

# Cartilage Lesion Score: Comparison of a Quantitative Assessment Score with Established Semiquantitative MR Scoring Systems<sup>1</sup>

Hamza Alizai, MD  
Warapat Virayavanich, MD<sup>2</sup>  
Gabby B. Joseph, PhD  
Lorenzo Nardo, MD<sup>3</sup>  
Felix Liu, MS  
Hans Liebl, MD  
Michael C. Nevitt, PhD  
John A. Lynch, PhD  
Charles E. McCulloch, PhD  
Thomas M. Link, MD, PhD

<sup>1</sup>From the Department of Radiology and Biomedical Imaging (H.A., W.V., G.B.J., L.N., H.L., T.M.L.) and Department of Epidemiology and Biostatistics (F.L., M.C.N., J.A.L., C.E.M.), University of California—San Francisco, 185 Berry St, Suite 350, San Francisco, CA 94107; and Department of Radiology, University of Texas Health Science Center at San Antonio, San Antonio, Tex (H.A.). Received September 25, 2012; revision requested November 9; revision received July 18, 2013; accepted August 6; final version accepted October 15. Address correspondence to H.A. (e-mail: [alizai@uthscsa.edu](mailto:alizai@uthscsa.edu)).

The Osteoarthritis Initiative is a public-private partnership comprised of five contracts (N01-AR-2-2258, N01-AR-2-2259, N01-AR-2-2260, N01-AR-2-2261, N01-AR-2-2262) funded by the National Institutes of Health, a branch of the Department of Health and Human Services, and conducted by the OAI Study Investigators. Private funding partners include Merck Research Laboratories, Novartis Corporation, GlaxoSmithKline, and Pfizer. Private sector funding for the OAI is managed by the Foundation for the National Institutes of Health. This manuscript was prepared using an OAI public use data set and does not necessarily reflect the opinions or views of the OAI investigators, the NIH, or the private funding partners.

<sup>2</sup>Current address: Department of Diagnostic and Therapeutic Radiology, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.

<sup>3</sup>Current address: Department of Radiology, University of Brescia, Brescia, Italy.

© RSNA, 2014

## Purpose:

To describe a scoring system for quantification of cartilage lesions (Cartilage Lesion Score [CaLS]), to determine its reproducibility, to examine the association of CaLS-detected longitudinal change with known risk factors for osteoarthritis (OA) progression by comparing a group of subjects with OA risk factors with a group of subjects without OA risk factors, and to compare the CaLS system with the established semiquantitative Whole-Organ Magnetic Resonance Imaging Score (WORMS) and Boston-Leeds Osteoarthritis Knee Score (BLOKS) systems in terms of detection of cartilage defect progression.

## Materials and Methods:

All subjects provided written informed consent, and the local institutional review board approved this HIPAA-compliant study. Fifty-two subjects with and 25 subjects without risk factors for knee OA were randomly selected from the Osteoarthritis Initiative. Inclusion criteria were age of 45–60 years, body mass index of 19–27 kg/m<sup>2</sup>, and no knee pain or OA on radiographs at baseline. Baseline and 24-month follow-up right knee 3-T magnetic resonance images were analyzed with WORMS, BLOKS, and CaLS systems. Progression of cartilage lesions with each scoring system was compared by using multilevel mixed-effects linear-regression models.  $\kappa$  values were calculated to determine reliability.

## Results:

Intraclass coefficient values for inter- and intraobserver reliability of the CaLS system were 0.86 and 0.91, respectively. Interobserver  $\kappa$  value range for individual features was 0.81–0.94. The CaLS system enabled significantly higher detection of cartilage lesion progression than did WORMS or BLOKS systems ( $P < .001$ ); 51.8% (56 of 108), 17.6% (19 of 108), and 13.0% (14 of 108) of the lesions progressed when analyzed with the CaLS, WORMS, and BLOKS systems, respectively. With the CaLS system, subjects with OA risk factors had significantly higher odds of progression than did subjects without risk factors (odds ratio, 2.78;  $P = .005$ ).

## Conclusion:

The CaLS system is a reproducible scoring system for cartilage lesions that yields an improved detection rate for monitoring progression when compared with detection rates of semiquantitative WORMS and BLOKS systems.

© RSNA, 2014

Osteoarthritis (OA) is the most prevalent chronic joint disease in the United States, and its incidence is increasing due to the aging population and the obesity pandemic (1). The failure to develop effective OA-specific therapies has created a greater need to develop imaging tools that are more sensitive and specific for OA; these will be invaluable because they will enhance our knowledge of OA pathophysiology, enable us to diagnose OA at an early stage when secondary prevention is still possible, and allow us to effectively monitor disease progression.

The current diagnostic criteria for knee OA rely on clinical and radiologic features (2). Plain radiography, considered the imaging reference standard, uses joint space as a surrogate for hyaline cartilage, enabling crude assessment of the disease process (3). However, radiography is not very sensitive to disease progression (4). Furthermore, there is a growing awareness that the OA disease process involves a

complex interplay of all tissues in the affected joint. A Food and Drug Administration initiative (5) to define the disease state of OA recommended that additional objective imaging criteria beyond plain radiography are needed to assess early onset of structural abnormalities, including changes to the bone, cartilage, menisci, synovium, and other soft tissues of the joint, in patients with OA (6–10).

Several magnetic resonance (MR) imaging-based semiquantitative scoring systems exist for the purpose of whole organ assessment of the OA joint. These include the Whole-Organ Magnetic Resonance Imaging Score (WORMS) (11), Boston-Leeds Osteoarthritis Knee Score (BLOKS) (12), and MR Osteoarthritis Knee Score (MOAKS) (13). However, relatively recent studies have highlighted problems with scaling of items, especially in early OA cohorts in which only the low end of the scales can be used, and consequent concerns about responsiveness have been raised (14,15). To improve detection capabilities of the scales, some groups have used methods to detect within-grade changes when using the WORMS system (16). Each of these systems has its strengths and weaknesses in terms of assessment of different features of the joint with OA (eg, meniscal abnormalities, cartilage lesions) (17,18).

