Magnetic Resonance Imaging (MRI) of the Knee as an Outcome Measure in Juvenile Idiopathic Arthritis: An OMERACT Reliability Study on MRI Scales

Robert Hemke, Nikolay Tzaribachev, Charlotte M. Nusman, Marion A.J. van Rossum, Mario Maas, and Andrea S. Doria

ABSTRACT. Objective. There is increasing evidence that early therapeutic intervention improves longterm joint outcome in juvenile idiopathic arthritis (JIA). Given the existence of highly effective treatments, there is an urgent need for reliable and accurate measures of disease activity and joint damage in JIA. Our objective was to assess the reliability of 2 magnetic resonance imaging (MRI) scoring methods: the Juvenile Arthritis MRI Scoring (JAMRIS) system and the International Prophylaxis Study Group (IPSG) consensus score, for evaluating disease status of the knee in patients with JIA.

Methods. Four international readers independently scored an MRI dataset of 25 JIA patients with clinical knee involvement. Synovial thickening, joint effusion, bone marrow changes, cartilage lesions, bone erosions, and subchondral cysts were scored using the JAMRIS and IPSG systems. Further, synovial enhancement, infrapatellar fat pad heterogeneity, tendinopathy, and enthesopathy were scored. Interreader reliability was analyzed by using the generalized κ, ICC, and the smallest detectable difference (SDD).

Results. ICC regarding interreader reliability ranged from 0.33 (95% CI 0.12–0.52, SDD = 0.29) for enthesopathy up to 0.95 (95% CI 0.92–0.97, SDD = 3.19) for synovial thickening. Good interreader reliability was found concerning joint effusion (ICC 0.93, 95% CI 0.89–0.95, SDD = 0.51), synovial enhancement (ICC 0.90, 95% CI 0.85–0.94, SDD = 9.85), and bone marrow changes (ICC 0.87, 95% CI 0.80–0.92, SDD = 10.94). Moderate to substantial reliability was found concerning cartilage lesions and bone erosions (ICC 0.55–0.72, SDD 1.41–13.65).

Conclusion. The preliminary results are promising for most of the scored JAMRIS and IPSG items. However, further refinement of the scoring system is warranted for unsatisfactorily reliable items such as bone erosions, cartilage lesions, and enthesopathy. (J Rheumatol First Release June 1 2017; doi:10.3899/jrheum.160821)

Key Indexing Terms:
JUVENILE IDIOPATHIC ARTHRITIS   MAGNETIC RESONANCE IMAGING
OUTCOMES   REPRODUCIBILITY OF RESULTS   KNEE JOINT

Persistent synovitis can lead to osteochondral abnormalities that are responsible for disability and impaired quality of life in patients with juvenile idiopathic arthritis (JIA)[1,2,3,4]. The main goal of treatment in JIA is the complete suppression of joint inflammation. Therefore, outcome measures in daily practice and clinical trials must include sensitive and reliable measures of disease activity[5].

Magnetic resonance imaging (MRI) holds potential to become an important outcome measure for assessment of joints in patients with JIA[5,6]. Although MRI is the preferred imaging modality for detection of synovial inflammation, and early destructive bone changes in JIA[7,8,9], experiences on the use of MRI for assessing JIA is limited. Consequently, this method is underused in both clinical practice and research. One of the most important causes for underuse of MRI for assessment of JIA is the absence of standardized protocols and scales for data acquisition and interpretation[5,6,10]. The use of such protocols and scales by different centers across the globe is crucial for meaningful and comparable results acquired in multicenter clinical trials.

