORMS score of less than 40, there was a 75% chance of a KL grade of zero or one (no OA). However, for WORMS summed scores of 40 or greater, there was a 94%

chance of a KL grade of two or greater (OA present). We therefore calculated 15% of this score (40), which was six, as the minimal change that would be expected to be seen on a radiograph. These criteria were met by 173 knees. Of these, 75 individuals were randomly selected.

1.3.2 Measurements

1.3.2.1 Kellgren-Lawrence Grades—Standing PA fixed-flexion radiographs were assessed for KL grades by two blinded expert readers from MOST [10,11]. Baseline and follow-up films were scored while viewed simultaneously, with the chronological order of the images known to the readers [10].

KL grades were assigned to each knee: 1-doubtful narrowing of joint space and possible osteophytic lipping; 2-definite osteophytes and possible joint space narrowing; 3-moderate multiple osteophytes, definite narrowing of joint space and some sclerosis and possible deformity of bone ends; 4-large osteophytes, marked narrowing of joint space, severe sclerosis and definite deformity of bone ends [2,11,16]. The MOST study used a modified version of the KL grades (inter-rater reliability: Cohen's kappa 0.66) [11] For knees with a KL grade of four, a lateral radiograph was also viewed. A grade of 3.5 was assigned if bone-on-bone cartilage erosion was seen on the PA radiograph but residual joint space was seen on a lateral radiograph [11]. We wished to test the original scale, therefore KL grades of 3.5 were changed to 4 for the current study.

1.3.2.2 Osteoarthritis Research Society International Joint Space Narrowing Grades—OARSI JSN grades were assessed by expert readers from MOST, with the same procedures used for KL grades [10]. OARSI JSN grades (0-normal, 1-mild, 2-moderate, 3-severe) were given for the most severely-affected TF compartment following the radiograph examples in the OARSI Radiographic Atlas (intra-rater reliability: Cohen's weighted kappa 0.67; inter-rater reliability: Cohen's weighted kappa 0.48) [4,17]. Additionally, in the MOST protocol, if there was clear evidence of JSN worsening but not enough to assign the next grade, a half-grade was given for the second radiograph [11]. To assess the original OARSI JSN grading scale, all one-half grades were changed to the lower integer for the current study.

1.3.2.3 Compartmental Grading Scale for Osteoarthritis—One experienced reader assessed the CG grades following the same procedures used for the KL grades. CG features of the most severely-affected TF compartment were analyzed [JSN (0-3), presence and size of femoral osteophytes (0-3), presence and degree of tibial attrition (0-4) and evidence of subluxation (0-3)] for a total score from 0-13 (intra-rater reliability: Cohen's weighted kappa 0.65-0.75; test-retest reliability: Cohen's weighted kappa 0.64) [5,18].

1.3.2.4 Whole-organ Magnetic Resonance Imaging Scores—WORMS scoring was performed on pairs of MRIs by expert readers from MOST, using the procedures described for KL grades [10]. Results for the most severely-affected TF compartment were used. The five articular features of the WORMS [(cartilage morphology (scored out of 6), tibial and femoral osteophytes (7), bone attrition (3), bone marrow lesions (3) and subchondral cysts

(3)] were used to create a WORMS composite score different from the one used for participant selection. The WORMS composite score included the score for the worst of the tibial and femoral sub-regions for each articular feature, for a maximum score of 22.

Individual features of the WORMS were also correlated to corresponding features of the CG scale and to the OARSI JSN scale. JSN grades were correlated with the WORMS cartilage morphology score for the worst of the tibial and femoral sub-regions. CG femoral osteophyte grades were correlated with the WORMS osteophyte score for the worst of the femoral sub-regions. CG tibial attrition grades were correlated with the WORMS bone attrition score for the worst of the tibial sub-regions. CG subluxation grades were correlated with the WORMS meniscal extrusion score.

1.3.3 Procedure

The KL and OARSI JSN grades and WORMS scores had been recorded by MOST. To obtain CG grades, each PA fixed-flexion knee image was assessed using an imaging analysis program, Surveyor™ 3.1 (Orthopedic Alignment & Imaging Systems Inc., Kingston, ON) [5]. Once the images were graded, unblinded data were released by MOST for each participant, including demographic data, KL and OARSI JSN grades, and WORMS composite scores.

