Title
Development of a Comprehensive Osteochondral Allograft MRI Scoring System (OCAMRISS) With Histopathologic, Micro-Computed Tomography, and Biomechanical Validation

Permalink
https://escholarship.org/uc/item/8q93t2cb

Journal
Cartilage, 5(1)

ISSN
1947-6035

Authors
Chang, EY
Pallante-Kichura, AL
Bae, WC
et al.

Publication Date
2014

DOI
10.1177/1947603513514436

Peer reviewed
Development of a Comprehensive Osteochondral Allograft MRI Scoring System (OCAMRISS) With Histopathologic, Micro–Computed Tomography, and Biomechanical Validation

Eric Y. Chang1,2, Andrea L. Pallante-Kichura3, Won C. Bae2, Jiang Du2, Sheronda Statum2, Tanya Wolfson2, Anthony C. Gamst2, Esther Cory3, David Amiel4, William D. Bugbee4,5, Robert L. Sah3,4, and Christine B. Chung1,2

Abstract
Objective: To describe and apply a semiquantitative MRI scoring system for multifeature analysis of cartilage defect repair in the knee by osteochondral allografts and to correlate this scoring system with histopathologic, micro–computed tomography (µCT), and biomechanical reference standards using a goat repair model. Design: Fourteen adult goats had 2 osteochondral allografts implanted into each knee: one in the medial femoral condyle and one in the lateral trochlea. At 12 months, goats were euthanized and MRI was performed. Two blinded radiologists independently rated 9 primary features for each graft, including cartilage signal, fill, edge integration, surface congruity, calcified cartilage integrity, subchondral bone plate congruity, subchondral bone marrow signal, osseous integration, and presence of cystic changes. Four ancillary features of the joint were also evaluated, including opposing cartilage, meniscal tears, synovitis, and fat-pad scarring. Comparison was made with histologic and µCT reference standards as well as biomechanical measures. Interobserver agreement and agreement with reference standards was assessed. Cohen’s κ, Spearman’s correlation, and Kruskal-Wallis tests were used as appropriate. Results: There was substantial agreement (κ > 0.6, P < 0.001) for each MRI feature and with comparison against reference standards, except for cartilage edge integration (κ = 0.6). There was a strong positive correlation between MRI and reference standard scores (ρ = 0.86, P < 0.01). Osteochondral allograft MRI scoring system was sensitive to differences in outcomes between the types of allografts. Conclusions: We have described a comprehensive MRI scoring system for osteochondral allografts and have validated this scoring system with histopathologic and µCT reference standards as well as biomechanical indentation testing.

Keywords
cartilage repair, osteochondral allografts, MRI scoring system

Introduction
Localized articular cartilage lesions in young patients are relatively common,1 and treatment remains a challenge. A number of repair strategies exist, including microfracture, osteochondral autograft transfer, and cell-based transplantation techniques, but each has its own limitations.2-8 Osteochondral allograft transplantation is an attractive technique that allows for the simultaneous resurfacing of large articular defects and the correction of underlying bone abnormalities. In some ways, osteochondral allografts are ideal for transplantation because cartilage is a relatively immunoprivileged tissue,9 mature living chondrocytes can survive for many years after transplantation without tissue matching or immunosuppression,10 and bone has the potential to heal and remodel.11-13
reliable comparison between different therapeutic approaches. Unlike histopathologic and surgical scoring systems, however, there are a paucity of studies which have validated MRI scoring systems. Recent studies have shown a lack of strong evidence for MRI to predict clinical outcome, including with the widely used Magnetic Resonance Observation of Cartilage Repair Tissue (MOCART) score. Validation of the individual components of a noninvasive scoring system is a logical first step to understanding this discord.

We have used and evaluated a number of osteochondral allografts in our practice and have recognized the need for accurate noninvasive characterization. Furthermore, while imaging grading systems that are used to compare various cartilage repair techniques are useful, we have found that they may not be optimized for comparing features unique to a certain technique. For instance, grading systems that are optimized to address cartilage defect repair with cell-based transplantation techniques (such as the MOCART score) are not optimal to evaluate techniques that involve osteochondral transfer en bloc. This is further highlighted with the introduction of the 3D MOCART score, which includes “bone interface,” which describes integration of the transplant to the native subchondral bone as well as integration of a possible periosteal flap and “chondral osteophytes,” which are a complication described with cartilage transplantation and bone marrow stimulating techniques. Additionally, to our knowledge, there are no studies that systematically compare a comprehensive grading system with reference standard validation. In part this is because of the fact that most means for validation are invasive, including second look arthroscopy and biopsy. Animal models are useful to circumvent these limitations.