Over the years, different MRI scoring methods have been

developed for assessment of joints. Two of them focus on the assessment of large pediatric joints. The Juvenile Arthritis MRI Scoring (JAMRIS) system[11] focuses on JIA and the International Prophylaxis Study Group (IPSG) MRI Scoring system[12] focuses on hemophilic arthropathy. Both the JAMRIS and IPSG systems consist of 2 domains: 1 soft tissue and 1 osteochondral (bone and cartilage items). Whereas the soft tissue domain of the JAMRIS system consists of synovial hypertrophy, the soft tissue domain of the IPSG system consists of synovial hypertrophy and joint effusion. In the JAMRIS system, the bone items include bone marrow edema (BME), bone erosions, and cartilage lesions/loss. In the IPSG system, the bone item includes bone erosions, subchondral cysts, and cartilage loss. Subchondral cysts are items of the IPSG MRI system, but not of the JAMRIS system. Nevertheless, whereas in the JAMRIS system the final score for each item represents the sum of maximum scores in pre-established sites of the knee, in the IPSG system the final score for each item represents the maximum score of any area of the knee without summing up scores of regions.

Regardless of their conceptual methodological differences, both scoring methods are limited because they are not internationally validated for the evaluation of disease activity in JIA[6]. Prior to the Outcome Measures in Rheumatology (OMERACT) 11 meeting, a Special Interest Group (SIG) was formed focusing on MRI in JIA. One of the aims of this SIG was to standardize and develop an MRI outcome measure to objectively assess disease activity, including large joints such as the knee[10]. To develop a unique consensus, MRI scale for the assessment of large pediatric joints, we used items from existing MRI scales and assessed the reliability of the aforementioned MRI scoring systems. Therefore, the objective of our study was to assess the reliability of the JAMRIS and the IPSG scoring systems for evaluating disease activity and osteochondral changes in JIA of the knee.

MATERIALS AND METHODS

Design and patients. To evaluate the interobserver reliability of the JAMRIS and IPSG systems, 4 readers independently scored MRI datasets of 25 JIA patients with current or previous knee involvement attending the pediatric rheumatology clinic of 2 tertiary centers from 2011 to 2014 (Academic Medical Center, Amsterdam, the Netherlands, and the Hospital for Sick Children, Toronto, Ontario, Canada). Contrast-enhanced MRI were selected from examinations of clinically active patients with JIA obtained in Amsterdam, the Netherlands (n = 20) and Toronto, Ontario, Canada (n = 5). Readers consisted of 1 musculoskeletal radiologist (MM, 19 yrs of experience), 1 pediatric radiologist (ASD, 11 yrs of experience), 1 pediatric rheumatologist (NT, 9 yrs of experience), and 1 radiology trainee (RH, 6 yrs of experience). All readers were blinded to clinical history.

Patients fulfilled the International League of Associations for Rheumatology criteria for JIA[13]. Institutional Review Board approval was granted for the current study, with informed consent being waived in both centers [Academic Medical Center (NL41846.018.12), Hospital for Sick Children (10000422167)]. The study was performed in accordance with the Declaration of Helsinki.

MRI protocol. The MRI sequences acquired in Amsterdam were obtained using an open-bore 1.0T magnet (Panorama HFO, Philips Medical Systems) and those derived from Toronto were obtained using a standard 1.5T magnet (Avanto-Fit, Siemens Healthcare). The nonsedated children were placed in the supine position with the knee joint lying centrally in the magnetic field in a dedicated knee coil (16 channel and 8 channel knee coil in Amsterdam and Toronto, respectively).

Sequences included sagittal T2-weighted fat-saturated images, coronal T2-weighted fat-saturated images, axial T2-weighted fat-saturated images, sagittal T1-weighted (obtained in Amsterdam) and sagittal proton-density (obtained in Toronto), and sagittal and axial T1-weighted fat-saturated images obtained after intravenous (IV) contrast injection (obtained in both centers). Postcontrast axial images were obtained in the early phase (< 5 min) after IV injection of gadolinium. Additional sagittal 3-D double-echo steady state images of the knee were obtained in Toronto for cartilage assessment.

MRI scoring system. MRI datasets were scored by using the JAMRIS and the IPSG systems, complemented with 4 additional MRI features: infra-patellar fat pad heterogeneity, enthesopathy, tendinopathy, and synovial enhancement. The rationale for including these additional features in our analysis was to determine the effect of them on the reliability of the scales. A prior study has assessed the value of these features in the content validity of the JAMRIS system[14].

An easy-to-use scoring form is depicted in the Supplementary Data (available with the online version of this article).