1.3.4 Data Analysis

1.3.4.1 Concurrent Validity—Spearman's rank correlation coefficients were used to correlate the radiographic grades with the WORMS composite score. Values ≥ 0.80 are considered to indicate a very high correlation between two features for Pearson's correlation coefficients [19]. Similarly, values between 0.60 - 0.79 indicate high validity, 0.30 - 0.59 indicate moderate validity and < 0.30 indicate low validity [19]. Confidence intervals were used to compare correlation coefficients.

The Spearman's rank correlation coefficient was also calculated for medial versus lateral TF compartment involvement and right versus left knee involvement. Individual components of the OARSI JSN and CG grading scales were also correlated with comparable components of the WORMS composite score.

1.3.4.2 Sensitivity to Change—Spearman's rank correlation coefficients were used to correlate the change in each grading scale from baseline to 30 months with the change in the WORMS composite score. Also, change in OARSI JSN grades and the individual features of the CG grading scale were correlated with change for the comparable features of the WORMS composite score.

Analyses were performed using Minitab (version 15.1.30.0, Minitab Inc., State College, PA) and MedCalc (version 12.2.1.0, MedCalc Software, Mariakerke, Belgium). Statistical significance was set at $\alpha = 0.05$.

1.4 Results

1.4.1 Participants

Participant samples are described in Table 1. Two participants in sample one and three in sample two had both right and left knees assessed. Summaries of the radiographic grades and WORMS composite scores are found in Table 2. A summary of knees that showed change over 30 months on the radiographic grading scales relative to the WORMS score is found in Table 3.

For sample two, 29/75 knees had the lateral TF compartment designated most-affected. Of these, only eight showed definite progression of OA in the lateral TF compartment. In the remaining 21 knees, there were either definite medial TF compartment changes, or very little change. Both the medial and lateral TF compartments were assessed with the CG scale for these knees and only the TF compartment which changed the most on the WORMS scoring was included in the analyses. If both compartments changed the same amount, the lateral TF compartment was used. This changed the proportion of knees identified with medial and lateral TF compartments most-affected from 46:29 to 55:20. Overall, 66 knees showed change with WORMS scoring in the most-affected TF compartment.

In the second sample, one knee had a KL grade of 3.5 assigned at 30 months, which was changed to a grade of 4 for analysis. A half-grade was assigned for 9 knees at 30 months for the OARSI JSN scale; these were changed to the lower integer.

1.4.2 Concurrent Validity

To ensure independence, one knee was randomly excluded for the two individuals who had both knees selected (one right and one left). Spearman's rank correlations between the radiographic measures of knee OA severity and WORMS composite scores were high to very high (Figure 1 and Table 4). The confidence intervals overlapped considerably, showing that no scale was preferred. Correlations of OARSI JSN and CG JSN grades to WORMS cartilage morphology scores were also very high, however correlations for the other CG features were less robust.

There were no differences in Spearman's rank correlations between right and left knees or between medial and lateral TF compartments, with one exception. The Spearman's rank correlation coefficient for the association of the CG femoral osteophyte grade with the WORMS femoral osteophyte score was 0.61 ($p < 0.0001$) for the medial TF compartment and 0.36 ($p = 0.0467$) for the lateral TF compartment.

1.4.3 Sensitivity to Change

One knee was randomly excluded for the three individuals who had both knees selected (one right and two left). Spearman's rank correlation coefficients for change in knee OA severity over 30 months seen on the knee OA radiographic grading scales relative to the WORMS composite score show moderate sensitivity to change (Table 5). The confidence intervals overlapped considerably suggesting that no scale was more sensitive to change than the others. Change for the individual radiographic OA features was moderately associated with

the corresponding change in WORMS features, although for the association of CG subluxation and WORMS meniscal extrusion, the association was surprisingly negative ($r = -0.409$, $p = 0.0004$), which suggests that an increase in subluxation is moderately associated with a decrease in meniscal extrusion.