Cartilage degradation is the hallmark of OA, and it is considered an irreversible progressive process (19); this is unlike other joint features associated with OA whose natural course can be variable (20). This makes cartilage lesions ideal for monitoring subtle progression of disease. Thus, the purpose

of our study was to (a) describe a scoring system for quantification of cartilage lesions (Cartilage Lesion Score [CaLS]), (b) determine its reproducibility, (c) determine the association of CaLS-detected longitudinal change in cartilage lesions with known risk factors for OA progression by comparing progression in a healthy cohort with progression in a cohort consisting of subjects with OA risk factors, and (d) compare the CaLS system with established semiquantitative WORMS and BLOKS systems in terms of detection of cartilage defect progression.

### Advances in Knowledge

- A quantitative scoring system for cartilage lesions has been proposed, and it has been shown to be reproducible (interobserver  $\kappa$  values ranged from 0.81 to 0.94 for individual features of the score).
- Cartilage Lesion Score (CaLS) enables better detection of change in cartilage lesions than does Whole-Organ Magnetic Resonance Imaging Score or Boston-Leeds Osteoarthritis Knee Score, and this detected change is associated with known risk factors for osteoarthritis (OA) progression; by using CaLS, the subjects with risk factors for OA had significantly higher odds ratios of progression than did subjects without OA risk factors (odds ratio, 2.78; 95% confidence interval: 1.36, 5.7;  $P = .005$ ).
- CaLS showed more change in subjects with OA risk factors than in those without risk factors.

### Implication for Patient Care

- The quantitative measurement of cartilage lesions and the better detection of change in lesions over time that are enabled by the new CaLS system may help clinicians in the follow-up and management-related decision-making process for eventual treatment of cartilage degeneration.

### Materials and Methods

#### Subjects

The Osteoarthritis Initiative (OAI) is a large-scale multicenter longitudinal cohort study of OA that provides annual MR images of subjects (21). The data

#### Published online before print

10.1148/radiol.13122056 Content codes: MK BQ

Radiology 2014; 271:479–487

#### Abbreviations:

BLOKS = Boston-Leeds Osteoarthritis Knee Score  
 BMI = body mass index  
 CaLS = Cartilage Lesion Score  
 MOAKS = MR Osteoarthritis Knee Score  
 OA = osteoarthritis  
 OAI = Osteoarthritis Initiative  
 WORMS = Whole-Organ Magnetic Resonance Imaging Score

#### Author contributions:

Guarantors of integrity of entire study, H.A., L.N., T.M.L.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; literature research, H.A., G.B.J., L.N., H.L., J.A.L., T.M.L.; clinical studies, H.A., W.V., G.B.J., L.N., T.M.L.; statistical analysis, W.V., G.B.J., L.N., F.L., J.A.L., C.E.M.; and manuscript editing, H.A., W.V., G.B.J., L.N., F.L., H.L., J.A.L., C.E.M., T.M.L.

#### Funding:

This research was supported by the National Institutes of Health (grants U01 AR059507, P50 AR060752, N01-AR-2-2258, N01-AR-2-2259, N01-AR-2-2260, N01-AR-2-2261, N01-AR-2-2262) and the Osteoarthritis Initiative.

Conflicts of interest are listed at the end of this article.

are available for public access (<http://www.oai.ucsf.edu/>). All subjects provided written informed consent, and this Health Insurance Portability and Accountability Act-compliant study was approved by the University of California-San Francisco institutional review board.

A subset of 52 subjects (31 women; age range, 45–60 years; mean age, 51.3 years; 21 men; age range, 47–58 years; mean age, 51.6 years) with two or more risk factors for OA but no symptoms or radiographic evidence of OA (Kellgren-Lawrence score of 1 or less) in the study knee was randomly selected from the OAI incidence cohort. The specific OA risk factors included: (a) knee symptoms (pain, aching, stiffness, use of pain medication), (b) overweight or obesity, (c) prior knee injury, (d) prior knee surgery, (e) family history of knee replacement, (f) Heberden nodes, and (g) frequent knee bending activity. Twenty-five additional subjects (20 women; age range, 47–54 years; mean age, 50.7 years; five men; age range, 47–55 years; mean age, 51.4 years) with no risk factors for OA and no symptomatic or radiographic OA also were randomly selected from the healthy cohort of the OAI. The exclusion criteria for the OAI included rheumatoid arthritis, severe joint space narrowing in both knees, and a positive pregnancy test.

The specific inclusion criteria for the presented analyses were as follows: (a) age range of 45–60 years, (b) body mass index (BMI) range of 19–27 kg/m<sup>2</sup>, (c) no pain in either knee (Western Ontario and McMaster University score of zero), and (d) Kellgren-Lawrence score of 1 or less on right (study) knee radiographs at baseline. These specific inclusion criteria were used to identify patients with early degenerative changes in whom a scoring system that can be used to detect subtle changes would be most advantageous.

#### MR Imaging Protocol

MR images were obtained by using identical 3.0-T (Magnetom Trio; Siemens, Erlangen, Germany) units and quadrature transmit-receive coils (USA Instruments, Aurora, Ohio) at four sites (Ohio State University, Columbus;

University of Maryland School of Medicine, Baltimore; University of Pittsburgh, Pa; and Memorial Hospital of Rhode Island, Pawtucket). The following sequences were performed: sagittal two-dimensional intermediate-weighted fast spin-echo sequence (repetition time msec/echo time msec, 3200/30; spatial resolution, 0.357 × 0.511; section thickness, 3.0 mm), sagittal three-dimensional dual-echo in steady state sequence (16.3/4.7; spatial resolution, 0.365 × 0.456 mm; section thickness, 0.7 mm), coronal two-dimensional intermediate-weighted fast spin-echo sequence (3700/29; spatial resolution, 0.365 × 0.456 mm; section thickness, 3.0 mm), and a three-dimensional fast low-angle shot sequence with selective water excitation (20/7.57; spatial resolution, 0.313 × 0.313 mm; section thickness, 1.5 mm) (22).