The purpose of this study was to (1) describe and apply a semiquantitative MRI scoring system for multifeature analysis of cartilage defect repair by osteochondral allografts and (2) to correlate this scoring system with histopathologic, micro–computed tomography (µCT), and biomechanical reference standards using a goat repair model.

**Methods**

Tissues analyzed were from a previously published study in an adult Boer goat model. Although the previous study used 15 goats, only 14 were imaged and included in this study. All experiments were carried out in accordance with protocols approved by the Institutional Animal Care and Use Committee. In brief, each goat was operated in one knee and each knee received 2 orthotopic osteochondral allografts measuring 8 × 5 mm (diameter × depth). Grafts were implanted into 7.5 × 5 mm (diameter × depth) defects in the medial femoral condyle (MFC) and lateral trochlea (LT). Four different groups of graft storage were analyzed, including fresh (n = 7), 14-day storage at 4 °C (n = 7), 28-day storage at 4 °C (n = 7), and frozen grafts (n = 7). At 12 months, goats were euthanized and MRI was performed and compared with histopathologic, µCT, and biomechanical analysis.

**MRI Technique**

Imaging was performed on a 3T clinical MRI scanner (Signa Twinspeed, GE Healthcare, Milwaukee, WI) with a wrist coil. Hardware modification included an addition of a custom transmit–receive switch to the receiver preamplifiers for rapid switching after the end of a radiofrequency excitation pulse, which allows for detection of signal as early as 8 µs. Imaging parameters are listed in Table 1. In brief, 2D imaging sequences consisted of 3-plane intermediate-weighted fast spin echo (FSE) sequences, sagittal and axial images.
STIR and proton density-weighted conventional spin echo sequences, and sagittal T1-weighted FSE sequences. Additionally, sagittal 2D and 3D multi-echo ultrashort echo time (UTE) sequences with fat saturation were obtained with subtraction of the second echo from the first to generate images that highlight short T2* tissues. However, for this study, the whole UTE data set was not used, but rather only the subtraction images for evaluation of the deep cartilage of the graft (feature 5, see below). Although the 3D UTE source images are higher in spatial resolution and sensitive to both long and short T2* tissues, we felt there may have been limited translatability had we used the source images for comprehensive MRI feature analysis since not all investigators have access to this particular sequence. Therefore, the MRI readers were blinded to the UTE source images.

Our imaging resolution parameters were modified from our typical clinical knee examination to account for the relatively smaller dimensions of the goat knee. Specifically, we calculated that volumes of goat distal femoral epiphyses used in our study were approximately 24% of the volume of human distal femoral epiphyses. Therefore, voxels used in this study were scaled to 24% of that used in our clinical knee protocols.

**MRI Analysis**

Images were evaluated independently by 2 blinded fellowship-trained musculoskeletal radiologists (reader A [EYC] and reader B [CBC], with 2 and 15 years of experience, respectively). Based on our experience with osteochondral allografts, we have created the semiquantitative OsteoChondral Allograft MRI Scoring System (OCAMRISS), which includes 9 primary features of the graft (5 cartilage and 4 bone) and 4 ancillary features of the joint, as shown in Table 2. Each MRI-determined primary feature also had a corresponding reference standard, determined either histopathologically or on µCT. Specifically, primary features for cartilage included (1) cartilage signal of the graft relative to adjacent host cartilage, (2) cartilage “fill” of the graft (percentage volume), (3) cartilage edge integration at the host-graft junction, (4) cartilage surface congruity of the graft and host-graft junction, (5) calcified cartilage integrity of the graft, and for bone included (6) subchondral bone plate congruity of the graft and host-graft junction, (7) subchondral bone marrow signal intensity of the graft relative to epiphyseal bone, (8) osseous integration at the host-graft junction, and (9) presence of cystic changes of the graft and host-graft junction. Ancillary features of the joint were evaluated using similar criteria as in multiple prior studies and included detecting abnormalities of (10) opposing medial tibial plateau cartilage, (11) meniscal tears, (12) synovitis, and (13) infrapatellar fat pad scarring.