1.5 Discussion

The KL, OARSI JSN and CG ordinal grading scales were all highly or very highly associated with WORMS composite scores of articular damage due to knee OA. Furthermore, they are considered equally-well correlated to the WORMS composite scale. KL grades have previously been correlated to cartilage defects [Spearman's $r = 0.55$, $p < 0.01$; Pearson's r of up to 0.52 (medial femoral condyle), depending on location, $p < 0.05$] [8,20] and osteophytes (Pearson's $r = 0.66$ in the medial TF compartment, $p < 0.05$) [20] as seen on MRI. The greater association observed in our study may be due to the inclusion of several selected knee OA features in our WORMS composite scale.

JSN assessed on a radiograph showed a strong association with the related WORMS feature of cartilage morphology. This was expected, since articular cartilage makes up a considerable proportion of the joint space. Since the meniscus also contributes to the observed joint space, meniscal subluxation or degeneration may contribute to the variance between the observation of JSN on fixed-flexion radiographs and cartilage morphology as seen on MRI [21-24].

The CG total score performed similarly to the other radiographic scales. While the other grading scales demonstrated a “ceiling effect” when the severity of knee OA measured by the WORMS custom composite scale was between 12 and 18, the CG grading scale did not, which suggests that it might continue to be sensitive in individuals with more severe presentations of knee OA. Testing of the CG grading scale on a cohort with more severe knee OA would be confirmatory. The CG individual feature scores were correlated to corresponding WORMS OA features in order to explore the content validity of the CG grading scale. Surprisingly, the CG JSN grades were more highly correlated to WORMS cartilage morphology scores than the CG total score was to the WORMS composite score; however, the inclusion of all four OA features in the CG score provides a more complete picture of OA change in the TF compartments. CG femoral osteophyte grades were moderately correlated to the WORMS femoral osteophyte scores and CG tibial attrition grades were highly correlated to the WORMS tibial bone attrition scores. Because radiographs offer a two-dimensional representation of the bony structure, osteophytes and bone attrition can often only be appreciated on the edges of the bones. Osteophytes which overlap may also not be appreciated. These differences from a three-dimensional MRI representation of the same bone may contribute to variance between the CG femoral osteophytes and tibial attrition grades and the corresponding WORMS scores. The CG subluxation grades did not correlate significantly with the WORMS meniscal extrusion scores. It is likely that meniscal extrusion contributes more to JSN than to subluxation [25,26]. Ligament laxity and bone attrition may contribute more to subluxation and should be studied in the future.

Correlations between radiographic and MRI OA features for left and right knees were the same, as expected. However, it was expected that radiographic OA features of the medial TF compartment, particularly JSN, would be more-highly associated to WORMS scores than those of the lateral TF compartment, since the fixed-flexion radiograph protocol emphasizes the positioning of the medial tibial plateau parallel to the x-ray beam. While we did not find this difference for JSN, there was a large difference between the medial and lateral TF compartments for the correlation of the CG femoral osteophyte grade to the WORMS femoral osteophyte score. Anecdotally, the readers reported that osteophytes on the lateral femoral condyle were more difficult to see than those on the medial condyle.

Changes seen on all three radiographic ordinal grading scales were moderately correlated with changes seen with the WORMS composite scale for progression of TF compartment OA severity. We did not find previous studies that reported the correlation between ordinal measures of change in radiographic knee OA severity and MRI measures. However, several authors have tested the association between change in the continuous variable of joint space width (JSW) measured from a radiograph in millimeters and change in cartilage volume measured from MRI, and have determined that there was no correlation (Spearman's rank correlation $r = -0.11$ and 0.19, $p > 0.05$) [27-29]. In a similar study, JSW was moderately associated with WORMS cartilage morphology of the whole knee (Spearman's rank correlation $r = 0.41$, $p = 0.039$) [30]. Although continuous scales measuring JSW are often used for clinical trials of potentially disease-modifying OA drugs [31], we show that ordinal scales for JSN appear to have a higher association to MRI findings of articular cartilage degeneration than continuous scales. This finding is similar to that of Nevitt et al. [32] and suggests that JSN could be an alternative for JSW as an outcome measure for change in TF compartment OA severity. To confirm this, JSW and JSN should be directly compared for sensitivity to change in a future study.