#### Image Analyses

MR images of the right knee obtained at baseline and after 2 years were reviewed at picture archiving and communication system workstations (Agfa, Ridgefield Park, NJ) by two radiologists (W.V, T.M.L.; 8 and 25 years of experience in musculoskeletal imaging, respectively). As part of their initial training, the two radiologists analyzed 10 MR imaging studies in consensus to calibrate thresholds for grading with the different scoring systems. The two radiologists then independently analyzed the MR imaging studies for all subjects by using the semiquantitative WORMS and BLOKS systems and the quantitative CaLS system for quantification of cartilage defects. The grading with each scoring system was performed at 3-week intervals to minimize recall. WORMS scoring was performed first and was followed by CaLS scoring and then BLOKS scoring. The order of the subjects was randomized for each analysis to further prevent recall. The radiologists were blinded to subject identifiers, demographics, clinical history, and the OAI cohort (incidence vs normal). The radiologists were not blinded to the sequence of images (ie, baseline or 2-year follow-up). The grading of a subset of 10 randomly selected

subjects was timed to assess the average time required to perform measurements with the three scoring systems.

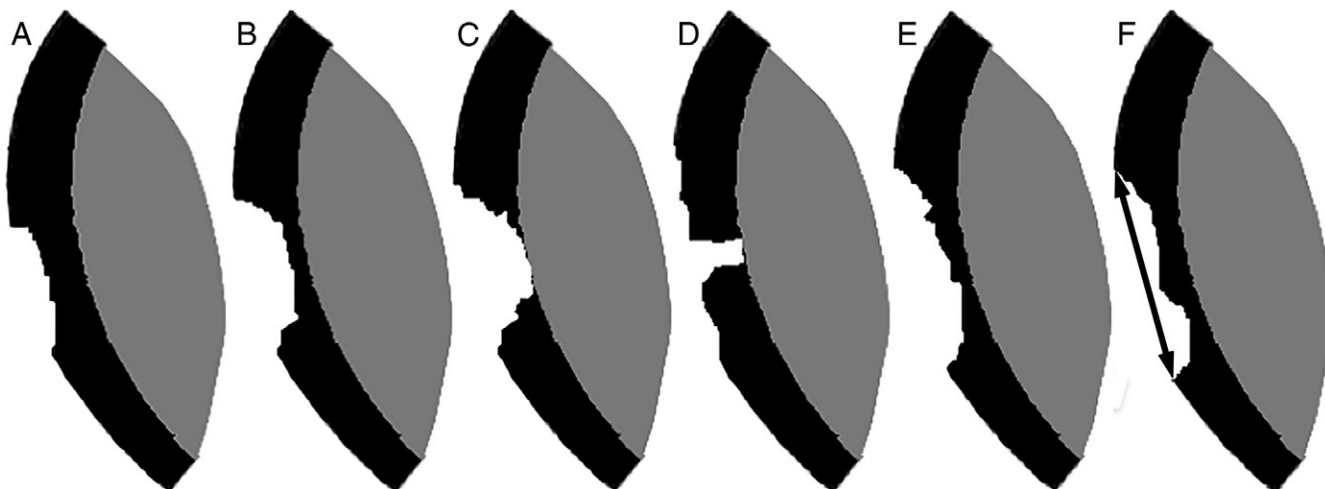
#### Semiquantitative Morphologic Assessment

**WORMS scoring for cartilage.**—Cartilage lesions are scored with the WORMS system on an eight-point scale, as follows: 0 indicates normal cartilage; 1, increased signal with fluid-sensitive intermediate-weighted sequences; 2, partial-thickness defect less than 1 cm in greatest width; 2.5, full-thickness defect less than 1 cm in greatest width; 3, multiple areas of partial-thickness (grade 2) defects intermixed with areas of normal thickness or a partial-thickness defect wider than 1 cm but less than 75% of the region; 4, diffuse ( $\geq 75\%$  of the region) partial-thickness loss; 5, multiple areas of full-thickness loss (grade 2.5) or a full-thickness defect wider than 1 cm but less than 75% of the region; and 6, diffuse ( $\geq 75\%$  of the region) full-thickness loss (11). Cartilage abnormalities were assessed by using a modified WORMS system (11), in which the number of anatomic compartments was reduced from 15 to six (patella, trochlea, medial femur, medial tibia, lateral femur, and lateral tibia) (23). This modified version of the WORMS system was developed to more efficiently grade cartilage lesions in subjects with relatively mild cartilage abnormalities. To examine the effect of reducing compartments on the detection of cartilage lesions with the WORMS system, the original WORMS (15 compartments) scoring was performed in the same studies after an interval of 5 months, and the results were compared.

In this article, we do not report cartilage signal abnormalities (grade 1) because these cannot be measured with the quantitative scoring system.