**Corresponding Reference Standards**

As part of a prior study, µCT, histopathology, and biomechanical indentation testing was performed on retrieved tissue after euthanization. The MRI results were compared with modified Mankin (MM) scores from that study. In brief, sagittal slices through the central portion of the grafts were prepared and stained at 7 µm thickness and the MM score was used to assess the graft cartilage with a minimum score of 0, representing normal cartilage, and a maximum score of 15, representing the highest degree of degeneration. The MRI results were also compared with cartilage loadbearing function data from that prior study. In brief, indentation testing was performed at the center of each graft and stiffness was calculated (expressed in units of MPa).

For the current study, a blinded reviewer (ALP) evaluated the µCT images for (1) subchondral bone plate integrity and congruity, (2) osseous integration, and (3) presence of intraosseous cystic changes. (1) Subchondral bone plate was scored as intact within the graft and flush at host-graft junction (score 0) or disrupted within the graft or offset greater than one subchondral bone plate thickness (score 1). (2) Osseous integration at the host-graft junction was scored as normal if trabeculae crossed the circumferential interface (score 0) or abnormal if there was a region without crossing trabeculae (score 1). (3) Intraosseous cystic changes within the graft or at the host-graft junction, defined as trabeculae voids with a maximum diameter of greater than 2.0 mm, were scored as absence of cystic changes (score 0) or presence of cystic changes (score 1).

A blinded reviewer (ALP) also scored the hematoxylin and eosin (H&E) slides for degree of cartilage filling, similar to the corresponding MRI variable where cartilage “fill” of the graft (percentage volume) was scored as 76% to 100% (score 0), 51% to 75%, greater than 100% (score 1), and less than 50% (score 2). Cartilage edge integration at host-graft junction was scored as no discernible fissure (score 0), discernible fissure (score 1), and fissure >1 mm (score 2). The cartilage surface congruity of the graft and host was evaluated and designated as entirely flush (score 0), less than 50% offset compared with host cartilage thickness (score 1), or greater than 50% offset compared with the host cartilage thickness (score 2). Calcified cartilage integrity was designated as present and intact (score 0) or thinned/absent (score 1). Subchondral bone was designated as normal (score 0) or abnormal (score 1) if there were histologic features typically associated with bone marrow edema pattern lesions, including swollen fat cells surrounded by eosinophilic staining, fibrous tissue, or increased trabeculae.

**Statistical Analysis**

Statistical analyses were performed with R (v2.15, 2012, R Foundation for Statistical Computing, Vienna, Austria).
Unweighted $\kappa$ was used to assess interobserver agreement for individual MRI features. Features 1 to 13 were added together and designated as a 13-feature Total Score for MRI (TS13-MRI). Intraclass correlation coefficient (ICC) was used to assess interobserver agreement for TS13-MRI.

Unweighted $\kappa$ was used to assess agreement between individual MRI features and reference standards where available and the two kappa values (one for each reader) were averaged in order to maintain the same scale as the reference standard. For the remainder of analyses, MRI scores of the 2 readers were averaged and the average was used. Spearman’s correlation was used to examine the ordinal relationship for feature 1 MRI score (cartilage signal) versus MM score. Features 1 to 9 were added together and designated as a 9-feature Total Score for MRI (TS9-MRI) values and as a 9-feature Total Score for reference standard (TS9-REF) values based on histopathology and $\mu$CT. Spearman’s correlation was used to evaluate the relationship between TS9-MRI and TS9-REF.

Spearman’s correlation was also used to assess the relationship between TS13-MRI scores and indentation stiffness for MFC and LT grafts both as a group and separately. Kruskal-Wallis test was performed to detect significant differences between groups based on type of graft storage for TS9-MRI.

Table 2. Grading Scheme.