Attempts were made by MOST to increase the sensitivity to change of the KL and OARSI JSN grading scales [10,11,33]. In the participant sample for sensitivity to change, KL grade 3.5 was assigned only once by MOST. If the MOST definition of the KL grading scale was used, the correlation with the WORMS composite score was 0.473 ($p < 0.0001$); therefore the modified grading scale did not provide any increased sensitivity to change over the original KL scale. Increased sensitivity to change was greater for the modified OARSI JSN grades. In our participant sample, the OARSI JSN grade increased by a half-grade for nine knees as recorded by MOST. If the MOST definition of the OARSI JSN grading scale had been used, the correlation with the WORMS composite score would have been 0.563 ($p < 0.0001$), showing a beneficial effect of the revisions with respect to sensitivity to change.

One limitation of this study might be the unusually high number of participants in sample one with designated lateral TF OA, resulting from the compartment selection criteria used by MOST. This might have caused the correlations between all three radiographic grading scales and the WORMS composite score to be attenuated because the radiographic protocol favors the assessment of OA features in the medial TF compartment [34].

A second limitation was the participant selection criteria for sample two. Individuals were chosen based on a minimum amount of change on a WORMS-derived scale scored out of

164. The intent was that this score would give a global sense of the severity of the articular features in both TF compartments and would allow selection of a range of presentations of knee OA. We then correlated change in the three radiographic grading scales against a smaller WORMS composite score, which scored the articular features of OA only in the most-affected TF compartment. Unfortunately when the custom summed WORMS score had picked up “change”, this change was not always in the designated most-affected TF compartment. This occurred most often when there was no noticeable JSN at baseline and neutral alignment and required that the designated most-affected compartment be changed. A participant selection score that focused on choosing a single TF compartment would have prevented this confusion. Even so, because we ended up changing the most-affected TF compartment according to the WORMS composite score when required, there should be no bias against the application of the radiographic grading scales.

A related, and potential limitation is that the selection criteria for sample two may be seen as defining a certain amount of sensitivity to change, which is the outcome being measured. It was necessary to choose only knees with change because many knees in the MOST database had not changed over 30 months and a random selection would have given inconclusive results. The correlation between change in radiographic scales and change in WORMS scores was around 0.50, which shows that the sample did contain many knees whose change was not picked up by the radiographic scoring methods (Table 3).

These results have established the validity of the OARSI JSN and CG grading scales, and the sensitivity to change of all three scales, outcomes that have not previously been reported. We conclude that since the KL, OARSI JSN and CG grading scales are all highly correlated to OA joint changes seen on MRI, these grading scales are equally valid and suitable to assess knee OA severity. Furthermore, all three radiographic scoring methods are moderately and equally sensitive to change for knee OA severity over 30 months. The moderate results suggest that caution must be taken when using ordinal radiographic grading scales to monitor change in knee OA severity over time.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Chris Wale, who customized the Surveyor™ image analysis program for the study.

Role of the funding source: The MOST study was funded by the National Institutes of Health and the National Institute of Aging under the following grant numbers: Felson (Boston University) – AG18820, Lewis (University of Alabama at Birmingham – AG18947, Torner (University of Iowa) – AG18832 and Nevitt (University of California at San Francisco) – AG19069. The Ancillary Study presented in this paper was unfunded.

