Methodologies for Semiquantitative Evaluation of Hip Osteoarthritis by Magnetic Resonance Imaging: Approaches Based on the Whole Organ and Focused on Active Lesions

Jacob L. Jaremko, Robert G.W. Lambert, Veronika Zubler, Ulrich Weber, Damien Loeuille, Frank W. Roemer, Jolanda Cibere, Marcus Pianta, David Gracey, Philip Conaghan, Mikkel Ostergaard, and Walter P. Maksymowych

ABSTRACT. ***Objective.*** As a wider variety of therapeutic options for osteoarthritis (OA) becomes available, there is an increasing need to objectively evaluate disease severity on magnetic resonance imaging (MRI). This is more technically challenging at the hip than at the knee, and as a result, few systematic scoring systems exist. The OMERACT (Outcome Measures in Rheumatology) filter of truth, discrimination, and feasibility can be used to validate image-based scoring systems. Our objective was (1) to review the imaging features relevant to the assessment of severity and progression of hip OA; and (2) to review currently used methods to grade these features in existing hip OA scoring systems.

Methods. A systematic literature review was conducted. MEDLINE keyword search was performed for features of arthropathy (such as hip + bone marrow edema or lesion, synovitis, cyst, effusion, cartilage, etc.) and scoring system (hip + OA + MRI + score or grade), with a secondary manual search for additional references in the retrieved publications.

Results. Findings relevant to the severity of hip OA include imaging markers associated with inflammation (bone marrow lesion, synovitis, effusion), structural damage (cartilage loss, osteophytes, subchondral cysts, labral tears), and predisposing geometric factors (hip dysplasia, femoral-acetabular impingement). Two approaches to the semiquantitative assessment of hip OA are represented by Hip OA MRI Scoring System (HOAMS), a comprehensive whole organ assessment of nearly all findings, and the Hip Inflammation MRI Scoring System (HIMRISS), which selectively scores only active lesions (bone marrow lesion, synovitis/effusion). Validation is presently confined to limited assessment of reliability.

Conclusion. Two methods for semiquantitative assessment of hip OA on MRI have been described and validation according to the OMERACT Filter is limited to evaluation of reliability. (First Release Nov 15 2013; J Rheumatol 2014;41:359–69; doi:10.3899/jrheum.131082)

Key Indexing Terms:
OSTEOARTHRITIS HIP MAGNETIC RESONANCE IMAGING
OMERACT SCORING SYSTEM

From the Department of Radiology and Diagnostic Imaging, University of Alberta, Edmonton, Alberta, Canada; Department of Radiology and Department of Rheumatology, Balgrist University Hospital, Zurich, Switzerland; Department of Medicine, CHU de NANCY-Brabois, Vandoeuvre, France; Department of Radiology, Klinikum Augsburg, Augsburg, Germany; Quantitative Imaging Center (QIC), Department of Radiology, Boston University School of Medicine, Boston, Massachusetts, USA; Department of Medicine, University of British Columbia and Research Scientist, Arthritis Research Centre of Canada, Vancouver, British Columbia, Canada; St. Vincent's Hospital, Victoria, Australia; Radiology Department, Craigavon Area Hospital, Southern Health and Social Care Trust, Portadown, Northern Ireland; National Institute for Health Research (NIHR) Leeds Musculoskeletal Biomedical Research Unit, Chapel Allerton Hospital, Leeds, UK; Department of Rheumatology, Copenhagen University Hospital at Glostrup, Copenhagen, Denmark; Department of Medicine, University of Alberta, Edmonton, Alberta, Canada.

J.L. Jaremko, MD, PhD, Assistant Professor; R.G.W. Lambert, MB, FRCPC, Professor, Department of Radiology and Diagnostic Imaging, University of Alberta; V. Zubler, MD, Assistant Professor, Department of Radiology, Balgrist University Hospital; U. Weber, MD, Assistant Professor, Department of Rheumatology, Balgrist University Hospital; D. Loeuille, MD, Professor of Rheumatology, Department of Medicine, CHU de NANCY-Brabois; F.W. Roemer, MD, Associate Professor of Radiology, Department of Radiology, Klinikum Augsburg and QIC, Department of Radiology, Boston University School of Medicine; J. Cibere, MD, PhD, Associate Professor, Department of Medicine, University of British Columbia and Research Scientist, Arthritis Research Centre of Canada; M. Pianta, MBBS, MMed, FRANZCR, Consultant Radiologist, St. Vincent's Hospital; D. Gracey, MB, ChB, MRCPCH, FRCR, Consultant Radiologist (Musculoskeletal Lead) Radiology Department, Craigavon Area Hospital, Southern Health and Social Care Trust; P. Conaghan, MD, FRACP, Professor of Musculoskeletal Medicine, NIHR Leeds Musculoskeletal Biomedical Research Unit, Chapel Allerton Hospital; M. Ostergaard, MD, PhD, Professor of Rheumatology, Department of Rheumatology, Copenhagen University Hospital at Glostrup; W.P. Maksymowych, MB, FRCPC, Professor, Department of Medicine, University of Alberta.

Address correspondence to Dr. Jaremko, Department of Radiology and Diagnostic Imaging, Faculty of Medicine, University of Alberta, 2A2.41 WMC, 8440 112 Avenue NW, Edmonton, Alberta T6G 2B7, Canada. E-mail: jjaremko@ualberta.ca.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2014. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org

Given recent developments in the understanding of the pathophysiology of osteoarthritis (OA), and the widening array of oral, intraarticular, and surgical treatments for this condition, there is increasing interest in evaluating the effectiveness of therapeutic agents. Validated tools are required for assessment of disease severity and activity. Conventional approaches rely on patient self-reported measures, which lack objectivity, and radiographic assessment, which lacks responsiveness, particularly in regard to the proportion of patients who progress in the time frame of clinical trials. The Outcome Measures in Rheumatology (OMERACT) filter and methodological framework have been developed to validate outcome instruments that include imaging modalities aimed at assessment of disease severity according to 3 criteria: truth (i.e., face, content, construct, and criterion validity), discrimination (reliability and sensitivity to change), and feasibility[1]. At the OMERACT 11 conference May 13-18, 2012, in North Carolina, the Hip Magnetic Resonance Imaging (MRI) Osteoarthritis Working Group assessed 2 different hip OA scoring systems, the Hip OA MRI Scoring System (HOAMS) and the Hip Inflammation MRI Scoring System (HIMRISS) according to the OMERACT filter. The first major activity of the group was to conduct a systematic literature search to (1) define the features of hip OA on MRI considered relevant to disease severity; (2) determine what is known regarding associations with clinical characteristics of disease severity; (3) identify and describe existing methods for quantifying disease severity on MRI; and (4) assess the extent of their validation according to the OMERACT filter.

METHODS

A 2-part literature review was performed by the first author (JJ), an MD-PhD radiologist in practice for 5 years, and results reviewed with the other co-authors in consensus.