The JAMRIS system. The JAMRIS scoring method has been described before in detail[12]. Briefly, synovial thickening was scored when thickness of the contrast-enhanced synovial membrane was ≥ 2 mm. Synovial thickening was scored at 6 locations. The maximal thickness of any slice at each site was graded as follows: grade 0 = < 2 mm, grade 1 = ≥ 2–4 mm, and grade 2 = > 4 mm, resulting in a minimum score of 0 and a maximum of 12.

Bone marrow changes, cartilage lesions, and bone erosions were scored at 8 locations. BME was scored semiquantitatively based on the subjectively estimated percentage of involved bone volume at each site as follows: grade 0 = none, grade 1 = < 10%, grade 2 = ≥ 10–25%, and grade 3 = > 25% of the whole bone volume or of the region of the cartilage surface area, resulting in a minimum score of 0 and a maximum of 24.

MRI definitions of BME, cartilage lesions, and bone erosions were previously described[12]. For example, a bone erosion was defined as a sharply marginated bone lesion with correct juxtaarticular localization, typical signal characteristics, and visible in 2 planes with a cortical break in at least 1 plane. On T1-weighted images, there is a loss of the normal low signal intensity of cortical bone and loss of the normal high signal intensity of trabecular bone.

The IPSG scoring system. The IPSG score has been developed and validated for use in hemophilic arthritis and described before[11,15]. Briefly, synovial thickening was scored when thickness of the contrast-enhanced synovial membrane was > 1 mm. In the IPSG score, the location with the maximal score does count[16]. Synovial thickening was graded semiquantitatively as follows: grade 0 = none, grade 1 (mild) = > 2–3 mm, grade 2 (moderate) = > 3–5 mm, and grade 3 (severe) = > 5 mm.

Joint effusion was scored when the maximal diameter of the largest pocket was ≥ 3 mm as previously described[16].

Erosions and subchondral cysts were graded semiquantitatively according to the area involved. Bone erosions were scored by evaluating the presence of any surface erosion. For bone erosions, 1 score was given for any surface erosion, and 1 score was given for half or more of the articular surface eroded in at least 1 bone, yielding a maximum of 2. The presence of subchondral cysts was scored when there was at least 1 subchondral cyst anywhere in the knee. For subchondral cysts, 1 score was given for at least 1 subchondral cyst, and 1 score was given for subchondral cysts in at least 2 bones, yielding a maximum of 2. Cartilage lesions/loss were scored by evaluating the presence of any loss of the normal cartilage. Normal values for cartilage thickness in healthy children was used for the evaluation of cartilage thinning[17]. For cartilage loss, 1 score was given for any cartilage damage, 1 score was given for loss of half or more of the total volume of

joint cartilage in at least 1 bone, 1 score was given for full thickness loss of joint cartilage in at least some area in at least 1 bone, and 1 score was given for full thickness loss of joint cartilage including at least 1 half of the joint surface in at least 1 bone yielding a maximum of 4 scores per joint. MRI definitions of bone erosions, subchondral cysts, and cartilage loss were previously described[16].

Additional MRI features. Four additional MRI features were scored as well, including synovial enhancement, infrapatellar fat pad heterogeneity, enthesopathy, and tendinopathy.

The degree of synovial enhancement was scored per JAMRIS region as described by Damasio, *et al*[18] as follows: 0 = normal synovial enhancement as compared to neighboring muscle, 1 = mildly increased synovial enhancement, and 2 = moderately to severely increased synovial enhancement.

The infrapatellar fat pad was scored for heterogeneity (0 = absent, 1 = present) being caused by water infiltration or by scar tissue[14].

Tendinopathy of enthesopathy of the patellar and quadriceps tendons were assessed for increased signal intensity on T2-weighted images within the tendon or the enthesis, respectively (0 = absent, 1 = present).