References

1. Mathers, CD.; Bernard, C.; Iburg, KM.; Inoue, M.; Fat, DM.; Shibuya, K., et al. Annex Table 4: GBD2000 cause categories, sequelae and case definitions. World Health Organization; 2003. Global Burden of Disease in 2002: data sources, methods and results. Global Programme on Evidence for Health Policy Discussion Paper NO. 54. Report.: 54

2. Kellgren J, Lawrence J. Radiological assessment of osteo-arthrosis. Annals of the Rheumatic Diseases. 1957; 16:494–502. [PubMed: 13498604]
3. Altman RD, Hochberg M, Murphy WA Jr, Wolfe F, Lequesne M. Atlas of individual radiographic features in osteoarthritis. Osteoarthritis and Cartilage. 1995; 3(Suppl A):3–70. [PubMed: 8581752]
4. Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. Osteoarthritis and Cartilage. 2007; 15(Suppl A):A1–56. [PubMed: 17320422]
5. Cooke TDV, Kelly BP, Harrison L, Mohamed G, Khan B. Radiographic grading for knee osteoarthritis. A revised scheme that relates to alignment and deformity Journal of Rheumatol. 1999; 26:641–644.
6. Peterfy CG, Guermazi A, Zaim S, Tirman PF, Miaux Y, White D, et al. Whole-Organ Magnetic Resonance Imaging Score (WORMS) of the knee in osteoarthritis. Osteoarthritis and Cartilage. 2004; 12:177–190. [PubMed: 14972335]
7. Harada Y, Tokuda O, Fukuda K, Shiraishi G, Motomura T, Kimura M, et al. Relationship between cartilage volume using MRI and Kellgren-Lawrence radiographic score in knee osteoarthritis with and without meniscal tears. American Journal of Roentgenology. 2011; 196:W298–W304. [PubMed: 21343478]
8. Link TM, Steinbach LS, Ghosh S, Ries M, Lu Y, Lane N, et al. Osteoarthritis: MR imaging findings in different stages of disease and correlation with clinical findings. Radiology. 2003; 226:373–381. [PubMed: 12563128]
9. Segal NA, Nevitt MC, Gross KD, Hietpas J, Glass NA, Lewis CE, et al. The Multicenter Osteoarthritis Study: opportunities for rehabilitation research. PM R. 2013; 5:647–654. [PubMed: 23953013]
10. Multicenter Osteoarthritis Study Public Data Sharing. http://most.ucsf.edu/default.asp [2011 Available from: URL:http://most.ucsf.edu/default.asp
11. Felson DT, Nevitt MC, Yang M, Clancy M, Niu J, Torner JC, et al. A new approach yields high rates of radiographic progression in knee osteoarthritis. Journal of Rheumatology. 2008; 35:2047–2054. [PubMed: 18793000]
12. Cohen J. A power primer. Psychology Bulletin. 1992; 112:155–159.
13. Cooke D, Scudamore A, Li J, Wyss U, Bryant T, Costigan P. Axial lower-limb alignment: comparison of knee geometry in normal volunteers and osteoarthritis patients. Osteoarthritis and Cartilage. 1997; 5:39–47. [PubMed: 9010877]
14. Felson DT, Niu J, Gross KD, Englund M, Sharma L, Cooke TD, et al. Valgus malalignment is a risk factor for lateral knee osteoarthritis incidence and progression: Findings from the multicenter osteoarthritis study and the osteoarthritis initiative. Arthritis and Rheumatism. 2013; 65:355–362. [PubMed: 23203672]
15. Niu J, Zhang YQ, Torner J, Nevitt M, Lewis CE, Aliabadi P, et al. Is obesity a risk factor for progressive radiographic knee osteoarthritis? Arthritis and Rheumatism. 2009; 61:329–335. [PubMed: 19248122]
16. Kellgren, J.; Jeffrey, MR.; Ball, J. The epidemiology of chronic rheumatism: atlas of standard radiographs of arthritis. Oxford: Blackwell Scientific; 1963.
17. Gossec L, Jordan JM, Mazzuca SA, Lam MA, Suarez-Almazor ME, Renner JB, 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. Osteoarthritis and Cartilage. 2008; 16:742–748. [PubMed: 18417373]
18. Sheehy L, Cooke TDV, McLean L, Niu J, Lynch J, Segal NA, et al. Reliability of the uni-compartmental scale for the radiographic evaluation of knee osteoarthritis: Data from the Multicenter Osteoarthritis Study. Osteoarthritis and Cartilage. 2013; 21(Suppl):S201.
19. Handbook of Research Methods for Nursing and Health Science. 2nd. Frenchs Forest, NSW: Pearson Education Australia; 2004.
20. Sowers M, Karvonen-Gutierrez CA, Jacobson JA, Jiang Y, Yosef M. Associations of anatomical measures from MRI with radiographically defined knee osteoarthritis score, pain, and physical functioning. Journal of Bone and Joint Surgery American. 2011; 93:241–251.
21. Buckland-Wright C. Which radiographic techniques should we use for research and clinical practice? Best Practices and Research in Clinical Rheumatology. 2006; 20:39–55.