Search (A) sought the set of features potentially visible on imaging known to reflect OA, and search (B) sought hip OA scoring systems at MRI. Search (A), on PubMed MEDLINE (1948 to March 2013), included initial query of MESH headings "Osteoarthritis, Hip", subheadings "Epidemiology" and "Etiology" (991 articles). We limited this to review articles (149) and then to core clinical journals (27 articles). Each feature detectable on imaging described in abstracts and selected full text of these major reviews was then assessed by specific search. For example, bone marrow lesion (BML) was queried in Medline by keywords (bone marrow AND (edema OR lesion) AND hip AND osteoarthritis), giving 19 articles, 6 of which were relevant on review of abstracts. From within these, secondary manual search of citations was also then performed to assess other aspects of each imaging feature.

Search (B), seeking papers describing OA scoring systems applied to hip MRI, is summarized in Figure 1. Initial search on PubMed MEDLINE (1948 to March 2013) gave 36 results, 10 relevant, of which 5 contained unique MRI scoring systems for OA and 2 were review articles. Use of the same search strategy in several other biomedical databases added 70 results, of which 9 were relevant, all conference abstracts applying the scoring systems already retrieved at Medline search. Manual search of the bibliographies of the 10 retrieved full-text papers also found no additional comprehensive scoring systems. Critical appraisal of the selected articles considered (1) whether the scoring system was described in adequate detail to be reproduced by us, and (2) to what extent the components of the OMERACT filter were addressed.

RESULTS AND DISCUSSION

Radiographic Grading

The original radiographic scoring system for OA, the Kellgren-Lawrence scale (Table 1), is simple and remains frequently used to stratify patients. In terms of the OMERACT filter, the scale may lack validity given the implicit assumption that features such as osteophyte formation and articular narrowing progress together[2,3]; and it certainly lacks discrimination for short-term interval change, making it insufficiently sensitive for use in evaluating therapeutic agents.

MRI Grading Systems Available for Hip OA

In 2003, Schmid, *et al* developed a systematic grading system for hip cartilage[4]. The earliest comprehensive hip OA scoring system we identified, from Neumann, *et al*[5], focused on cartilage, bone marrow signal changes, labral tears, osteophytes, and subchondral cysts. More recently, HOAMS was developed by consensus among hip orthopedic surgeons, rheumatologists, and radiologists[6]. HOAMS assesses the same features of arthropathy graded by Neumann, *et al* and adds scoring of synovitis and other articular and periarticular findings. Of note, the version of HOAMS described in this report has been slightly modified and updated since the prior publication. HIMRISS has recently been developed using data from a trial of intraarticular steroid therapy for OA[7,8], based on the concept that scoring of hip OA most usefully emphasizes evidence of active inflammation and omits measurement of existing and untreatable structural damage. Unlike the other systems, HIMRISS measures only 3 features (bone marrow lesion, synovitis, and effusion), but does so across more and smaller subregions of the joint.

HOAMS uses proton density (PD) base sequences, which are suitable for detecting abnormal water content on a background of homogeneous marrow fat (such as seen in adult knees), but abnormal water signal is much harder to distinguish from erythropoietic components commonly seen in marrow at the hip on PD sequences. This gives potential for confusion between normal red marrow and edema. HIMRISS uses a short-tau inversion recovery (STIR) base sequence, which simplifies the distinction between red marrow and edema.

Most recently, in 2012 Stelzeneder, *et al*[9] assessed cartilage, labrum, osteophytes, paralabral cysts, and bone cysts, on radial reformats of a 3D true-FISP MRI sequence while investigating relations between acetabular morphology and OA. This system scores image features on 7 zones, but the locations of these zones are based on specialized reformats from 3D sequences not frequently obtained clinically.

The 3 most comprehensive and generally applicable scoring systems are compared in Table 2. The remainder of

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2014. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org

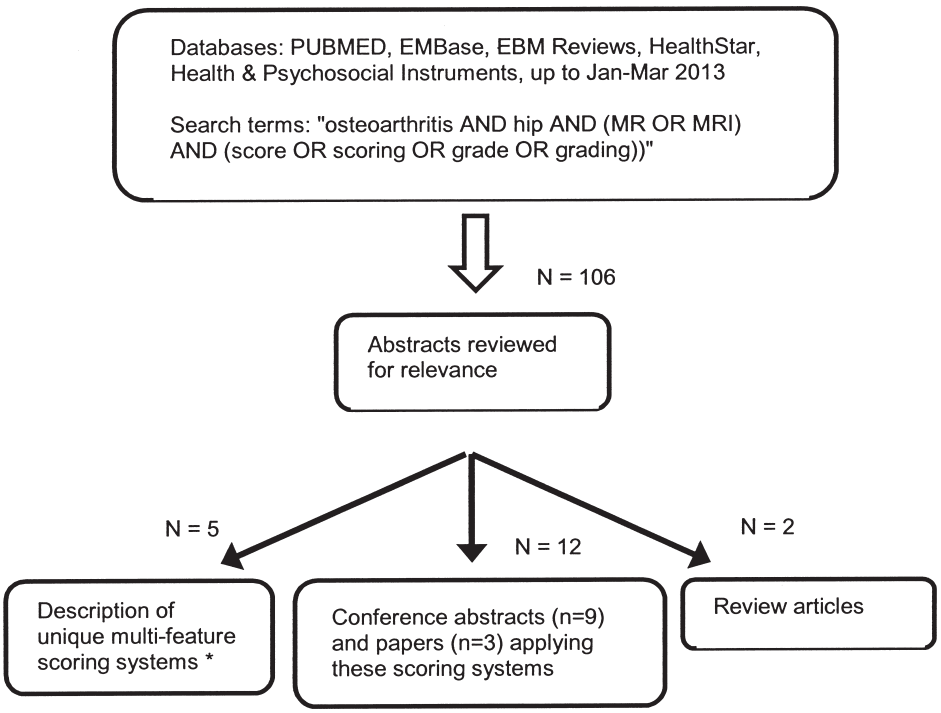


Figure 1. Search strategy for magnetic resonance imaging (MRI) hip osteoarthritis (OA) scoring systems. The 5 scoring systems are reported[4,5,6,8,9].

Table 1. Kellgren-Lawrence osteoarthritis scale (for radiographs)[3].

0	Normal
1	Doubtful joint space narrowing, possible osteophytes
2	Possible joint space narrowing, definite osteophytes
3	Definite joint space narrowing, multiple osteophytes, some sclerosis, possible deformity of bone ends
4	Marked joint space narrowing, large osteophytes, severe sclerosis, definite deformity of bone ends

this article will compare these scoring systems for each of the major features of arthropathy in terms of their validation according to the OMERACT filter.

Cartilage

Rationale for feature inclusion in OMERACT filter. A traditional measure of OA progression is extent of damage to articular cartilage, which eventually translates into joint space loss on radiographs. This has face validity in that normal joints have thick smooth cartilage plates while severely arthritic joints show extensive cartilage loss, while presence of focal cartilage lesions and/or change in cartilage signal has at least some correlation to patient pain[10]. Because of normal variation with age, sex, and body habitus, quantitative cartilage measurement is more valuable in longitudinal followup than cross-sectional study[2]. The main limitation of cartilage volume analysis is poor discrimination: cartilage loss occurs slowly, from 0–7% per year, giving limited sensitivity to change in time frames less than several years[1,11,12]. Focal cartilage defects may be more symptomatic than diffuse cartilage loss, and these are often scored as in arthroscopy, by a 4- to 8-point semiquantitative scale[13,14,15,16], or quantified by mapping MRI signal features such as the rate of T2 or T2* tissue relaxation after radiofrequency pulse, or accumulation of gadolinium contrast[2]. Feasibility, the third component of the OMERACT filter, is a concern at the hip given that cartilage imaging is much more difficult and error prone at the hip than the knee because of thinner and more tightly apposed cartilage plates.