Statistics. Descriptive statistics were calculated for each reader. Reliability was comprehensively evaluated with 3 statistical methods. The average measure ICC was classified as follows: ICC < 0.40 = poor, ≥ 0.40–0.60 = moderate, > 0.60–0.80 = substantial, and > 0.80 = excellent reliability[19]. Second, the smallest detectable difference (SDD) was assessed deriving from the limits of agreement method. The SDD was calculated for all the summated scores using the residual error variance from repeated measures ANOVA[20,21]. Moreover, the SDD was calculated as a percentage of the highest score reached. For categorical data, the generalized κ statistics were used and classified as follows: κ < 0.20 = poor, > 0.20–0.40 = fair, > 0.40–0.60 = moderate, > 0.60–0.80 = substantial, and > 0.80 = excellent reliability.

RESULTS

Patients. In our reliability study, MRI datasets of 25 patients with JIA were scored (60% women; mean age 12.3 yrs, SD 2.7; mean disease duration 5.6 yrs, SD 4.8). Frequencies of JIA subtypes were as follows: 8/25 (32%) persistent oligoarthritis, 4/25 (16%) extended oligoarthritis, 8/25 (32%) polyarthritis, 2/25 (8%) psoriatic arthritis (PsA), and 3/25 (12%) enthesitis-related arthritis. Of these patients, 4 (16%) were clinically inactive, while 21 (84%) had clinically active disease activity.

Reliability study. Descriptive statistics (mean, minimum and maximum scores) per reader of the different item scores are shown in Table 1.

Interreader reliability concerning MRI features focusing on disease activity was excellent for all scores (ICC 0.87–0.95; Table 2). Synovial thickening according to the JAMRIS and IPSG scores were ICC 0.95 (SDD 3.19) and 0.94 (SDD 0.39), respectively. Synovial enhancement was ICC 0.90 (SDD 9.85), joint effusion was ICC 0.93 (SDD 0.51), and bone marrow changes was ICC 0.87 (SDD 10.94). An example of synovial thickening with a good agreement between readers is depicted in Figure 1, whereas Figure 2 illustrates a case with good agreement on the scoring of joint effusion.

Interreader reliability for the MRI features focusing on destructive joint changes was moderate to substantial (ICC 0.55–0.72; Table 2). Assessment of cartilage lesions according to the JAMRIS and ISPG scores demonstrated ICC 0.55 (SDD 13.65) and 0.60 (SDD 8.30), respectively. Evaluation of bone erosions according to the JAMRIS and ISPG scores showed ICC 0.72 (SDD 1.41) and 0.57 (SDD 4.10), respectively. Figure 3 depicts an example of a bone erosion that was not scored by all readers. The generalized κ of subchondral cysts according to the IPSG score was poor (κ = 0.01).

Generalized κ concerning the additional scored MRI features (infrapatellar fat pad heterogeneity, enthesopathy, and tendinopathy) were relatively poor (range 0.01–0.27; Table 2).

Table 1. Descriptive statistics per reader of the different items scored from 25 juvenile idiopathic arthritis patients with MRI of the knee. Values are mean (interquartile range) unless otherwise specified.

MRI Feature	Reader 1	Reader 2	Reader 3	Reader 4
JAMRIS item				
Synovial thickening, 0–12	2.8 (0–10)	2.6 (0–9)	2.8 (0–12)	3.2 (0–12)
Bone marrow changes, 0–24	3.1 (0–19)	2.7 (0–11)	2.4 (0–20)	2.4 (0–15)
Cartilage lesions, 0–24	1.2 (0–9)	0.9 (0–6)	0.7 (0–8)	2.6 (0–18)
Bone erosions, 0–24	0.6 (0–4)	0.4 (0–4)	0.8 (0–10)	0.7 (0–3)
IPSG item				
Synovial thickening, 0–3	1.4 (0–3)	1.3 (0–3)	1.2 (0–3)	1.3 (0–3)
Joint effusion, 0–3	1.1 (0–3)	1.1 (0–3)	1.1 (0–3)	1.1 (0–3)
Cartilage lesions, 0–4	0.9 (0–3)	0.7 (0–3)	0.4 (0–3)	0.8 (0–4)
Bone erosions, 0–2	0.3 (0–2)	0.2 (0–2)	0.2 (0–2)	0.1 (0–2)
Subchondral cysts, 0–2	0.0 (0–0)	0.0 (0–0)	0.0 (0–1)	0.0 (0–1)
Synovial enhancement, 0–12*	4.6 (0–12)	4.2 (0–12)	4.0 (0–12)	5.2 (0–12)
Infrapatellar fat pad heterogeneity,				
% positive	44 (11/25)	40 (10/25)	28 (7/25)	52 (13/25)
Enthesopathy, % positive	12 (3/25)	20 (5/25)	4 (1/25)	36 (9/25)
Tendinopathy, % positive	16 (4/25)	4 (1/25)	4 (1/25)	28 (7/25)