<table>
<thead>
<tr>
<th>MRI Feature</th>
<th>MRI Score</th>
<th>Reference Standard</th>
<th>Reference Standard Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Cartilage signal of graft</td>
<td>0: Normal 1: Altered intensity (either hypointense or hyperintense, but not fluid) 2: Fluid signal intensity on all sequences</td>
<td>Histopathology</td>
<td>Modified Mankin (0-15)</td>
</tr>
<tr>
<td>2. Cartilage “fill” of graft (percentage of volume)</td>
<td>0: 76% to 100% 1: 51% to 75% or &gt;100% 2: &lt;50%</td>
<td>Histopathology</td>
<td>0: 76% to 100% 1: 51% to 75% or &gt;100% 2: &lt;50%</td>
</tr>
<tr>
<td>3. Cartilage edge integration at host-graft junction</td>
<td>0: No discernible boundary 1: Discernible boundary 2: Discernible fissure &gt;1 mm</td>
<td>Histopathology</td>
<td>0: No discernible fissure 1: Discernible fissure 2: Discernible fissure &gt;1 mm</td>
</tr>
<tr>
<td>4. Cartilage surface congruity of graft and host-graft junction</td>
<td>0: Flush 1: &lt;50% offset of host cartilage 2: &gt;50% offset of host cartilage</td>
<td>Histopathology</td>
<td>0: Flush 1: &lt;50% offset 2: &gt;50% offset</td>
</tr>
<tr>
<td>5. Calcified cartilage integrity of graft</td>
<td>0: Intact, thin, and smooth 1: Altered (disrupted, thickened, or blurred)</td>
<td>Histopathology</td>
<td>0: Present and intact 1: Thinned or absent</td>
</tr>
<tr>
<td>6. Subchondral bone plate congruity of graft and host-graft junction</td>
<td>0: Intact and flush 1: Disrupted or not flush by &gt;1 subchondral thickness</td>
<td>Histopathology</td>
<td>0: Intact and flush 1: Disrupted or offset by &gt;1 subchondral thickness</td>
</tr>
<tr>
<td>7. Subchondral bone marrow signal intensity of graft relative to epiphyseal bone</td>
<td>0: Normal 1: Abnormal (bone marrow edema pattern or hypointensity on all sequences)</td>
<td>$\mu$CT</td>
<td>0: &lt;33% area with swollen fat cells, fibrous tissue, or increased trabeculae 1: &gt;33% abnormal area</td>
</tr>
<tr>
<td>8. Osseous integration at host-graft junction</td>
<td>0: Crossing trabeculae 1: Discernible cleft</td>
<td>$\mu$CT</td>
<td>0: Circumferentially crossing trabeculae 1: Region without crossing trabeculae</td>
</tr>
<tr>
<td>9. Presence of cystic changes of graft and host-graft junction</td>
<td>0: Absent 1: Present</td>
<td>$\mu$CT</td>
<td>0: Absent 1: Cyst &gt;1.0 mm present</td>
</tr>
<tr>
<td>10. Opposing cartilage</td>
<td>0: Normal 1: Abnormal</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>11. Meniscal tears</td>
<td>0: Absent 1: Present</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>12. Synovitis</td>
<td>0: Normal 1: Abnormal</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>13. Fat pad scarring</td>
<td>0: Normal 1: Abnormal</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table includes 9 primary features with reference standard validation (5 cartilage [features 1-5] and 4 bone [features 6-9]) and 4 ancillary features of the joint (features 10-13) without reference standard validation.
and TS9-REF scores. To adjust for within-knee dependence, bootstrap-based techniques were used to compute 95% confidence intervals (CI) for the ICCs and correlation coefficients, and to assess p-values as appropriate. Kappa statistics were interpreted as 0 to 0.2, slight; 0.21 to 0.4, fair; 0.41 to 0.6, moderate; 0.61 to 0.80, substantial; 0.81 to 1.0, almost perfect.42

Despite being underpowered for a subanalysis (and not likely to detect results), a purely exploratory subanalysis was performed to determine which individual features of OCAMRISS would best predict biomechanical indentation stiffness. Each individual OCAMRISS feature was used as a predictor in a mixed-effects regression with stiffness as outcome, covarying for type of storage and with a subject-specific intercept fitted. The averages of both readers were used for each OCAMRISS MRI feature.

**Results**

The OCAMRISS method was successfully applied to a data set designed to have variable treatment outcome, ranging from outstanding to poor (Figs. 1-4).

**Interobserver Agreement**

There was substantial to almost perfect agreement for all MRI score components (κ > 0.7, P < 0.001; Table 3). Although there was perfect interobserver agreement for feature 8 (osseous integration), κ could not be calculated because of lack of variability since both readers scored all 28 grafts the same (score 0, evidence for crossing trabeculae). Reader agreement was excellent for TS13-MRI (ICC = 0.982, CI = [0.965, 0.992]).