**BLOKS scoring for cartilage.**—The BLOKS system consists of two separate features for the assessment of cartilage abnormalities; these are a subregional or cartilage I score and a cartilage II score that describe cartilage integrity in specific locations in a predefined coronal image section (12). For the purpose

Figure 1



**Figure 1:** Illustrations show definitions of depth and shape scores, as well as diameter, with the CaLS system. Black = cartilage, gray = subchondral bone. *A*, A superficial lesion (less than 50% deep) is scored as 1. *B*, A lesion deeper than 50% is scored as 2. *C*, A full-thickness lesion is scored as 3. *D*, If maximum depth occupies less than 50% of the lesion, it is scored as 0.5. *E*, If maximum depth occupies 50% or more of the lesion, it is scored as 1. *F*, The largest diameter (double arrow) is measured.

of consistency between analysis with the three scoring systems, we assessed only abnormalities with the cartilage I score, which assigns two separate scores (12,17). The first score enables us to assess size of any cartilage loss as a percentage of surface area as related to the size of each individual region: a score of 0 indicates no cartilage loss; a score of 1, less than 10% of the region of cartilage surface area; a score of 2, 10%–75% of the region of cartilage surface area; and a score of 3, more than 75% of the region of cartilage surface area. The second score enables us to assess the percentage of full-thickness cartilage loss of the region: a score of 0 indicates no cartilage loss; a score of 1, less than 10% of the region of cartilage surface area; a score of 2, 10%–75% of the region of cartilage surface area; and a score of 3, more than 75% of the region of cartilage surface area. The BLOKS system was used in eight of nine subregions (medial and lateral patella, trochlea, femur, tibia, and the tibial interspinous region). The tibial spines were devoid of cartilage.

#### Quantitative Morphologic Assessment

The CaLS system measures the three-dimensional volume of cartilage defects

(WORMS grade >2); the extent of visible cartilage defects is quantified by multiplying the five features of the lesion assessed, as follows:  $CaLS = L \times N \times T \times D \times S$ , where  $L$  is the largest diameter (in millimeters),  $N$  is the number of sections,  $T$  is the section thickness including the section gap (in millimeters),  $D$  is the depth, and  $S$  is the shape factor. To evaluate depth of the lesion, we visually dissected the cartilage into two equal halves (Figs 1, 2). If the maximum lesion depth was less than 50%, it was assigned a value of 1. If the maximum lesion depth exceeded 50%, it was assigned a value of 2. A full-thickness lesion was assigned a value of 3. The shape factor was 1 if the maximum depth occupied more than 50% of the lesion surface diameter (assessed in the section with the largest diameter); a shape factor of 0.5 was assigned if the maximum depth occupied less than half of the lesion surface diameter.

We did not encounter any patients who had multiple lesions per compartment in this study; however, we propose that observers should assign separate scores for individual lesions and use the sum of these scores to

obtain a total compartment score in such patients.

#### Assessment of Reliability of CaLS System

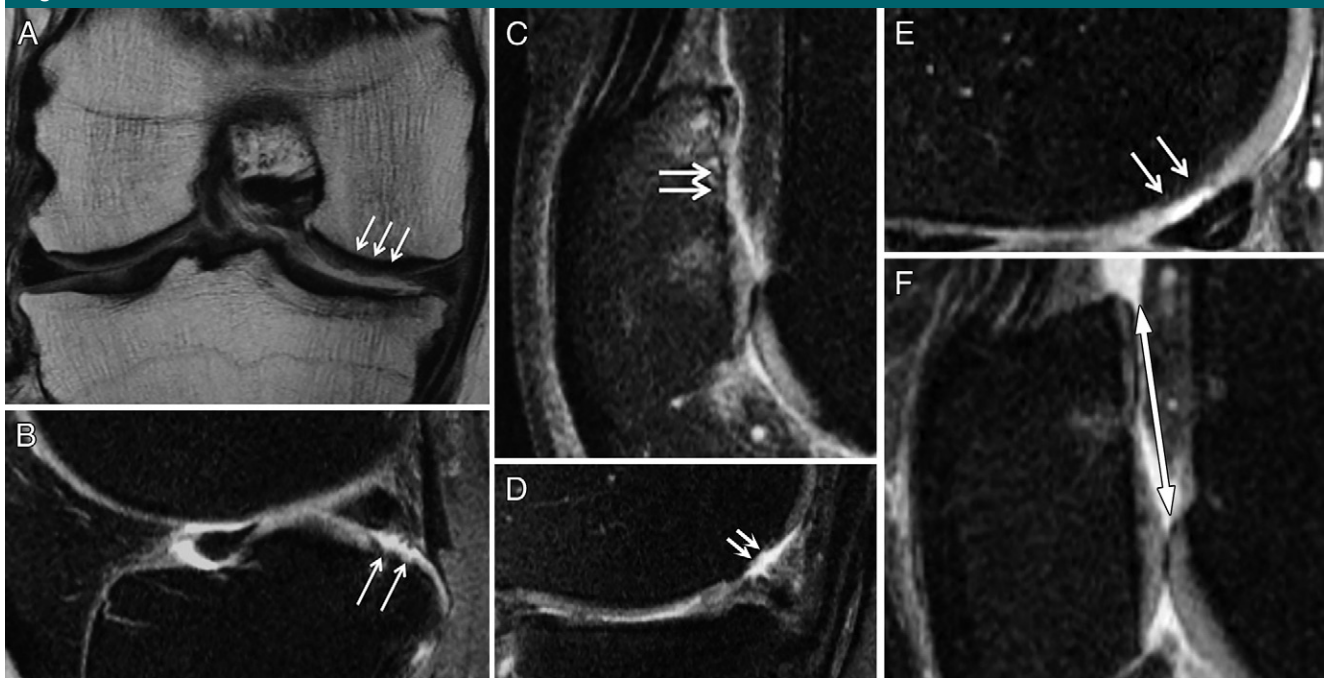
Two radiologists (W.V, T.M.L; 8 and 25 years of experience in musculoskeletal imaging, respectively) independently graded 20 randomly selected studies twice by using the CaLS system. After a 1-month interval, the grading was repeated to determine intraobserver reliability. An intraclass correlation coefficient was calculated to determine inter- and intraobserver reliability of the total CaLS system (24). The Cohen  $\kappa$  statistic was calculated to determine inter- and intraobserver reliability of individual features of the CaLS system (25).

#### Statistical Analyses

Statistical analysis was performed by using statistical software (Stata, version 11; Stata, College Station, Tex). Descriptive statistics were calculated for all subjects. The Wilcoxon Mann-Whitney test was used to compare mean age and BMI between groups.