* According to Damasio, *et al*[18]. MRI: magnetic resonance imaging; JAMRIS: Juvenile Arthritis MRI Scoring; IPSG: International Prophylaxis Study Group.

Table 2. Generalized κ, ICC, and SDD statistics of the different items scored from 25 juvenile idiopathic arthritis patients with MRI of the knee.

MRI Feature	ICC or κ	95% CI	SDD	SDD%
JAMRIS item				
Synovial thickening, 0–12	ICC 0.95	0.92–0.97	3.19	27
Bone marrow changes, 0–24	ICC 0.87	0.80–0.92	10.94	55
Cartilage lesions, 0–24	ICC 0.55	0.37–0.69	13.65	76
Bone erosions, 0–24	ICC 0.72	0.59–0.81	1.41	35
IPSG item				
Synovial thickening, 0–3	ICC 0.94	0.91–0.96	0.39	13
Joint effusion, 0–3	ICC 0.93	0.89–0.95	0.51	17
Cartilage lesions, 0–4	ICC 0.60	0.44–0.73	8.30	69
Bone erosions, 0–2	ICC 0.57	0.40–0.70	4.10	41
Subchondral cysts, 0–2	κ 0.01	NA	NA	NA
Synovial enhancement, 0–12*	ICC 0.90	0.85–0.94	9.85	82
Infrapatellar fat pad heterogeneity, 0–1	κ 0.27	NA	NA	NA
Enthesopathy, 0–1	κ 0.19	NA	NA	NA
Tendinopathy, 0–1	κ 0.01	NA	NA	NA

* According to Damasio, *et al*[18]. MRI: magnetic resonance imaging; SDD: smallest detectable difference; NA: not applicable; JAMRIS: Juvenile Arthritis MRI Scoring; IPSG: International Prophylaxis Study Group.
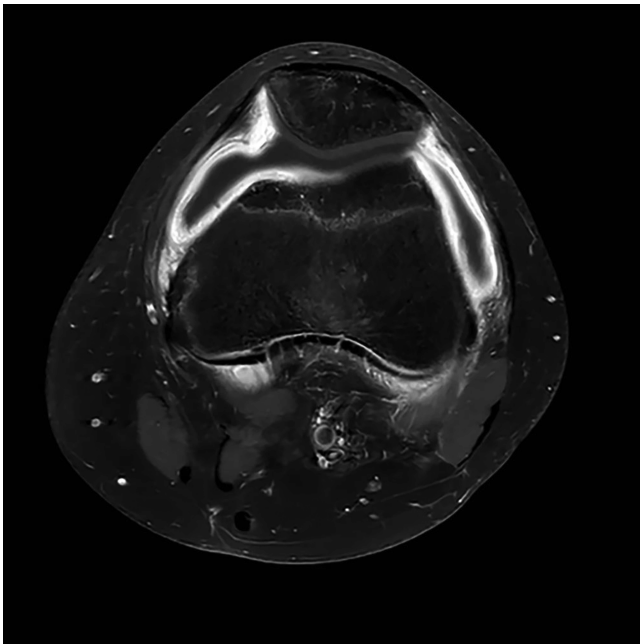