22. Gale DR, Chaisson CE, Totterman SM, Schwartz RK, Gale ME, Felson D. Meniscal subluxation: association with osteoarthritis and joint space narrowing. Osteoarthritis and Cartilage. 1999; 7:526–532. [PubMed: 10558850]

23. Hunter DJ, Zhang YQ, Tu X, Lavalley M, Niu JB, Amin S, et al. Change in joint space width: hyaline articular cartilage loss or alteration in meniscus? Arthritis and Rheumatism. 2006; 54:2488–2495. [PubMed: 16868968]

24. Hunter DJ, Buck R, Vignon E, Eckstein F, Brandt K, Mazzuca SA, et al. Relation of regional articular cartilage morphometry and meniscal position by MRI to joint space width in knee radiographs. Osteoarthritis and Cartilage. 2009; 17:1170–1176. [PubMed: 19401243]

25. Adams JG, McAlindon T, DiMasi M, Carey J, Eustace S. Contribution of meniscal extrusion and cartilage loss to joint space narrowing in osteoarthritis. Clinics in Radiology. 1999; 54:502–506.

26. Berthiaume MJ, Raynauld JP, Martel-Pelletier J, Labonte F, Beaudoin G, Bloch DA, et al. Meniscal tear and extrusion are strongly associated with progression of symptomatic knee osteoarthritis as assessed by quantitative magnetic resonance imaging. Annals of the Rheumatic Diseases. 2005; 64:556–563. [PubMed: 15374855]

27. Raynauld JP, Martel-Pelletier J, Berthiaume MJ, Beaudoin G, Choquette D, Haraoui B, et al. Long term evaluation of disease progression through the quantitative magnetic resonance imaging of symptomatic knee osteoarthritis patients: correlation with clinical symptoms and radiographic changes. Arthritis Research and Therapy. 2006; 8:R21. [PubMed: 16507119]

28. Raynauld JP, Martel-Pelletier J, Berthiaume MJ, Labonte F, Beaudoin G, de Guise JA, et al. Quantitative magnetic resonance imaging evaluation of knee osteoarthritis progression over two years and correlation with clinical symptoms and radiologic changes. Arthritis and Rheumatism. 2004; 50:476–487. [PubMed: 14872490]

29. Cicuttini F, Hankin J, Jones G, Wluka A. Comparison of conventional standing knee radiographs and magnetic resonance imaging in assessing progression of tibiofemoral joint osteoarthritis. Osteoarthritis and Cartilage. 2005; 13:722–727. [PubMed: 15922634]

30. Stahl R, Jain SK, Lutz J, Wyman BT, Le Graverand-Gastineau MP, Vignon E, et al. Osteoarthritis of the knee at 3.0 T: comparison of a quantitative and a semi-quantitative score for the assessment of the extent of cartilage lesion and bone marrow edema pattern in a 24-month longitudinal study. Skeletal Radiology. 2011; 40:1315–1327. [PubMed: 21479518]

31. Conrozier T, Mathieu P, Piperno M, Favret H, Colson F, Vignon M, et al. Selection of knee radiographs for trials of structure-modifying drugs in patients with knee osteoarthritis: a prospective, longitudinal study of Lyon Schuss knee radiographs with the definition of adequate alignment of the medial tibial plateau. Arthritis and Rheumatism. 2005; 52:1411–1417. [PubMed: 15880526]

32. Nevitt MC, Peterfy C, Guermazi A, Felson DT, Duryea J, Woodworth T, et al. Longitudinal performance evaluation and validation of fixed-flexion radiography of the knee for detection of joint space loss. Arthritis and Rheumatism. 2007; 56:1512–1520. [PubMed: 17469126]