Progression of cartilage defects for each score system was defined as a difference of greater than 0 from baseline to 24-month follow-up for

Figure 2



**Figure 2:** MR images show definitions of depth and shape scores, as well as diameter, with the CaLS system. *A*, A superficial lesion (<50% deep) (arrows) was given a depth score of 1 in a 53-year-old man. *B*, A lesion deeper than 50% (arrows) was given a depth score of 2 in a 55-year-old woman. *C*, A full-thickness lesion (arrows) was given a depth score of 3 in a 53-year-old woman. *D*, If maximum depth (arrows) occupies less than 50% of the lesion, as in this 55-year-old man, it is scored as 0.5. *E*, Otherwise, maximum depth (arrows) is scored as 1, as in this 52-year-old woman. *F*, Largest diameter (double arrow) is measured in a 56-year-old woman.

the individual knee compartments. The progression measured with the three scoring systems was compared by using multilevel mixed-effects logistic regression models, with repeated measures by knee compartment. Adjustments were performed for the compartment of the knee but not for subject characteristics (age, BMI, sex), as these do not affect direct comparison of the three scoring systems. Odds ratios were calculated for the progression of lesions in the subjects with risk factors for OA and those without risk factors for OA when using the three different scoring systems.

## Results

### Baseline Subject Characteristics

Subject characteristics are described in detail in Table 1. The mean age of the 25 subjects in the healthy cohort

was 51.5 years, and the mean BMI was 23.9 kg/m<sup>2</sup>. The mean age of the 52 subjects with risk factors for OA was 51.0 years, and the mean BMI was 23.8 kg/m<sup>2</sup>. There was no significant difference between the two subject groups in terms of age or BMI. To observe the difference in progression between the subjects with risk factors and those without, we did not adjust for OA risk factors.

### Quantitative and Semiquantitative Assessment of Cartilage Lesions

For each scoring system, 462 compartments of the knee were studied. At baseline, a total of 24 compartments (16.0%) in healthy subjects without OA risk factors and 84 compartments (26.9%) in subjects with OA risk factors had cartilage lesions. Patellar lesions were most common at baseline; 42 patellar compartments had one lesion each when we assessed all 77 subjects. Table 2 shows the number of

compartments that progressed from baseline to 24-month follow-up in all subjects, as assessed with the CaLS, WORMS, and BLOKS systems. There was no significant difference between the modified (six compartments) and original WORMS systems in terms of detection of lesion progression ( $P > .99$ ); the results presented were obtained by using the modified WORMS system. There was substantially more progression with the CaLS system than with the WORMS or BLOKS systems. When measured with the CaLS system, the largest percentage of progression was observed in the medial femur.

High-grade lesions were rare in the cohort studied; at baseline, 3.7% of all lesions were WORMS grade 4, 12.0% were WORMS grade 5, and none were WORMS grade 6.

Average times of 8.8, 9.35, and 12.55 minutes were required when we used the BLOKS, CaLS, and WORMS systems, respectively.

**Table 1****Subject Characteristics**

Characteristic	Subjects without OA Risk Factors ( <i>n</i> = 25)	Subjects with OA Risk Factors ( <i>n</i> = 52)
<b>Sex</b>		
Female	20	31
Male	5	21
BMI (kg/m <sup>2</sup> )*	23.9	23.8
Age (y)*	51.4 (47–55)	51.0 (45–60)
BMI ≥ 24 kg/m <sup>2</sup>	12	26
Prior knee injury	0	26
Prior knee surgery	0	13
Family history of knee replacement	0	12
Heberden nodes	0	9
Frequent knee bending activity	16	40

Note.—Unless otherwise indicated, data are numbers of patients.

\* Data are the mean. Data in parentheses (if any) are the range.

**Table 2****Total Number of Lesions in All 77 Subjects Progressing from Baseline to 24-month Follow-up as Assessed with CaLS, WORMS, and BLOKS Systems**

Compartment	Total No. of Compartments with Lesions at Baseline	No. of Lesions that Progressed		
		CaLS	WORMS	BLOKS
Patella	42	17	3	3
Trochlea	18	12	4	2
Medial femur	17	12	5	3
Lateral tibia	15	8	3	4
Lateral femur	13	5	1	0
Medial tibia	3	2	3	2

**Detection of Cartilage Lesion Progression by Using Quantitative versus Semiquantitative Scoring Systems**

Table 2 shows the number of compartments that progressed when the MR images of all subjects were assessed with the three different scoring systems: the CaLS system demonstrated higher detection of cartilage defect progression than did the semiquantitative scores in all individual compartments of the knee except the medial tibia. There were only three medial tibias with lesions among all subjects. When we assessed progression across all compartments in subjects with and those without OA risk factors, 51.8% (56 of 108), 17.6% (19 of 108), and 13.0% (14 of 108) of the lesions progressed when analyzed

with the CaLS, WORMS, and BLOKS systems, respectively; the CaLS system was found to have a higher detection of cartilage lesion progression than either the WORMS system or the BLOKS system ( $P < .001$ ). Figure 3 shows a patellar cartilage lesion progressing from baseline to 24-month follow-up as measured with the CaLS system but not with the BLOKS or WORMS systems. The difference between BLOKS and WORMS systems regarding measurement of progression was not significant ( $P = .101$ ).

**Reproducibility of CaLS System**

Intraclass correlation coefficients calculated for intra- and interobserver agreement of the quantitative analysis

of images based on the CaLS system were 0.86 and 0.91, respectively. Table 3 comprises the intra- and interobserver Cohen  $\kappa$  measurements for the individual features of the CaLS system. All features demonstrated excellent reproducibility, with interobserver  $\kappa$  values ranging from 0.81 to 0.94; shape factor and diameter measurements had the highest level of agreement.