**MRI Scores Compared With Reference Standards**

For feature 1 (cartilage signal), there was only one disagreement between the 2 readers and there was a strong positive correlation of the MRI scores with MM score (ρ = 0.68, P < 0.0001). Agreement of MRI score for feature 3 (cartilage edge integration) compared with the reference standard was moderate (κ = 0.6). For all other features, there was substantial to almost perfect agreement between MRI and reference standard scores (κ > 0.6, P < 0.0001). Again, κ could not be calculated for feature 8, but there was perfect agreement. Figure 5 shows the relationship between TS9-MRI versus TS9-REF (ρ = 0.855, CI = [0.708, 0.928]).

Correlation between TS13-MRI with indentation stiffness was significantly negative (ρ = −0.528, CI = [−0.746, −0.149]) and this relationship strengthened when evaluating for only MFC grafts (ρ = −0.788, CI = [−0.948, −0.374]), in part due to a wider range of LT graft stiffness (Fig. 6). The Kruskal-Wallis test detected significant differences between storage subgroups of TS9-MRI and TS9-REF with the frozen group performing worse than the other 3 groups (P = 0.007 and P = 0.001, respectively; Fig. 7).

The purely exploratory subanalysis was performed on all OCAMRISS features except feature 8 where there was no variability. The mixed-effects regression showed that the only observed difference was in feature 5, with higher stiffness when the calcified cartilage was normal on UTE subtraction images (P = 0.0239). As this was an exploratory subanalysis, results from this mixed-effects regression require validation with independent data.
Discussion

In this study, we have described a comprehensive, semi-quantitative MRI scoring system for analysis of cartilage defect repair by osteochondral allografts, which we have termed OCAMRISS (OsteoChondral Allograft MRI Scoring System). The OCAMRISS method shows substantial interobserver agreement, substantial agreement with histopathologic and μCT reference standards for nearly all primary features, and significant negative correlation with biomechanical indentation testing. It is known from multiple prior reports that fresh and refrigerated osteochondral allografts perform better than frozen allografts, likely because of differences in cell viability, and OCAMRISS was sensitive to the differences in outcomes between these groups.

The OCAMRISS method was designed and validated on osteochondral allografts. However, where appropriate, we incorporated features from one of the most widely used MRI scoring systems after cartilage repair, the MOCART score. The MOCART score includes 9 variables: the degree of filling of the defect, integration to the border, calcified cartilage integrity, subchondral bone plate congruity, subchondral bone marrow signal, osseous integration, presence of cystic changes, and the degree of filling of the defect, integration to the border, and cartilage repair. The OCAMRISS method shows substantial agreement with the MOCART score for nearly all primary features, and significant negative correlation with biomechanical indentation testing. It is known from multiple prior reports that fresh and refrigerated osteochondral allografts perform better than frozen allografts, likely because of differences in cell viability, and OCAMRISS was sensitive to the differences in outcomes between these groups.

Figure 2. Medial femoral condyle allograft stored at 4 °C × 14 days with outstanding performance (OCAMRISS TS9-MRI 1 point and TS9-REF 7 points; cartilage stiffness = 5.1 MPa). Sagittal proton density (PD)–weighted image (A), sagittal 3D ultrashort echo time (UTE) subtraction image (B), hematoxylin and eosin stain (C), and micro–computed tomography (D) demonstrate features as listed in the accompanying table (E).