Figure 1. Synovial thickening. Axial fat-saturated contrast-enhanced T1-weighted image obtained in a 12-year-old girl showing retropatellar and parapatellar synovial thickening.
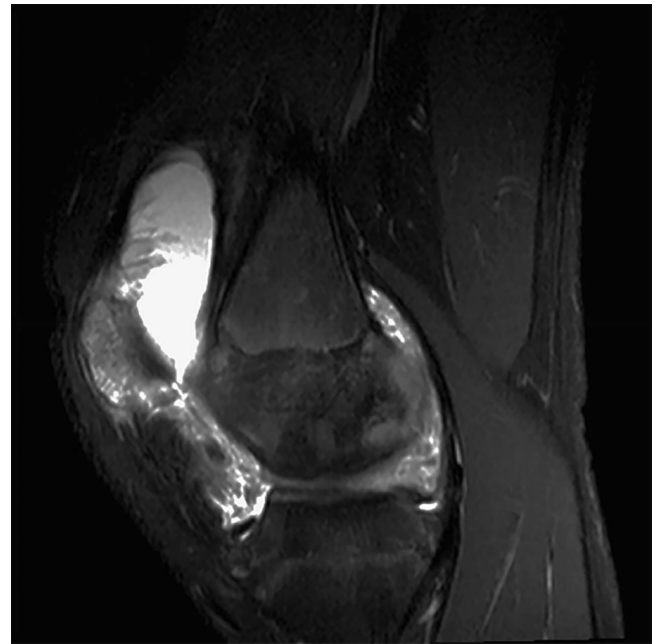

Figure 2. Joint effusion. Sagittal fat-saturated T2-weighted image obtained in a 14-year-old girl showing extensive joint effusion in the suprapatellar recess.

DISCUSSION
In our study, we assessed the interreader agreement of the JAMRIS and the IPSG scoring systems for evaluating disease activity of the knee in patients with JIA. Overall the inter-reader reliability was excellent for MRI scores focusing on active disease (e.g., synovial hypertrophy, effusion, and bone marrow changes). Interreader reliability focusing on osteo-chondral changes (e.g., cartilage lesions, bone erosions) was moderate to substantial.

Out of 2 MRI scales designed to assess morphologic changes of the knee in pediatric patients[11,12], the JAMRIS system was purposely developed for the use in JIA and the IPSG scoring system for the use in hemophilic arthritis. Although the pathophysiology, clinical presentation, and genetic background of hemophilic arthritis differ from those of JIA, there is a considerable structural overlap of MRI characteristics of affected joints that may present with (reactive) synovial hypertrophy, bone marrow changes, and

Figure 3. Bone erosion. Sagittal T1-weighted postcontrast magnetic resonance image obtained in a 17-year-old girl with clinically active juvenile idiopathic arthritis shows a bone erosion in the tibial plateau with loss of normal high signal intensity of the trabecular bone and irregular cortical outlines.

destruction of cartilage and bone. In our current study, we were intentionally broadly inclusive of scoring items that could identify a wide range of MRI features in these 2 types of arthritis. Therefore, items of each score were assessed independently. The scoring methods used have some conceptual methodological differences. For example, in the JAMRIS system, the final score for each item represents the sum of maximum scores in preestablished sites of the knee, and in the IPSG system, the final score for each item represents the maximum score of any area of the knee without summing up scores of regions. By doing so, the JAMRIS system has a significantly wider range of score compared with the method used in the IPSG score, making the JAMRIS system, in all probability, more sensitive to change compared with the IPSG method. On the other hand, a simpler score with a lower range of score (such as the IPSG method) could be more reliable when focusing on less common abnormalities. The sensitivity to change over time of the methods used should be evaluated in future studies.

Although abnormalities of the infrapatellar fat pad and tendons might be associated with disease activity in JIA, in our present study, most discrepancies were noted about the assessment of infrapatellar fat pad heterogeneity, enthesopathy, and tendinopathy (κ 0.01–0.27). Therefore, we should consider improving the definitions for characterization of these findings or excluding them from a final score in the future if appropriate.

MRI has some technical limitations for its use in children.

Limitations include, among others, the necessity for sedation in young children and the limited number of joints that can be evaluated during 1 imaging session. Moreover, MRI using an IV contrast agent is indispensable for the sensitive differentiation of joint effusion and synovial hypertrophy[8]. Despite these limitations, MRI is deemed the most sensitive technique for the evaluation of disease status in patients with JIA[6]. The appropriateness and feasibility of different imaging modalities differs with age. Therefore, the use of ultrasound can be helpful in patients with JIA as well.