Feature scoring. Schmid, *et al*[4], using MR arthrograms, graded cartilage degeneration as present or absent in 5 hip regions (4 acetabular and one representing the entire femoral head; Figure 2a), with a field for readers to explain in free text why they felt a given area had cartilage degeneration or not. Not unexpectedly, this subjective approach without predefined criteria had relatively poor interobserver reliability, as evidenced by the wide variation reported between 2 observers in sensitivity for cartilage defects observed on arthroscopy (SN = 79% for 1 reader, 50% for the other), and poor to fair agreement by kappa statistic (0.2 to 0.31). Use of this system would have limited discrimination for progressive change, since it scores only presence or absence of degeneration in relatively wide swathes of cartilage. Scoring by this approach is unlikely to meet the criteria of the OMERACT filter.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2014. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org

Table 2. Magnetic resonance imaging (MRI) based scoring systems for hip osteoarthritis.

	Neumann, et al[5]	HOAMS[6]	HIMRISS[8]
Concept	Overall grading of hip degeneration, as statistical combination of maximum changes in each of 5 zones	Score of 14 features of osteoarthritis in the hip, developed by multidisciplinary consensus committee	Sum of signs of inflammation (edema, synovitis, effusion) in portions of femoral head, acetabulum and adjacent tissue, across each slice of MRI
Base sequence	MR arthrogram	Proton density fat saturated (PD FS), small field of view (1 hip), axial and sagittal	Coronal inversion recovery (STIR), can be large field of view (full pelvis) or 1 hip
Zones: femoral head	1	5 sectors in central hip: central-central, central-superior, central-inferior, central-medial, central-lateral	Central 5 slices: 8 equal sectors of arc + 1 central zone, per slice
		4 zones in anterior and posterior 9 mm of femoral head: anterosuperior, anteroinferior, posterosuperior, posteroinferior	2 zones (superior and inferior) per slice up to 5 anterior slices and up to 5 posterior slices
Zones: acetabulum	4	3 sectors in central hip (central-central, central-superior, central-inferior); 2 sectors in anterior and posterior 9 mm (superior and inferior)	3 sectors per slice for central 5 slices; 2 sectors per slice for up to 5 anterior slices, 5 posterior
Zones: total	5 total	9 femoral, 5 acetabular, total 14	65 femoral, 35 acetabular, total 100
Planes to use		Coronal for central slices, sagittal for ant/post	Coronal only
Cartilage	0 normal	0 normal	Not scored
	1 = signal change only	1 focal partial thickness < 25% of region area	
	2 fissuring only, < 1 mm width	2 focal full thickness < 25% of region area	
	3 defect < 50% of thickness	3 several partial or 1 full > 25%	
	4 defect > 50% of thickness	4 several full thickness defects or one large, > 25%	
	5 exposed bone		
	Width of lesion		
	A = < 1 cm, B = 1–2 cm, C = > 2 cm per zone		
BML	0 = none	0 = none	0 = normal
	1 = < 0.5 cm from bone	1 = < 33% of region	1 = higher signal than reference signal (uninvolved or contralateral marrow)
	2 = < 1.5 cm from bone	2 = < 66% of region	
	3 = > 1.5 cm from bone	3 = > 66% of region	
Subchondral cysts	Grade whole joint: 0 none	0 = none	Not scored
	1 = 1–2 cysts	1 = < 33% of region	
	2 = 3–4 cysts	2 = < 66% of region	
	3 = 5+ cysts	3 = > 66% of region	
Osteophytes			
Locations	5 zones (ant/post/med/lat + central)	Femoral in each location	Not scored
Scoring	Per zone: 0 = absent; 1 = mild (< 2 mm); 2 = mod (< 5 mm); 3 = severe > 5 mm	5 locations: (ant/post on axial slices; superomedial/inferolateral/central intra-articular on coronal slices)	
	Whole-joint score, from sum of partial scores: 0 if 0; 1 if 1–2; 2 if 3–4; 3 if > 5+	0 absent; 1 = equivocal; 2 = small definite beak-like; 3 = intermediate; 4 = large proliferative	
Labrum			
Locations	Anterior/posterior/medial/lateral	Anterior on sagittal; superolateral/posteromedial on coronal; anterior/posterior on axial	Not scored
Scoring	0 = normal	0 = normal	
	1 = signal change or small labrum	1 = signal change or small labrum	
	2 = simple tear	2 = simple tear	
	3 = complex tear	3 = maceration	
Synovitis			
Locations		4 sites: ant/post/med/lat femoral head/neck junction	Scored once per slice × 15 slices
Scoring	NA	Synovium only; thickness 0 if < 2 mm; 1 if 2–4 mm; 2 if > 4 mm	Combined synovium + effusion: thickness 0 < 2 mm; 1 if 2–4 mm; 2 if > 4 mm
Effusion	NA	Single grade 0, 1, 2 (overall joint assessment)	Combined with synovitis

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2014. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org

Table 2. Continued.

	Neumann, et al[5]	HOAMS[6]	HIMRISS[8]
Other			
Loose bodies	NA	Present or absent	NA
Attrition (flattened articular surface)		1 if definite flattening/asphericity of femoral head	
Dysplasia (under-coverage of acetabular roof)		Measure centre-edge angle at mid coronal slice, if < 25*, score as dysplastic	
Bursitis/enthesopathy		Trochanteric bursitis, greater & lesser trochanteric tendinopathy each either present or absent	
Paralabral cysts		Present or absent	
Superolateral femoral head cyst		Present or absent ("herniation pit")	
Maximum score	Statistical combination scheme using max. individual score at each site	45 (cartilage) + 45 (BML) + 45 (cysts) + 40 (osteophytes) + 15 (labrum) + 8 (synovitis) + 10 (other) = 208	57 (femoral head BML) + 27 (acetabular BML) + 30 (synovitis/effusion) = 114

HOAMS: Hip OsteoArthritis Scoring System; HIMRISS: Hip Inflammation MRI Scoring System; BML: bone marrow lesion; NA: not applicable.

Neumann, *et al*[5] also used MR arthrograms, but divided the femoral head and acetabular cartilage into 5 zones each (total 10 zones; Figure 2b). The published report does not give specific instructions as to how to subdivide the regions. They scored cartilage defects on a 6-point scale based on signal change and quantity of cartilage loss (Table 2). Prevalence of cartilage defects in young patients with mechanical hip pain was high, at 76%. For the small subgroup of patients who also had arthroscopy (n = 23), this MRI scoring system had SN = 81–89%, specificity (SP) = 66% for cartilage defects, suggesting a tendency to overcall defects on MRI and/or missed small cartilage lesions on arthroscopy. Interreader reliability was improved compared to Schmid, *et al*, with kappa values in anterior femoral head and medial acetabulum both excellent at 0.83. However, agreement was lower in other areas, and the medial acetabulum is not a frequent site of focal cartilage loss. Since all imaged patients had mechanical hip pain, correlation to clinical findings was unclear. This scoring system might better meet OMERACT filter criteria for cartilage damage but has not been fully tested.