Figure 3. Medial femoral condyle allograft with frozen storage and poor performance (OCAMRISS TS9-MRI 12 points and TS9-REF 25 points; cartilage stiffness = 0.2 MPa). Sagittal short TI inversion recovery (STIR) image (A), sagittal 3D ultrashort echo time (UTE) subtraction image (B), hematoxylin and eosin stain (C), and micro–computed tomography (D) demonstrate features as listed in the accompanying table (E).
### Table 3. Results of Interobserver Agreement and Agreement With Reference Standards.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Reader 1 Scores, Score (Number of Samples, %)</th>
<th>Reader 2 Scores, Score (Number of Samples, %)</th>
<th>Interobserver Agreement</th>
<th>Reference Standard Scores, Score (Number of Samples, %)</th>
<th>Agreement With Reference Standard (Reader A)</th>
<th>Agreement With Reference Standard (Reader B)</th>
<th>Average κ for Readers A and B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Cartilage signal of graft</td>
<td>0 (14, 50%)</td>
<td>0 (14, 50%)</td>
<td>0.94</td>
<td>Mean modified Mankin 7.2 (SD 3.9, range 1-15)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>1 (10, 36%)</td>
<td>1 (9, 32%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 (4, 14%)</td>
<td>2 (5, 18%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Cartilage “fill” of graft (percentage of volume)</td>
<td>0 (17, 61%)</td>
<td>0 (18, 64%)</td>
<td>0.93</td>
<td>0 (18, 64%)</td>
<td>0.93</td>
<td>1</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>1 (4, 14%)</td>
<td>1 (3, 11%)</td>
<td></td>
<td>1 (3, 11%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 (7, 25%)</td>
<td>2 (7, 25%)</td>
<td></td>
<td>2 (7, 25%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Cartilage edge integration at host-graft junction</td>
<td>0 (4, 14%)</td>
<td>0 (3, 11%)</td>
<td>0.71</td>
<td>0 (3, 11%)</td>
<td>0.58</td>
<td>0.62</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>1 (20, 71%)</td>
<td>1 (18, 64%)</td>
<td></td>
<td>1 (21, 75%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 (4, 14%)</td>
<td>2 (7, 25%)</td>
<td></td>
<td>2 (4, 14%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Cartilage surface congruity of graft and host-graft junction</td>
<td>0 (12.43%)</td>
<td>0 (12.43%)</td>
<td>0.94</td>
<td>0 (11.39%)</td>
<td>0.88</td>
<td>0.94</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>1 (12.43%)</td>
<td>1 (13.46%)</td>
<td></td>
<td>1 (14.50%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 (4.14%)</td>
<td>2 (3.11%)</td>
<td></td>
<td>2 (3.11%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Calcified cartilage integrity of graft</td>
<td>0 (11.39%)</td>
<td>0 (13.46%)</td>
<td>0.86</td>
<td>0 (13.46%)</td>
<td>0.71</td>
<td>0.86</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>1 (17.61%)</td>
<td>1 (15.54%)</td>
<td></td>
<td>1 (15.54%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Subchondral bone plate congruity of graft and host-graft junction</td>
<td>0 (13.46%)</td>
<td>0 (14.50%)</td>
<td>0.93</td>
<td>0 (13.46%)</td>
<td>1.00</td>
<td>0.93</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>1 (15.54%)</td>
<td>1 (14.50%)</td>
<td></td>
<td>1 (15.54%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Subchondral bone marrow signal intensity of graft relative to</td>
<td>0 (15.54%)</td>
<td>0 (14.50%)</td>
<td>0.93</td>
<td>0 (18.64%)</td>
<td>0.64</td>
<td>0.57</td>
<td>0.61</td>
</tr>
<tr>
<td>epiphyseal bone</td>
<td>1 (13.46%)</td>
<td>1 (14.50%)</td>
<td></td>
<td>1 (10.36%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Osseous integration at host-graft junction</td>
<td>0 (28.100%)</td>
<td>0 (28.100%)</td>
<td>Perfect agreementa</td>
<td>0 (28.100%)</td>
<td>Perfect agreementa</td>
<td>Perfect agreementa</td>
<td>Perfect agreementa</td>
</tr>
<tr>
<td></td>
<td>1 (0.0%)</td>
<td>1 (0.0%)</td>
<td></td>
<td>1 (0.0%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Presence of cystic changes of graft and host-graft junction</td>
<td>0 (7.25%)</td>
<td>0 (9.32%)</td>
<td>0.83</td>
<td>0 (7.25%)</td>
<td>1.00</td>
<td>0.83</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>1 (21.75%)</td>
<td>1 (19.68%)</td>
<td></td>
<td>1 (21.75%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Opposing cartilage</td>
<td>0 (20.71%)</td>
<td>0 (20.71%)</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>1 (8.29%)</td>
<td>1 (8.29%)</td>
<td></td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Meniscal tears</td>
<td>0 (20.71%)</td>
<td>0 (20.71%)</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>1 (8.29%)</td>
<td>1 (8.29%)</td>
<td></td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Synovitis</td>
<td>0 (26.93%)</td>
<td>0 (26.93%)</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>1 (2.7%)</td>
<td>1 (2.7%)</td>
<td></td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Fat pad scarring</td>
<td>0 (26.93%)</td>
<td>0 (26.93%)</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>1 (2.7%)</td>
<td>1 (2.7%)</td>
<td></td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*aSpearman’s correlation was used instead (refer to text).  
*bKappa could not be calculated because of lack of variability (refer to text).  
*cReference standards for features 10 to 13 were not available.
unable to regenerate across a physical gap, unlike defect for a snug fit, cartilage of osteochondral grafts is additionally, despite oversizing the graft relative to the cartilage signal intensity) and 2 (cartilage “fill”) in our system. repair tissue) may be better suited for cell-based repair techniques. For this reason, none of our samples demonstrated complete chondral integration on histology. Therefore, delineation of “integration to the border zone” (MOCART score feature 2 and corresponding feature 3 of our grading scheme) should be made with this in mind since the best possible MRI outcome (score 0, no discernible boundary) will show a boundary at histology. We also did not include MOCART score feature 8 (presence of adhesions) into our classification since it may be more important in cell-based repair techniques, where it can lead to the requirement for arthroscopic debridement in up to 5% to 10% of patients. Finally, evaluation of the calcified cartilage layer (feature 5) is unique to our system and made possible with the advent of UTE sequences. The calcified cartilage layer is metabolically active, remodels with changing loads, and plays an important role in overall function of the osteochondral unit.