**Association of CaLS-detected Longitudinal Change in Cartilage Lesions with Known Risk Factors for OA Progression**

With the CaLS system, the subjects with risk factors had a significantly higher odds ratios of progression than did subjects without risk factors (odds ratio, 2.78; 95% confidence interval: 1.36, 5.7;  $P = .005$ ). With the WORMS system, the difference in detection of progression between the subjects with and subjects without OA risk factors was not significant (odds ratio, 1.44; 95% confidence interval: 0.40, 5.13;  $P = .571$ ). With the BLOKS system, the subjects with risk factors did not have significantly different odds of progression than did subjects without risk factors (odds ratio, 7.09; 95% confidence interval: 0.92, 54.6;  $P = .060$ ). The small number of lesions that progressed with the BLOKS system were exclusively in the group with risk factors for OA, leading to the surprisingly high odds ratio; however, the confidence interval was very large, and this result was not significant.

**Discussion**

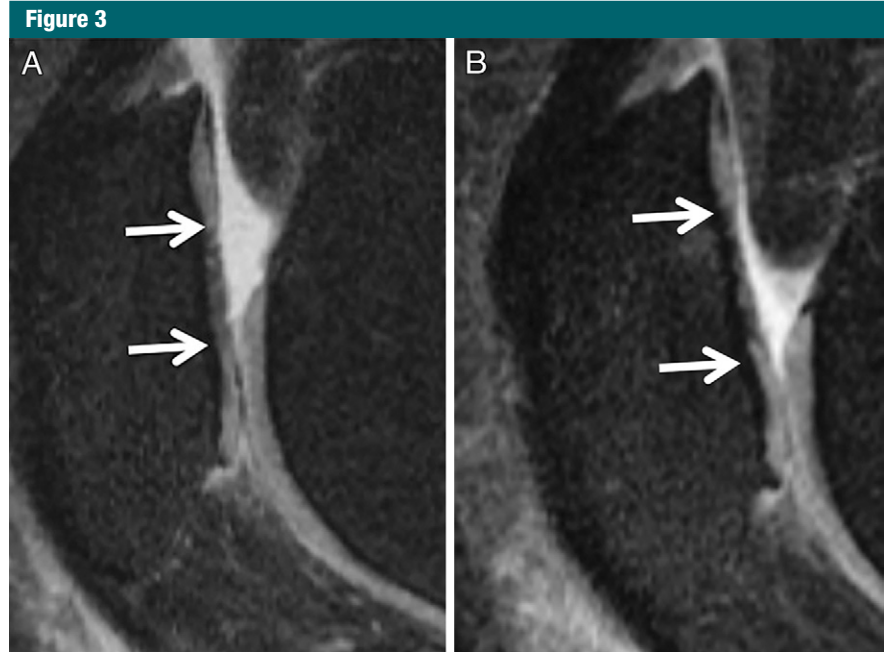
In this article, we describe a scoring system specifically designed for longitudinal follow-up of relatively mild cartilage defects on MR images. The score was found to have excellent reproducibility and demonstrated better detection of cartilage defect progression than the established semiquantitative WORMS and BLOKS grading systems. The clinical validity of the higher progression measured with the CaLS system was confirmed by comparing progression in a cohort of subjects with risk factors for OA and a

cohort of subjects without risk factors for OA; cartilage lesions assessed with the CaLS system in subjects with risk factors for OA had significantly higher odds of progression than did those in healthy subjects without any risk factors for OA, while odds ratios were not significant for either the WOMBS or BLOKS systems.

The Food and Drug Administration and other regulatory agencies still recommend joint space narrowing on radiographs in addition to pain and function as coprimary endpoints to establish the effectiveness of disease-modifying drugs (1). The delayed intervals required for measurement of clinically important changes on radiographs (4) and the discordance between radiographic and symptomatic manifestations (26) creates a need for a faster, more sensitive, and more reproducible instrument with which to evaluate OA disease progression. MR imaging allows unparalleled visualization of all tissues involved, and it has been shown to be more sensitive than radiography or arthroscopy after a short period of 1 year in the assessment of the progression of OA-related chondropathy (27).

The reliability and validity of MR imaging-based semiquantitative systems, such as the WOMBS (11) and BLOKS (12) systems, have been established. However, these grading systems are complex, and they have limitations when used to monitor changes in OA-related cartilage degeneration. Some research groups have used methods to detect within-grade changes to increase the sensitivity of these scales (16); these methods were published shortly after we completed the CaLS study presented herein. In contrast to categorical variables of semiquantitative scoring systems, the CaLS score provides continuous variables. Continuous variables may increase the possibility of detecting clinically unimportant change; however, on the other hand, use of categorical variables may result in loss of valuable information, which is particularly critical during longitudinal follow-up studies (28–30).

The feasibility of using a quantitative scoring system to evaluate cartilage defects has been a subject of previous



**Figure 3:** Sagittal intermediate-weighted fast spin-echo MR images of the right knee in a 49-year-old woman from the OAI incidence cohort show a patellar lesion that progressed with the CaLS system but not with the WOMBS or BLOKS systems. At both, *A*, baseline and, *B*, 24-month follow-up, the lesion was scored as WOMBS grade 3 and BLOKS grade 2. With the CaLS system, the lesion measured 11 mm at baseline and had a partial thickness of less than 50%; at 24-month follow-up, the lesion measured 14 mm and had a partial thickness of more than 50%.