Hip cartilage regions are even further subdivided in HOAMS[6], which uses PD fat saturated (PD FS) and MEDIC (Multi-echo Data Image Combination) base sequence without arthrogram. Unlike the earlier arthrographic systems, the thin femoral and acetabular cartilage plates are scored together in each region. To deal with volume averaging, the most anterior and posterior 9 mm of femoral head are scored on sagittal images (Figure 2c; 4 regions), and the remaining central portions are scored on coronal slices (Figure 2d; 5 regions). In each region, scoring is on a 5-point scale based on number and depth of defects seen (Table 2). Reliability of the HOAMS cartilage scoring system was tested on 15 hips, with substantial, but not excellent, maximum interobserver agreement (kappa = 0.65). This score had significant correlation to the Kellgren-Lawrence grade of OA on radiographs, but did not correlate to patient pain scores. No arthroscopic correlation has been performed. Thus, both the reliability and truth component of the HOAMS cartilage scoring system have had only limited assessment to date.

HIMRISS does not score cartilage morphology or signal. Cartilage is also not well assessed on whole-pelvis coronal STIR images used in HIMRISS, owing to limited spatial resolution.

Osteophytes

Rationale for feature inclusion in OMERACT filter. Osteophytes represent bony remodeling and are a traditional sign of OA on radiographs (Table 1)[3]. They are measurable and show progressive change, and at least in knee OA, growth of osteophytes had a higher standardized response mean than either changes in cartilage volume or marrow edema[11]. In a systematic review of hip OA studies, presence of femoral head osteophytes was one of the few strong predictors of OA progression[17]. Also, observed changes in osteophytes are slow (e.g., standardized response mean of 0.3 SD in 6 months at the knee[11]). Osteophytes at the hip are often smaller and more sessile than at the knee, further increasing difficulty in accurate measurement and reducing potential for discrimination.

Feature scoring. Neumann, *et al*[5] scored osteophytes using each of the same 5 regions as for cartilage or BML, assigning a score in each zone based on osteophyte maximum size and a whole-joint score based on the sum of regional scores (Table 2). HOAMS uses 5 locations (4 along femoral head-neck junction, and central intraarticular adjacent to the fovea), scored on axial and coronal

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2014. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org

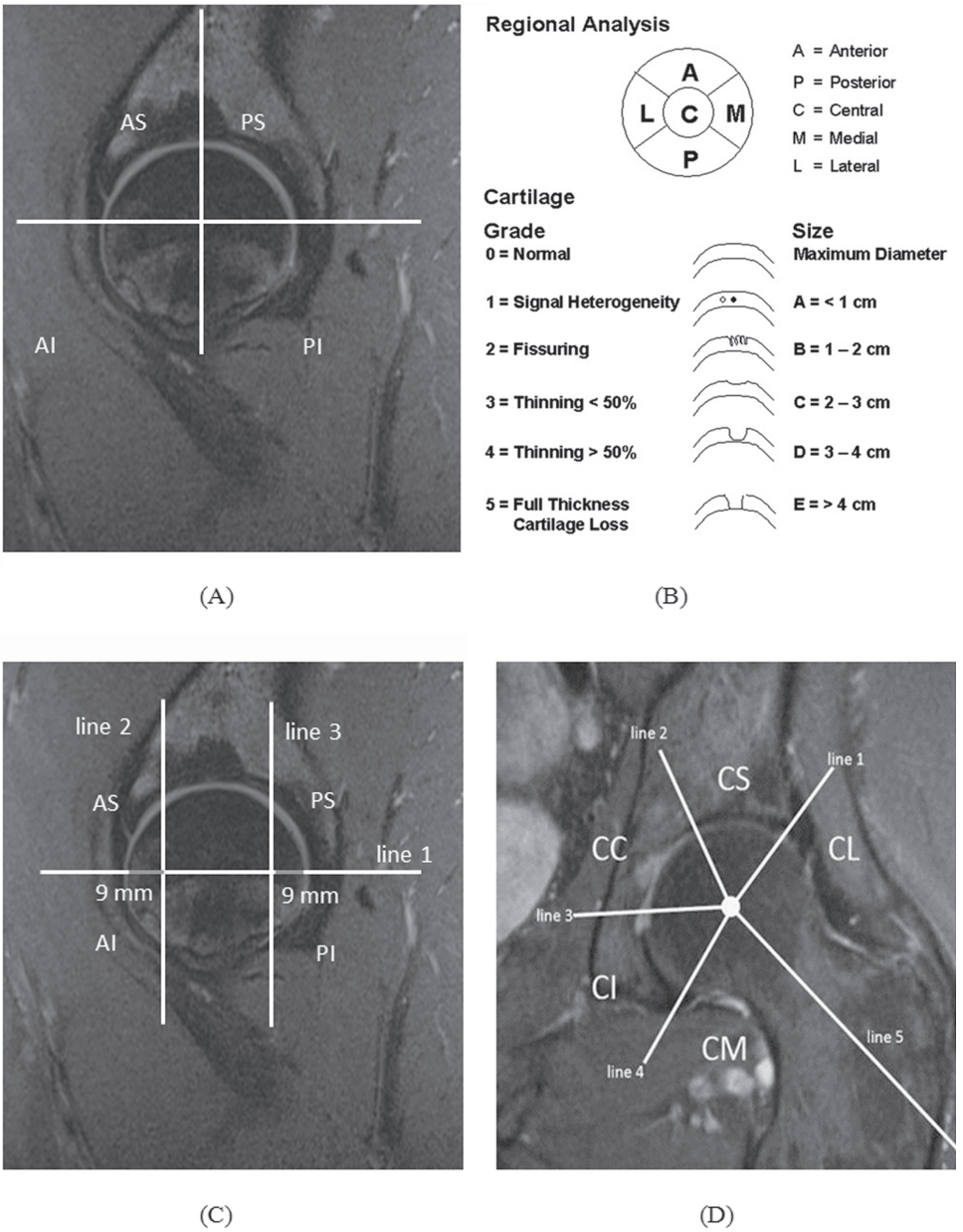


Figure 2. Assessment of cartilage lesions on magnetic resonance imaging (MRI). A. Method according to Schmid, *et al*[4]. Sagittal proton density (PD) fat saturated (FS) MRI. Acetabular zones are labeled anterosuperior (AS), anteroinferior (AI), posterosuperior (PS), posteroinferior (PI). Note that the PI zone is present only medially. B. Method according to Neumann, *et al*[5]. C, D. Method according to Roemer, *et al*[6], Hip OsteoArthritis MRI Scoring System. C: sagittal PD FS MRI. The edges of the joint (anterior and posterior 9 mm of femoral head) are scored in 4 zones: AS: anterior-superior, AI: anterior-inferior, PS: posterior-superior, PI: posterior-inferior. D: coronal PD FS MRI showing one of the central slices of joint (between anterior and posterior 9 mm). Scoring is in 5 zones: CL: central-lateral, CS: central-superior, CC: central-central, CI: central-inferior, CM: central-medial. From Roemer, *et al*. Osteoarthritis Cartilage 2011;19:946-62; adapted with permission.

proton density fat-saturated sequences. At each location a score is based on osteophyte size using a 5-point scale (Table 2). HIMRISS does not score osteophytes. In HOAMS, interobserver agreement on osteophyte score was moderate (best kappa 0.63), but no correlation to pain was observed[6].