In addition to the widely popular MOCART score, other authors have used various MRI scoring systems to characterize osteochondral grafts, but to our knowledge, no formal one-to-one comparison between scoring system components and corresponding reference standards has been made. Rather, prior studies have focused on the comparison between MRI and clinical outcomes, histologic cellular viability, histologic composition, or immunologic responses. Furthermore, both the OCAMRISS and the MOCART systems are more comprehensive in comparison with these prior studies.

Correlation between OCAMRISS and indentation stiffness was significantly negative, however the relationship strengthened when evaluating for only MFC grafts. This may have been because of several reasons. As we have previously noted, our sample demonstrated greater histopathologic and biomechanical variability at the LT compared with the MFC sites. This is consistent with several clinical studies which have shown worse outcomes at patellofemoral lesions. In addition to this, both MRI readers subjectively felt the MFC grafts were easier to score, in part due to increased volume averaging at the LT due to curvature of cartilage and because LT grafts were on average ~40% thinner (LT 0.72 ± 0.21 mm versus MFC 1.78 ± 0.54 mm). For this reason, typically two planes were required for LT graft evaluation whereas a single plane was sufficient for MFC grafts. Despite these differences, a post hoc analysis determined that there was no detectable adverse effect of cartilage thickness on interobserver agreement or correlations used in this study. Additionally, our purely exploratory subanalysis suggests that calcified cartilage integrity (feature 5) as seen on UTE subtraction images may be particularly useful to predict biomechanical outcome. We are hopeful that OCAMRISS may serve as a springboard for future studies to independently validate this finding.

Our study has a number of limitations. First, our sample size was small. However, our main goal was to correlate...
OCAMRISS with reference standard scores. Confidence intervals were reported to demonstrate the sample size-based estimate of the uncertainty around the correlation coefficient, its “projection onto reality”. Regarding the correlation of TS9-MRI versus TS9-REF (Figure 5), the interpretation is that we are 95% confident that the true value is covered by the interval (the worst case of which is 0.708). Additionally, we attempted to simulate a spectrum of appearances by analyzing equal numbers of fresh, refrigerated (both 14 and 28 days), and frozen grafts. Despite this, we note that there was no variability in feature 8 (osseous integration), likely due to small sample size. We maintain that this is an important feature as we have noticed, as well as prior authors, that graft osseous nonunion can occur.\(^{13,67}\) Second, caution must be exercised when applying these results to humans. Although we designed the scoring

Figure 6. Graphs of 13-feature MRI score (TS13-MRI) versus biomechanical indentation stiffness for both medial femoral condyle (MFC) and lateral trochlea (LT) grafts (A) and only for MFC grafts (B). Spearman’s ρ for combined MFC and LT grafts was significantly negative (ρ = −0.528, confidence interval [CI] = [−0.746, −0.149]) and the relationship strengthened when evaluating for only MFC grafts (ρ = −0.788, CI = [−0.948, −0.374]) as there was a wider range of stiffness for LT grafts. Red dots represent MFC grafts and blue dots represent LT grafts.