**Table 3**

**Reproducibility Assessed in 20 Subjects, with  $\kappa$  Values Provided for Individual Features of the CaLS System**

CaLS Feature	Intraobserver Agreement	Interobserver Agreement
Largest diameter	0.88 (0.86, 0.89)	0.93 (0.91, 0.93)
Depth	0.89 (0.83, 0.92)	0.81 (0.77, 0.89)
Shape factor	0.96 (0.75, 0.99)	0.94 (0.91, 0.98)
No. of sections	0.88 (0.87, 0.89)	0.87 (0.85, 0.89)

Note.—Data in parentheses are 95% confidence intervals.

studies (31,32). The CaLS system essentially calculates a three-dimensional volume of a cartilage defect based on the manual measurements obtained with MR imaging. The largest diameter of the defect on any given section and the depth of the lesion are features of the score that were most commonly observed to worsen after 2 years. The number of sections in which the lesion was visualized and the shape of the lesion were the factors that were least likely to change. The depth measured

on a scale of 0–3 can be considered continuous for statistical purposes; accurate manual measurement of a depth comprising roughly four to six pixels is impractical. Similarly, the shape factor can be considered a continuous variable, and it provides important information about the lesion; it describes the depth of the lesion as occupying more or less than 50% of the total surface. A benefit of the shape factor is that it can be used to distinguish small full-thickness lesions from large full-thickness

lesions. The average time required to perform CaLS assessment in one knee was 9.35 minutes.

The face validity of the increased progression measured with the CaLS system is easily recognizable; cartilage defects that progress over time may be scored as not changed if they are graded based on a scale such as that used in the WORMS or BLOKS system. Since early structural changes in the joint often precede clinical symptoms of OA, validation of a scoring system for cartilage degeneration by correlation with symptoms is not optimal (33). Clinical characteristics such as history of knee trauma or surgical procedure are validated predisposing factors for knee OA (34). Thus, we compared the three scoring systems in healthy subjects and subjects with risk factors for OA. When we used the CaLS system, the cartilage lesions in subjects with risk factors for OA had significantly higher odds of progression than did those in healthy subjects without any risk factors, while neither the WORMS system nor the BLOKS system had a significant odds ratio.

A limitation of our study was the relatively small sample size of both healthy subjects and subjects with risk factors for OA. Another limitation of the study was the fact that the majority of the cartilage lesions were discovered in the patella and the risk factors for patellofemoral cartilage degeneration may be different from those for OA. In addition, although the radiologists who analyzed the images were not blinded to order in terms of time, such blinding has been shown to improve sensitivity without affecting reliability (35). Comparison of the CaLS system with the other scoring systems was performed by individuals who developed the CaLS system; future comparisons should be performed by individuals who were not involved in the development of this system to reduce the possibility of bias. Signal inhomogeneity was not included in this study because these lesions are not measurable. Moreover, Saadat et al showed that signal changes have a

limited correlation with histopathologic findings (36). It should also be noted that there was no reference standard in the assessment of cartilage lesions, and none of these subjects underwent arthroscopy; however, it is unlikely that individuals with mild OA would undergo arthroscopy in the clinical setting. Moreover, sensitivity and specificity of 3-T MR imaging for cartilage lesions have been published previously (36). We modified the WORMS score to more efficiently grade early degenerative changes for the current and future studies. The reduction in the number of compartments analyzed had no significant effect on the detection of progression when compared with the original WORMS system. Relatively recently, an optimized semiquantitative MOAKS system was introduced (13) that was derived from the WORMS and BLOKS systems. An additional limitation of this study is that the CaLS system was not compared with the MOAKS system; however, it should be noted that only one article on the MOAKS system (13) was available when we submitted our manuscript.

In summary, the proposed CaLS system is a reproducible and valid scoring system for cartilage lesions. The CaLS system is an MR-based technique with improved detection of longitudinal change in cartilage lesions compared with the established WORMS and BLOKS systems and may be a preferred outcome measure with which to examine effectiveness of therapeutic interventions in clinical and epidemiologic studies of OA.

**Disclosures of Conflicts of Interest:** A.H. No relevant conflicts of interest to disclose. W.V. No relevant conflicts of interest to disclose. G.B.J. No relevant conflicts of interest to disclose. L.N. No relevant conflicts of interest to disclose. E.L. No relevant conflicts of interest to disclose. H.L. No relevant conflicts of interest to disclose. M.C.N. No relevant conflicts of interest to disclose. J.A.L. No relevant conflicts of interest to disclose. C.E.M. No relevant conflicts of interest to disclose. T.M.L. Financial activities related to the present article: none to disclose. Financial activities not related to the present article: is a consultant for GE, receives royalties from Springer, institution received a grant for GE. Other relationships: none to disclose.

## References

1. Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: an update with relevance for clinical practice. *Lancet* 2011;377(9783):2115–2126.
2. Altman R, Alarcón G, Appelrouth D, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheum* 1991; 34(5):505–514.
3. Hunter DJ, Zhang W, Conaghan PG, et al. Responsiveness and reliability of MRI in knee osteoarthritis: a meta-analysis of published evidence. *Osteoarthritis Cartilage* 2011;19(5):589–605.
4. Boegård TL, Rudling O, Petersson IF, Jansson 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):370–376.
5. Lane NE, Brandt K, Hawker G, et al. OARSI-FDA initiative: defining the disease state of osteoarthritis. *Osteoarthritis Cartilage* 2011; 19(5):478–482.
6. Felson DT, Chaisson CE, Hill CL, et al. The association of bone marrow lesions with pain in knee osteoarthritis. *Ann Intern Med* 2001; 134(7):541–549.
7. Felson DT, McLaughlin S, Goggins J, et al. Bone marrow edema and its relation to progression of knee osteoarthritis. *Ann Intern Med* 2003;139(5 Pt 1):330–336.
8. Oegema TR Jr, Carpenter RJ, Hofmeister F, Thompson RC Jr. The interaction of the zone of calcified cartilage and subchondral bone in osteoarthritis. *Microsc Res Tech* 1997; 37(4):324–332.
9. Sellam J, Berenbaum F. The role of synovitis in pathophysiology and clinical symptoms of osteoarthritis. *Nat Rev Rheumatol* 2010;6(11):625–635.
10. Hunter DJ, Zhang YQ, Niu JB, et al. The association of meniscal pathologic changes with cartilage loss in symptomatic knee osteoarthritis. *Arthritis Rheum* 2006;54(3): 795–801.
11. Peterfy CG, Guermazi A, Zaim S, et al. Whole-Organ Magnetic Resonance Imaging Score (WORMS) of the knee in osteoarthritis. *Osteoarthritis Cartilage* 2004;12(3): 177–190.
12. Hunter DJ, Lo GH, Gale D, Grainger AJ, Guermazi A, Conaghan PG. The reliability of a new scoring system for knee osteoarthritis MRI and the validity of bone marrow lesion assessment: BLOKS (Boston Leeds Osteoarthritis Knee Score). *Ann Rheum Dis* 2008;67(2):206–211.