Bone Marrow Lesion

Rationale for feature inclusion in OMERACT filter. Osteoarthritic joints frequently show signal changes in periarticular bone marrow (Table 3). The meaning of these changes is controversial, particularly the "edema-like" pattern of ill-defined increased T2/decreased T1 marrow

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2014. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org

Table 3. Signal changes in subarticular bone and bone marrow[18,19].

T1	T2	Post-gadolinium Enhancement	Other	Interpretation
Low	High	Homogeneous	Ill defined margins	Edema-like
Low	High	Heterogeneous	Well defined margins	Necrosis-like
Low	Low	No		Sclerosis
Low	Low	Maybe	Generally surrounded by edema-like	Sclerosis-like
Linear low signal line	Surrounding high		Subchondral bone deformity	Subchondral fracture
Low	Very high	Rim enhancement	Well defined, rounded	Cyst

signal. This is traditionally reported as "bone marrow edema" by most radiologists[18], has complex and controversial histologic meaning and clinical significance, prompting others to describe it as "bone marrow lesion"[19].

BML is present in patients with all grades of OA, with increasing frequency paralleling increasing OA severity in the knee joint[20]. It correlates strongly with the location of increased tracer uptake on radionuclide bone scan, i.e., a region of increased osteoblast activity[21]. Location of BML also correlates strongly with sites of increased loading, prompting Felson, *et al* to suggest that it is a marker of "ongoing bone trauma"[22]. Histologic findings at sites of bone marrow "edema" vary (Table 4), potentially because edema in extracellular tissue such as bone marrow is naturally transient, and may not be captured effectively *ex vivo*. One study, in which bone was decalcified prior to histologic examination, found much less extensive edema in specimens than on MRI[23], while another, where decalcification was not done, found the opposite, with the "edema-like" MRI marrow pattern having higher SP (0.95) than SN (0.8) for true histologic edema, implying that MRI was actually missing areas of edema[24]. This discrepancy highlights the potential for specimen processing to affect results and shows the limitations of a histologic gold standard.

In aggregate, the relevant studies all confirm that "edema-like" MRI marrow signal (i.e., BML) is not specific to "edema" but includes areas of other changes such as inflammatory infiltrate, fat necrosis, microfracture and healing, and microscopic cyst formation[23,24,25,26]. In conditions other than OA, ill-defined periarticular marrow signal changes can reflect "an entire spectrum of pathological conditions"[23], including pannus, vascularity and fluid associated with inflammatory cell infiltrates in inflammatory arthropathy[27]; necrosis and reorganizing scar in avascular necrosis; callus in fractures[26]; marrow replacing lesions such as infection or leukemia[28]. Poor correlations between MRI BML and histologic edema in earlier studies[23,26] have been improved by careful distinction between the different subtypes of marrow changes on MRI (Table 3)[24], and it is likely that OA scoring systems will be more successful when using these distinctions.

BML is strongly associated with cartilage loss. At hip MR arthrography, 100% of patients with BML had focal cartilage loss (although prevalence of cartilage loss in the study was high at 76%), and the grade of BML increased with the increasing grade of cartilage loss, r = 0.46[5]. The authors suggested that BML is present early in OA, and in hips its presence is a specific indicator that cartilage loss must be present. This is mechanically plausible: loss of protective overlying cartilage is likely to increase focal trauma to subchondral bone, resulting in BML and pain. In cases of rapidly destructive hip OA (i.e., chondrolysis), all patients had prominent BML with joint effusion, and rarely osteophytes[29].

Table 4. Studies of histologic findings in areas of marrow "edema" on magnetic resonance imaging (MRI) at the hip.

Authors	Design	Finding in Specimen at Site of MRI "Edema"	Extent of Finding on Specimen vs MRI
Hofmann, et al 1993[25]	Marrow from core decompression in transient osteoporosis of hip	"Pale to dark staining homogeneous fluid" Necrotic and reparative changes, ? early osteonecrosis	Specimens "filled" with fluid
Taljanovic, et al 2008[23]	Resected hip components post total arthroplasty. Decalcified specimens, MRI within 48 h pre-surgery	"Palely eosinophilic extracellular fluid"	Much less than on MRI: 8/20 patients, only 3% of subchondral bone
Leydet-Quilici, et al, 2010[24]	Resected femoral heads post total arthroplasty. Non-decalcified specimens, MRI within 1 mo pre-surgery	Edema (SN 80%, SP 95%) Combined with vascular marrow fibrosis (SN 69%, SP 80%) Some necrosis	MRI more specific than sensitive for true edema Kappa = 0.77

SN: sensitivity; SP: specificity.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2014. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org

Feature scoring. Schmid, *et al*[4] did not consider BML. Neumann, *et al*[5] scored marrow edema in the same 10 regions of femur and acetabulum as cartilage damage (Figure 2b), on a 3-point scale based on distance of extension of edema beneath subchondral bone in each region (Table 2).

In HOAMS[6], regions for BML are divided similarly to cartilage, except that femoral head and acetabulum are considered separately (Figure 3a-b). This gives 7 subarticular zones (4 femoral, 3 acetabular) on sagittal slices at anterior and posterior joint margins, and 8 zones (5 femoral, 3 acetabular) on the central coronal slices. In each region, a score is assigned based on the fraction of the volume of that region involved (Table 2), and there is variation in the volume of each region. In the central-superior and central-central regions of the acetabulum (CSA and CCA, respectively), which have no natural boundary away from the joint, the evaluable region is considered to have depth of 2 cm from the articular surface. The reader scrolls through all slices of the scan to assess volume of each region. The base sequences remain PD FS.

HIMRISS assesses BML over a larger number of small subregions (total 100), recording only a simple binary score (present/absent) in each. It uses coronal STIR base sequence with 3–4 mm slice spacing, more fluid-sensitive than the PD FS sequence in HOAMS, but often with lower spatial resolution. The key to proper HIMRISS scoring is correct region definition. These regions are about equal in volume, with some compromises made for ease of reading (Figure 3c). The reader identifies the "central slice" of the femoral head, either by coregistration with axial sequence if available, or as the slice on which the head appears largest, then fits an overlay template, either as a printed transparency or a semitransparent screen window. For typical adult femoral head size, the central 5 slices (centered on the "central slice") show most of the head volume. In each of these slices, 9 sectors are defined, 8 around a clock face (starting superomedially) and the last central. For up to 5 remaining anterior slices containing femoral head, a pair of regions (anterior-superior, anterior-inferior) are formed. This is also done for up to 5 posterior slices. Thus the femoral head has $5 \times 9 = 45$ "central" regions, and up to 5×2 anterior and 5×2 posterior regions, for a total of 65 regions (Table 2).