33. Felson DT, Niu J, Guermazi A, Sack B, Aliabadi P. Defining radiographic incidence and progression of knee osteoarthritis: suggested modifications of the Kellgren and Lawrence scale. Annals of the Rheumatic Diseases. 2011; 70:1884–1886. [PubMed: 21908453]

34. Peterfy C, Li J, Zaim S, Duryea J, Lynch J, Miaux Y, et al. Comparison of fixed-flexion positioning with fluoroscopic semi-flexed positioning for quantifying radiographic joint-space width in the knee: test-retest reproducibility. Skeletal Radiology. 2003; 32:128–132. [PubMed: 12605275]

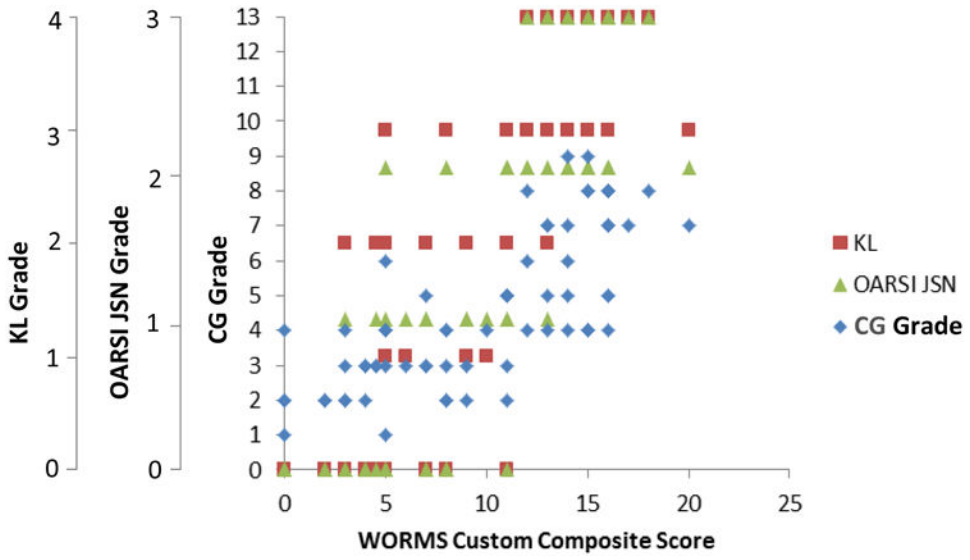

Figure 1.

Radiographic grade plotted against the WORMS composite score for 72 knees with a range of osteoarthritis severity.

KL – Kellgren-Lawrence grading scale

OARSI JSN – Osteoarthritis Research Society International joint space narrowing grading scale

CG – Compartmental osteoarthritis grading scale

WORMS – Whole organ magnetic resonance imaging score

Table 1
Description of participant samples [mean (standard deviation)]

	Concurrent Validity Sample	Sensitivity to Change Sample
Number of knees	72	75
Right : Left	40 : 32	46 : 29
Males : Females	38 : 34	22 : 50
Age (years)	63.2 (8.0)	62.3 (8.2)
Most-affected tibiofemoral compartment Medial : Lateral	40 : 32	55 : 20[1]
Body Mass Index (kg/m^2)	29.7 (4.7)	30.2 (4.8)
WOMAC Physical Function Subscale Score[2] (maximum score 68)	15.6 (12.1)	14.6 (11.8)
WOMAC Knee Pain Subscale Score[2] (affected knee, right or left, maximum score 20)	3.3 (2.9)	3.3 (3.1)
WOMAC Total Score[2] (affected knee, right or left, maximum score 96)	20.9 (15.3)	19.6 (14.0)
20 metre walk (average time of 2 trials, seconds)	16.6 (2.5)	16.5 (2.5)
5 chair stands (average time of 2 trials, seconds)	11.4 (4.7)	11.5 (4.3)

[1] ratio after analysis of most-affected compartment at 30 months completed

[2] WOMAC – Western Ontario and McMaster Universities Arthritis Index

Table 2

KL, OARSI JSN and CG grades and WORMS composite scores for concurrent validity and sensitivity to change.