13. Hunter DJ, Guermazi A, Lo GH, et al. Evolution of semi-quantitative whole joint assessment of knee OA: MOAKS (MRI Osteoarthritis Knee Score). *Osteoarthritis Cartilage* 2011;19(8):990–1002.
14. Hunter DJ, Conaghan PG, Peterfy CG, et al. Responsiveness, effect size, and smallest detectable difference of magnetic resonance imaging in knee osteoarthritis. *Osteoarthritis Cartilage* 2006;14(Suppl A):A112–A115.
15. Conaghan PG, Tennant A, Peterfy CG, et al. Examining a whole-organ magnetic resonance imaging scoring system for osteoarthritis of the knee using Rasch analysis. *Osteoarthritis Cartilage* 2006;14(Suppl A):A116–A121.
16. Roemer FW, Nevitt MC, Felson DT, et al. Predictive validity of within-grade scoring of longitudinal changes of MRI-based cartilage morphology and bone marrow lesion assessment in the tibio-femoral joint: the MOST study. *Osteoarthritis Cartilage* 2012;20(11):1391–1398.
17. Lynch JA, Roemer FW, Nevitt MC, et al. Comparison of BLOKS and WOMBS scoring systems. I. Cross sectional comparison of methods to assess cartilage morphology, meniscal damage and bone marrow lesions on knee MRI: data from the osteoarthritis initiative. *Osteoarthritis Cartilage* 2010;18(11):1393–1401.
18. Felson DT, Lynch J, Guermazi A, et al. Comparison of BLOKS and WOMBS scoring systems. II. Longitudinal assessment of knee MRIs for osteoarthritis and suggested approach based on their performance: data from the Osteoarthritis Initiative. *Osteoarthritis Cartilage* 2010;18(11):1402–1407.
19. Bay-Jensen A-C, Hoegh-Madsen S, Dam E, et al. Which elements are involved in reversible and irreversible cartilage degradation in osteoarthritis? *Rheumatol Int* 2010;30(4):435–442.
20. Blum A, Roch D, Loeuille D, et al. Bone marrow edema: definition, diagnostic value and prognostic value [in French]. *J Radiol* 2009;90(12):1789–1811.
21. Felson DT, Nevitt MC. Epidemiologic studies for osteoarthritis: new versus conventional study design approaches. *Rheum Dis Clin North Am* 2004;30(4):783–797, vii.
22. Peterfy CG, Schneider E, Nevitt M. The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee. *Osteoarthritis Cartilage* 2008;16(12):1433–1441.
23. Stehling C, Lane NE, Nevitt MC, Lynch J, McCulloch CE, Link TM. Subjects with higher physical activity levels have more severe focal knee lesions diagnosed with 3T MRI: analysis of a non-symptomatic cohort of the osteoarthritis initiative. *Osteoarthritis Cartilage* 2010;18(6):776–786.
24. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;86(2):420–428.
25. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159–174.
26. 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.
27. Pessis E, Drapé JL, Ravaud P, Chevrot A, Dougados M, Ayrat X. Assessment of progression in knee osteoarthritis: results of a 1 year study comparing arthroscopy and MRI. *Osteoarthritis Cartilage* 2003;11(5):361–369.
28. Altman DG. Problems in dichotomizing continuous variables. *Am J Epidemiol* 1994;139(4):442–445.
29. Dawson NV, Weiss R. Dichotomizing continuous variables in statistical analysis: a practice to avoid. *Med Decis Making* 2012;32(2):225–226.
30. Naggara O, Raymond J, Guilbert F, Roy D, Weill A, Altman DG. Analysis by categorizing or dichotomizing continuous variables is inadvisable: an example from the natural history of unruptured aneurysms. *AJNR Am J Neuroradiol* 2011;32(3):437–440.
31. Stahl R, Jain SK, Lutz J, 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 Radiol* 2011;40(10):1315–1327.
32. Stahl R, Luke A, Ma CB, et al. Prevalence of pathologic findings in asymptomatic knees of marathon runners before and after a competition in comparison with physically active subjects: a 3.0 T magnetic resonance imaging study. *Skeletal Radiol* 2008;37(7):627–638.
33. Manno RL, Bingham CO 3rd, Paternotte S, et al. OARSI-OMERACT initiative: defining thresholds for symptomatic severity and structural changes in disease modifying osteoarthritis drug (DMOAD) clinical trials. *Osteoarthritis Cartilage* 2012;20(2):93–101.
34. Blagojevic M, Jinks C, Jeffery A, Jordan KP. Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis. *Osteoarthritis Cartilage* 2010;18(1):24–33.
35. Gensburger D, Roux JP, Arlot M, Sornay-Rendu E, Ravaud P, Chapurlat R. Influence of blinding sequence of radiographs on the reproducibility and sensitivity to change of joint space width measurement in knee osteoarthritis. *Arthritis Care Res (Hoboken)* 2010;62(12):1699–1705.
36. Saadat E, Jobke B, Chu B, et al. Diagnostic performance of in vivo 3-T MRI for articular cartilage abnormalities in human osteoarthritic knees using histology as standard of reference. *Eur Radiol* 2008;18:2292–2302.