As in HOAMS, the acetabular marrow in HIMRISS is scored for a maximum depth of 2 cm from the joint. In the central 5 slices as described above, the acetabulum is divided into 3 regions containing a similar volume of subarticular bone (Figure 3c). In up to 5 slices anterior to the central slices, acetabulum is scored in 2 regions each (superior and inferior), and this is repeated for up to 5 slices posterior to central slices. This gives $5 \times 3 + 5 \times 2 + 5 \times 2 = 35$ regions. If no acetabulum is present, that slice is scored 0. Scoring is fastest if on each slice the reader identifies each BML and then maps this to the correct region(s), rather than checking each region for edema individually.

Reliability of BML assessment on HOAMS was better than cartilage assessment, with excellent interobserver agreement (best kappa 0.85) on 15 hips, representing the most reliable single feature in the HOAMS score[6]. Intrareader reliability for HIMRISS was excellent for detection of BML with ICC = 0.94 for status scores at both baseline and at 8 weeks. For change scores at 8 weeks after intraarticular administration of steroids, ICC was 0.86[8]. Regarding the truth component, HOAMS BML correlated significantly to Kellgren-Lawrence grade of radiographic OA, and approached significance for association with pain ($p = 0.09$)[6]. HIMRISS showed no significant correlation with WOMAC (Western Ontario and McMaster Universities Osteoarthritis Index) pain or patient global for either status or change scores[8]. There was no significant change in HIMRISS BML score 8 weeks after administration of intraarticular steroid. These systems show promise but neither has been tested in large datasets or directly validated against a tissue gold standard. Validation should be regarded as preliminary.

Subchondral Cysts

Rationale and scoring. Subchondral cysts are well known to be associated with OA, but represent structural damage (either intraosseous herniation of synovial fluid or fluid in a necrotic cavity) rather than the active inflammation thought to be associated with BML. Thus, these are scored separately from BML in the Neumann, *et al* and HOAMS systems, and are omitted entirely from HIMRISS. In HOAMS, cysts are scored in the same region as BML, based on the extent of cysts in each region (Table 2). In both HOAMS and HIMRISS the reader should also examine available T1 sequences to help distinguish cysts (well-defined low T1 signal) from BML (vaguely mildly decreased T1 signal). In HOAMS, agreement on subchondral cyst scoring was poor (best kappa 0.15) and there was no correlation to pain scores[6].

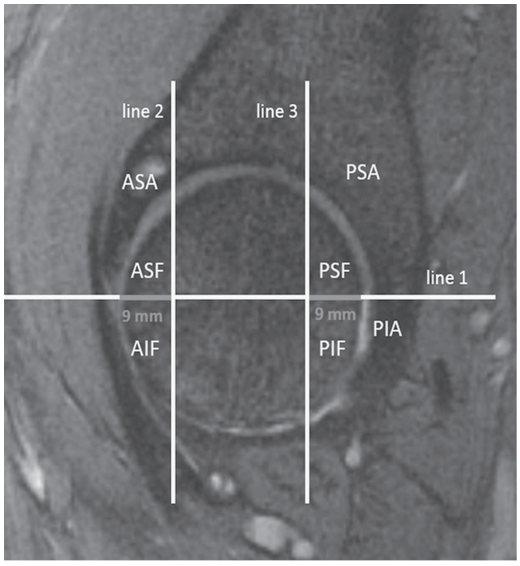
Synovitis/Effusion

Rationale for feature inclusion in OMERACT filter. Although OA can be distinguished from septic arthritis or other arthropathies such as rheumatoid arthritis by the lack of excess inflammatory cells in synovial fluid, there is increasing recognition that OA does have an inflammatory component. In a recent longitudinal study in knees, presence of baseline effusion-synovitis assessed on non-contrast enhanced MRI in knees without OA predicted cartilage loss at average 30 month followup[30]. In this study synovitis was also positively associated with presence and extent of BML. The relation between synovitis, OA, and pain has not been directly tested in the hip.

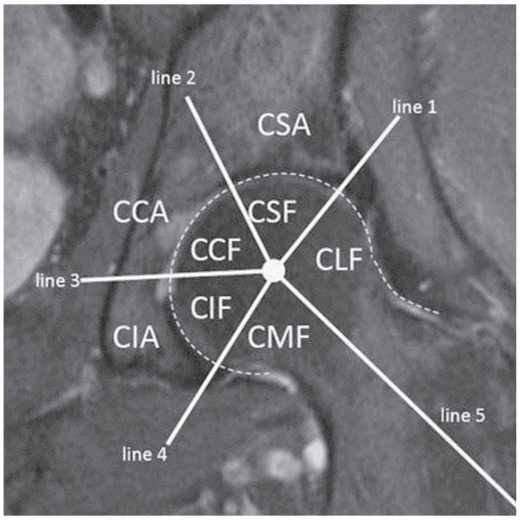
Feature scoring. Neumann, *et al* did not score synovitis or

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2014. All rights reserved.

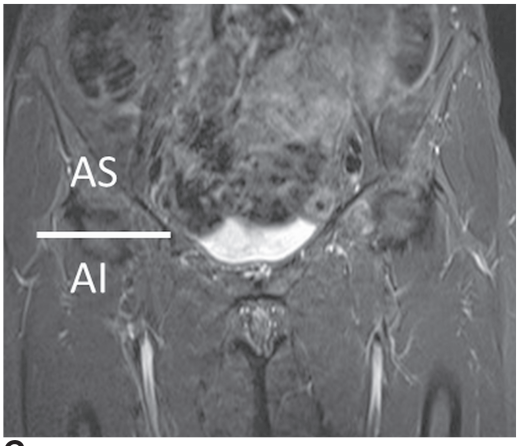
Downloaded on April 23, 2024 from www.jrheum.org