Figure 7. Boxplots of mean 9-feature MRI score (TS9-MRI) (A) and 9-feature reference standard score (TS9-REF) (B). The Kruskal-Wallis test detected significant differences for TS9-MRI and TS9-REF with the frozen group performing worse than the other three groups (p = 0.007 and p = 0.001, respectively).
system based on our experience with human osteochondral allografts, due to the ethical challenges that would prohibit validation in humans, we chose to validate with an animal model. Of note, the goat is a large animal model that demonstrates a similar limited intrinsic healing potential of cartilage and the 8-mm diameter grafts used in this study are larger than the critical size defect of 6 mm in goats.

Third, our semiquantitative classification system remains based on morphological MRI sequences. While some authors have found that conventional, qualitative MRI cannot predict histologic appearance (such as hyaline cartilage, fibrocartilage, or fibrous tissue), additional studies should be performed evaluating quantitative MR biomarkers and their possible inclusion into classification systems. Fourth, although the UTE sequence was only used for evaluation of Feature 5 of OCAMRISS (integrity of the deep cartilage of the graft), there is potential for the UTE sequence to influence grading of other features. Fifth, higher imaging resolution was used in this study compared with a clinical human knee protocol in order to compensate for the smaller tissues in goats. Unfortunately, this led to the requirement of a lower receiver bandwidth in order to achieve adequate signal-to-noise ratio. Chemical shift was approximately 2.7 pixels on the proton density–weighted images. Although we did not feel that this adversely affected our interpretation, the degree of chemical shift is higher than that typically seen in clinical practice. Finally, we did not include clinical outcome measures and prior studies have suggested that certain MRI variables, such as trabecular incorporation, correlate with outcome measures.

In conclusion, we described a comprehensive, semiquantitative MRI scoring system for the assessment of cartilage repair by osteochondral allografts, which we have termed OCAMRISS. We have demonstrated substantial interobserver agreement, substantial agreement with carefully selected reference standards for each component, and significant negative correlation with biomechanical indentation testing. The defined variables allow for an accurate description and may help standardize reporting of MR imaging findings after repair of cartilage defects with osteochondral allografts.

Acknowledgments and Funding
The authors thank Karen Bowden for technical histology assistance. This work was supported by the National Institutes of Health (grant numbers R01AR055637-02S1, R01DE022068). Eric Y. Chang receives salary support from a VA CSR&D Career Development Award (grant number 1IK2CX000749).

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical Approval
This study was approved by our institutional review board. All experiments were carried out in accordance with protocols approved by the Institutional Animal Care and Use Committee.

References

Table 4. Comparison Between MOCART Score and Osteochondral Allograft Score.

<table>
<thead>
<tr>
<th>MOCART Score</th>
<th>OCAMRISS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature 1: Degree of defect repair and filling of the defect</td>
<td>Feature 2: Cartilage “fill” of graft</td>
</tr>
<tr>
<td>Feature 2: Integration to border zone</td>
<td>Feature 3: Cartilage edge integration at host-graft junction</td>
</tr>
<tr>
<td>Feature 3: Surface of repair tissue</td>
<td>Feature 4: Cartilage surface congruity</td>
</tr>
<tr>
<td>Feature 4: Structure of repair tissue</td>
<td>Combination of feature 1 (cartilage signal of graft) and feature 2 (cartilage “fill” of graft)</td>
</tr>
<tr>
<td>Feature 5: Signal intensity of repair tissue</td>
<td>Feature 1: Cartilage signal of graft</td>
</tr>
<tr>
<td>Feature 6: Subchondral lamina</td>
<td>Feature 6: Subchondral bone plate congruity of graft and host-graft junction</td>
</tr>
<tr>
<td>Feature 7: Subchondral bone</td>
<td>Combination of feature 7 (subchondral bone marrow signal intensity of graft), feature 8 (osseous integration at host-graft junction), and feature 9 (presence of cystic changes of graft and host-graft junction)</td>
</tr>
<tr>
<td>Feature 8: Presence of adhesions</td>
<td>None</td>
</tr>
<tr>
<td>Feature 9: Presence of synovitis</td>
<td>Feature 12</td>
</tr>
</tbody>
</table>

MOCART, Magnetic Resonance Observation of Cartilage Repair Tissue; OCAMRISS, Osteochondral Allograft MRI Scoring System.


42. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159-74.