	Concurrent Validity				Sensitivity to Change							
	KL	OARSI JSN	CG	WORMS composite	KL		OARSI JSN		CG		WORMS composite	
					Baseline	Follow-up	Baseline	Follow-up	Baseline	Follow-up	Baseline	Follow-up
Mean (standard deviation)	2.0 (1.6)	1.4 (1.2)	4.6 (2.2)	9.9 (5.3)	1.4 (1.2)	2.3 (1.3)	0.9 (0.9)	1.5 (1.0)	1.8 (1.4)	3.1 (2.0)	7.3 (3.4)	10.8 (3.8)
Range	0-4	0-3	1-9	0-20	0-4	0-4	0-3	0-3	0-6	0-8	1-14	1-19
Median	2.5	1.5	4.0	11.0	1.0	3.0	1.0	2.0	2.0	3.0	7.0	11.0
Interquartile range	3.0	2.0	4.0	10.0	2.0	1.0	2.0	1.0	2.0	3.0	5.0	6.0

KL – Kellgren-Lawrence grading scale
OARSI JSN – Osteoarthritis Research Society International joint space narrowing grading scale
CG – Compartmental grading scale for OA
WORMS – Whole organ magnetic resonance imaging score

Table 3

Change in radiographic grade compared to whole-organ magnetic resonance imaging scores (WORMS) for the sensitivity to change sample (sample 2). Sixty six knees showed change in the most-affected compartment according to the WORMS scores.

	Progression over 30 months	No change over 30 months	Improved over 30 months
Kellgren-Lawrence	40	26	0
Osteoarthritis Research Society International Joint Space Narrowing	35	31	0
Compartmental Grading Scale	48	17	1

Table 4

Spearman's rank correlation coefficients (r) for concurrent validity of several methods of radiographic knee osteoarthritis assessment. (n = 70)

Correlates		Spearman's r p-value confidence interval
KL	WORMS composite score	0.836 <0.0001 0.748 to 0.895
OARSI JSN	WORMS composite score	0.840 <0.0001 0.754 to 0.898
CG	WORMS composite score	0.773 <0.0001 0.658 to 0.853
OARSI JSN	WORMS cartilage morphology	0.829 <0.0001 0.738 to 0.891
CG JSN	WORMS cartilage morphology	0.837 <0.0001 0.749 to 0.896
CG femoral osteophytes	WORMS femoral osteophytes	0.488 <0.0001 0.285 to 0.648
CG tibial attrition	WORMS tibial bone attrition	0.629 <0.0001 0.462 to 0.753
CG subluxation	WORMS meniscal extrusion	0.207 0.0851 -0.029 to 0.422

KL – Kellgren-Lawrence grading scale
OARSI JSN – Osteoarthritis Research Society International joint space narrowing grading scale
CG – Compartmental grading scale for OA
WORMS – Whole organ magnetic resonance imaging score

Table 5

Spearman's rank correlation coefficients (r) for sensitivity to change over 30 months of several methods of radiographic knee osteoarthritis assessment. (n = 73)

Correlates		**Spearman's r p-value confidence interval**
KL	WORMS composite score	0.501 <0.0001 0.304 to 0.656
OARSI JSN	WORMS composite score	0.525 <0.0001 0.334 to 0.675
CG	WORMS composite score	0.492 <0.0001 0.293 to 0.649
OARSI JSN	WORMS cartilage morphology	0.423 <0.0001 0.212 to 0.596
CG JSN	WORMS cartilage morphology	0.389 0.0007 0.173 to 0.570
CG femoral osteophytes	WORMS femoral osteophytes	0.300 0.0104 0.074 to 0.497
CG tibial attrition	WORMS tibial bone attrition	0.316 0.0069 0.091 to 0.510
CG subluxation	WORMS meniscal extrusion	-0.409 0.0004 -0.585 to -0.196

KL – Kellgren-Lawrence grading scale
OARSI JSN – Osteoarthritis Research Society International joint space narrowing grading scale
CG – Compartmental grading scale for OA
WORMS – Whole organ magnetic resonance imaging score