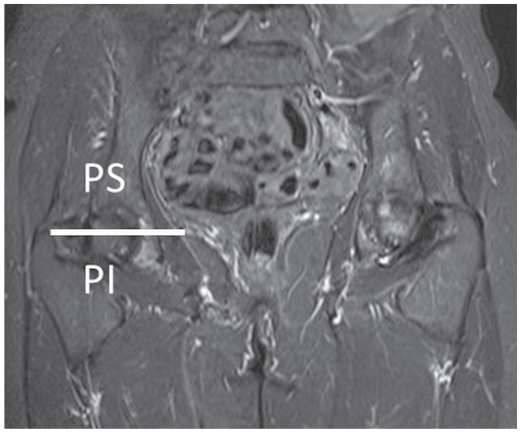
A



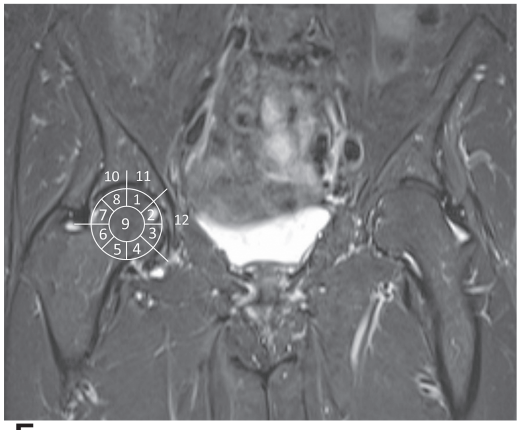
B



C



D



E

Figure 3. Assessment of bone marrow lesions on magnetic resonance imaging (MRI). A, B. Method according to Hip OsteoArthritis MRI Scoring System[6]. A: Sagittal proton density fat saturated (PD FS) MRI showing scoring for edges of joint, at anterior and posterior 9 mm of femoral head. Seven zones, including ASF: anterior-superior femoral, AIF: anterior-inferior femoral, PSF: posterior-superior femoral, PIF: posterior-inferior femoral, ASA: anterior-superior acetabulum, PSA: posterior-superior acetabulum, PIA: posterior-inferior acetabulum. B: Coronal PD FS MRI at one of the central slices of the joint (between anterior and posterior 9 mm). Three acetabular zones (CSA: central-superior acetabulum, CCA: central-central acetabulum, CIA: central-inferior acetabulum) and 5 femoral head zones (CLF: central-lateral femoral, CSF: central-superior femoral, CCF: central-central femoral, CIF: central-inferior femoral, CMF: central-medial femoral). C. Method according to Hip Inflammation MRI Scoring System (HIMRISS)[8]. Coronal short-tau inversion recovery MRI. The central 5 slices of femoral head (right image) have 8 circle sectors, numbered 1 = superomedial, clockwise to 8 = superolateral; center = 9. The central 5 slices of acetabulum have 3 sectors each, numbered 10 (lateral) to 12 (medial). Anterior slices (D) and posterior slices (E) have 2 regions for femoral head and acetabulum (e.g., anterior superior femoral, anterior superior acetabular). Edema in each region receives a score of 1, no edema = score of 0. AS: anterosuperior; AI: anteroinferior; PS: posterosuperior; PI: posteroinferior. See www.arthritisdoctor.ca for more detailed methodology. From Roemer, *et al*. Osteoarthritis Cartilage 2011;19:946-62; adapted with permission.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2014. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org

effusion. Using coronal and axial fat-saturated proton density images, HOAMS scores effusion as a single number (0, 1, or 2) indicating overall degree of capsular distention, and synovitis based on synovial thickness (0: < 2 mm, 1: 2–4 mm, 2: > 4 mm) at 4 locations along the femoral head-neck junction using contrast-enhanced T1 weighted fat saturated sequences: anterior, posterior, medial, and lateral[6]. HIMRISS instead measures the maximum combined thickness of synovium and fluid contacting femur, perpendicular to the bone, and repeats this measurement in each of the same 15 slices used for acetabular scoring. Maximum combined thickness at each slice is then given the same score (0–2) as in HOAMS (Table 2), for a total possible score of 2 × 15 = 30. This feature is therefore weighted much more heavily in HIMRISS than in HOAMS. Interobserver agreement on synovitis and effusion in HOAMS was only moderate (best kappa 0.60 and 0.65), and although both significantly correlated to radiographic OA grade, neither predicted pain. The synovitis/effusion component of HIMRISS has not been reported.

Labrum

Rationale for feature inclusion in OMERACT filter. Tears of the acetabular labrum, long thought to be the foremost cause of mechanical hip pain[31], are strongly associated with cartilage defects and clinical symptoms in MR arthrographic studies. Of 100 clinic patients with hip pain, 66 had labral tears, and 85% of these had cartilage abnormalities, 55% in the same region as the labral tear[5]. In a similar study, 100% of patients having arthroscopy had cartilage lesions, nearly all (37/42) in anterosuperior location adjacent to labrum, and of these, 67% had labral tears[4]. Labral tears and focal cartilage defects, together with focal bony protuberance, form the triad of MRI findings typical of cam type femoral-acetabular impingement[31,32]. There is a chicken-and-egg argument: do labral tears alter the mechanics of the hip joint, predisposing to OA, or are labral tears and cartilage injuries both sequelae of the same excessively traumatic loading[31]?

Feature scoring. Neumann, *et al* and HOAMS both scored the labrum in 4 quadrants on a 4-point scale capturing signal changes and tear complexity (Table 2). Neumann used MR arthrographic images, while HOAMS assesses labrum on high-resolution PD FS images without arthrogram. HOAMS also adds a fifth location: the anterior labrum on sagittal image (often the best view of a small labral tear). The labrum is not scored in HIMRISS. In HOAMS, interobserver agreement on labral score was moderate (best kappa 0.48), and correlation between presence of high-grade labral tear and pain approached significance ($p = 0.09$)[6].

Joint Geometry

Any incongruency of the hip joint — either acetabular undercoverage (hip dysplasia) or overcoverage (femoral-acetabular impingement, FAI) — strongly predisposes to development of early hip OA[33,34]. Both types of incongruency are potentially treatable. They can be assessed by numerous indices on radiographs, computed tomography, or MRI. These measurements become problematic in patients with established OA, because as joints become arthritic, alignment necessarily becomes less congruent as a result of asymmetric wear, and femoral osteophytes can develop in the same location as the lateral femoral head-neck junction bony protuberance seen in cam FAI, leading to confusion (is the osteophyte causing impingement, or did impingement cause the osteophyte to form, or both?).

Neumann, *et al* and HIMRISS do not include these features in scoring. HOAMS does not assess geometry in detail, but does add a single index of "dysplasia" if there is undercoverage of the acetabular roof based on measurement of lateral centre-edge angle. HOAMS also scores presence of a superolateral femoral head cyst, which may indicate impingement (Table 2). In HOAMS, there was moderate interobserver agreement (kappa 0.58) on cyst presence, and no observed dysplasia; these did not correlate to pain[6].

Other indices scored. HOAMS also scores miscellaneous features of hip OA such as presence of loose bodies, attrition/flattening of the femoral head, paralabral cysts, and adjacent bursitis and tendinopathy (Table 2). These are not scored by Neumann, *et al* or in HIMRISS, and showed no correlation to pain[6].

Other indices not scored. The major acquired external risk factors for hip OA are obesity (bilateral disease) and prior injury (unilateral disease)[35]. Markers of obesity such as subcutaneous fat thickness have not been incorporated into current scores, and none of the scores make note of the presence of healed fractures. The rectus femoris proximal tendon and iliopsoas tendon are intimately associated with hip joint capsule, but strain of these tendons is not assessed on any of the scoring systems.

The validation of hip OA MRI-based scoring methodologies is confined to 2 methods and the preliminary assessment of reliability by 2 readers with special expertise in the assessment of MRI scans from patients with OA. Assessment of responsiveness is limited by the lack of effective treatment for hip OA. Considerable validation is required to determine the truth and discrimination aspects of the OMERACT filter, which will require prospective studies and a multicenter approach. Priority should be addressed toward the assessment of reliability using the 2 methods that are currently available.

REFERENCES

1. Tugwell P, Boers M, Brooks P, Simon L, Strand V, Idzerda L. OMERACT: an international initiative to improve outcome measurement in rheumatology. Trials 2007;8:38.
2. Roemer FW, Crema MD, Trattnig S, Guermazi A. Advances in imaging of osteoarthritis and cartilage. Radiology 2011;260:332-54.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2014. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org

3. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494-502.
4. Schmid MR, Notzli HP, Zanetti M, Wyss TF, Hodler J. Cartilage lesions in the hip: diagnostic effectiveness of MR arthrography. Radiology 2003;226:382-6.
5. Neumann G, Mendicuti AD, Zou KH, Minas T, Coblyn J, Winalski CS, et al. Prevalence of labral tears and cartilage loss in patients with mechanical symptoms of the hip: evaluation using MR arthrography. Osteoarthritis Cartilage 2007;15:909-17.
6. Roemer FW, Hunter DJ, Winterstein A, Li L, Kim YJ, Cibere J, et al. Hip Osteoarthritis MRI Scoring System (HOAMS): reliability and associations with radiographic and clinical findings. Osteoarthritis Cartilage 2011;19:946-62.
7. Lambert RG, Hutchings EJ, Grace MG, Jhangri GS, Conner-Spady B, Maksymowych WP. Steroid injection for osteoarthritis of the hip: a randomized, double-blind, placebo-controlled trial. Arthritis Rheum 2007;56:2278-87.
8. Lambert RG, Maksymowych WP, Gracey D. Evaluation of bone marrow edema with the hip inflammation MRI scoring system (HIMRISS): is it reliable and do scores correlate with clinical status or response to therapy? [abstract]. Ann Rheum Dis 2011;70 Suppl 3:550.
9. Stelzeneder D, Mamisch TC, Kress I, Domayer SE, Werlen S, Bixby SD, et al. Patterns of joint damage seen on MRI in early hip osteoarthritis due to structural hip deformities. Osteoarthritis Cartilage 2012;20:661-9.
10. Baum T, Joseph GB, Arulanandan A, Nardo L, Virayavanich W, Carballido-Gamio J, et al. Association of magnetic resonance imaging-based knee cartilage T2 measurements and focal knee lesions with knee pain: data from the Osteoarthritis Initiative. Arthritis Care Res 2012;64:248-55.
11. Hunter DJ, Conaghan PG, Peterfy CG, Bloch D, Guermazi A, Woodworth T, et al. Responsiveness, effect size, and smallest detectable difference of Magnetic Resonance Imaging in knee osteoarthritis. Osteoarthritis Cartilage 2006;14 Suppl A:A112-5.
12. Jaremko JL, Cheng RW, Lambert RG, Habib AF, Ronsky JL. Reliability of an efficient MRI-based method for estimation of knee cartilage volume using surface registration. Osteoarthritis Cartilage 2006;14:914-22.
13. Outerbridge RE. The etiology of chondromalacia patellae. J Bone Joint Surg Br 1961;43B:752-7.
14. Hjelle K, Solheim E, Strand T, Muri R, Brittberg M. Articular cartilage defects in 1,000 knee arthroscopies. Arthroscopy 2002;18:730-4.
15. Noyes FR, Stabler CL. A system for grading articular cartilage lesions at arthroscopy. Am J Sports Med 1989;17:505-13.
16. Recht MP, Piraino DW, Paletta GA, Schils JP, Belhobek GH. Accuracy of fat-suppressed three-dimensional spoiled gradient-echo FLASH MR imaging in the detection of patellofemoral articular cartilage abnormalities. Radiology 1996;198:209-12.
17. Wright AA, Cook C, Abbott JH. Variables associated with the progression of hip osteoarthritis: a systematic review. Arthritis Rheum 2009;61:925-36.
18. Wilson AJ, Murphy WA, Hardy DC, Totty WG. Transient osteoporosis: transient bone marrow edema? Radiology 1988;167:757-60.
19. Felson DT, Chaisson CE, Hill CL, Totterman SM, Gale ME, Skinner KM, et al. The association of bone marrow lesions with pain in knee osteoarthritis. Ann Intern Med 2001;134:541-9.
20. 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;226:373-81.
21. Boegard T, Rudling O, Dahlstrom J, Dirksen H, Petersson IF, Jonsson K. Bone scintigraphy in chronic knee pain: comparison with magnetic resonance imaging. Ann Rheum Dis 1999;58:20-6.
22. Felson DT, McLaughlin S, Goggins J, LaValley MP, Gale ME, Totterman S, et al. Bone marrow edema and its relation to progression of knee osteoarthritis. Ann Intern Med 2003;139 Pt 1:330-6.
23. Taljanovic MS, Graham AR, Benjamin JB, Gmitro AF, Krupinski EA, Schwartz SA, et al. Bone marrow edema pattern in advanced hip osteoarthritis: quantitative assessment with magnetic resonance imaging and correlation with clinical examination, radiographic findings, and histopathology. Skeletal Radiol 2008;37:423-31.
24. Leydet-Quilici H, Le Corroller T, Bouvier C, Giorgi R, Argenson JN, Champsaur P, et al. Advanced hip osteoarthritis: magnetic resonance imaging aspects and histopathology correlations. Osteoarthritis Cartilage 2010;18:1429-35.
25. Hofmann S, Engel A, Neuhold A, Leder K, Kramer J, Plenk H Jr. Bone-marrow oedema syndrome and transient osteoporosis of the hip. An MRI-controlled study of treatment by core decompression. J Bone Joint Surg Br 1993;75:210-6.
26. Zanetti M, Bruder E, Romero J, Hodler J. Bone marrow edema pattern in osteoarthritic knees: correlation between MR imaging and histologic findings. Radiology 2000;215:835-40.
27. Appel H, Loddenkemper C, Grozdanovic Z, Ebhardt H, Dreimann M, Hempfing A, et al. Correlation of histopathological findings and magnetic resonance imaging in the spine of patients with ankylosing spondylitis. Arthritis Res Ther 2006;8:R143.
28. Ragab Y, Emad Y, Abou-Zeid A. Bone marrow edema syndromes of the hip: MRI features in different hip disorders. Clin Rheumatol 2008;27:475-82.
29. Boutry N, Paul C, Leroy X, Fredoux D, Migaud H, Cotten A. Rapidly destructive osteoarthritis of the hip: MR imaging findings. AJR Am J Roentgenol 2002;179:657-63.
30. Roemer FW, Guermazi A, Felson DT, Niu J, Nevitt MC, Crema MD, et al. Presence of MRI-detected joint effusion and synovitis increases the risk of cartilage loss in knees without osteoarthritis at 30-month follow-up: the MOST study. Ann Rheum Dis 2011;70:1804-9.
31. McCarthy JC, Busconi B. The role of hip arthroscopy in the diagnosis and treatment of hip disease. Orthopedics 1995;18:753-6.
32. Kassarjian A, Yoon LS, Belzile E, Connolly SA, Millis MB, Palmer WE. Triad of MR arthrographic findings in patients with cam-type femoroacetabular impingement. Radiology 2005;236:588-92.
33. Reid GD, Reid CG, Widmer N, Munk PL. Femoroacetabular impingement syndrome: an underrecognized cause of hip pain and premature osteoarthritis? J Rheumatol 2010;37:1395-404.
34. Fisher R, O'Brien TS, Davis KM. Magnetic resonance imaging in congenital dysplasia of the hip. J Pediatr Orthop 1991;11:617-22.
35. Cooper C, Inskip H, Croft P, Campbell L, Smith G, McLaren M, et al. Individual risk factors for hip osteoarthritis: obesity, hip injury, and physical activity. Am J Epidemiol 1998;147:516-22.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2014. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org