Our study has methodological limitations. The most important drawback is the lack of severely affected patients with JIA. The reliability of the JAMRIS and IPSG score has only been tested in patients with JIA visiting referral pediatric rheumatology centers with full access to current treatment. This has resulted in a population of studied patients with only mild to moderate disease activity. Consequently, the presence of destructive changes was relatively low. Additional research is necessary to evaluate the value of a scoring system as a sensitive measure regarding these destructive changes. For further validation of an objective and easy-to-use scoring system, international collaboration is warranted, especially with research centers with access to more severely affected patients with JIA.

Second, there is sparse information about normal MRI values in healthy knees of children at different ages[17,22]. For example, thinning of the articular cartilage can be either physiologic or pathologic, and bony depressions simulating joint damage are common in wrists of healthy children[23]. Early joint damage in children with JIA can, therefore, be masked on MRI[24]. International collaboration is required to solve these challenges in the interpretation of MRI. The development of MRI scores to measure JIA disease activity and joint damage in growing joints can be otherwise seriously limited, which affects the care of children with JIA.

Another limitation is the lack of a specific cartilage-sensitive MRI technique in a great number of examinations. The Amsterdam cohort had only proton-density sequence available for assessment of cartilage. This could have reduced the readers' ability to identify early cartilage changes. It is well known that proton density–weighted imaging is capable of depicting surface cartilaginous defects as well as abnormalities of internal cartilage composition. Nevertheless, institutions prefer to use intermediate-weighted sequences that combine the contrast advantage of proton density weighting with that of T2 weighting by using a TE of 33–60 ms. These sequences provide higher overall signal intensity in cartilage than standard T2-weighted sequences do, allowing better differentiation between cartilage and subchondral bone. In addition, they are less susceptible to the magic angle effects seen in proton density–weighted imaging with a shorter TE[25].

Previous studies have pointed out that with fat-suppressed proton density and T2-weighted fast spin echo sequences,

normal hyaline cartilage has intermediate signal intensity, and intraarticular fluid appears bright, thus displaying good contrast because of an "arthrographic" effect that identifies surface abnormalities as well as abnormalities in cartilage matrix[26]. Proton density-weighted sequences provide greater contrast in the cartilage structure and evaluate the cartilage-synovial fluid interface and the subchondral bone, being previously used for detecting chondral lesions[27,28,29].

There was a variance concerning the level of experience and training of the readers, different backgrounds and slightly different MRI protocols for data acquisition; this might have resulted in lower reliability values in the osteochondral domain compared with the Amsterdam reliability studies of the JAMRIS system[12]. However, our reliability scores of disease activity and osteochondral damage were comparable to the values obtained in reliability studies regarding JIA and rheumatoid arthritis MRI scores of the wrist[30,31]. On the other hand, with regard to the assessment of joint damage (e.g., cartilage lesions, bone erosions), in our study reliability scores were somewhat lower (ICC 0.55–0.72) than the scores of the wrist in patients with JIA as described by Malattia, *et al*[30] and the patients with PsA as described by the OMERACT SIG in Inflammatory Arthritis[31]. Our findings are in line with validation exercises focusing on the wrist in patients with JIA, in which it was shown that damage MRI scores can be relatively inaccurate[18,32,33,34]. Finally, the studied cohort was not consecutive, which may have caused selection bias and had mild osteochondral changes that reduced the generalizability of the results of our study to populations with more severe changes.

Moreover, reliability was evaluated using the ICC in a relatively small number of MRI datasets, and data were considered as continuous. This topic is controversial because some authors consider the variables of scoring systems whose maximum value is above a certain limit (e.g., 10) to be continuous rather than categorical data. Nevertheless, our intention was to compare the results of our study with results from the OMERACT studies conducted in adults that also used ICC and SDD[35,36].

To develop and validate an easy-to-use MRI scoring system, further formal expert consensus processes and discrete choice experiments are required to determine the relevance of items of the scale and to refine definitions of items that demonstrated poor interreader reliability in our study. Another question is whether to use the extensive JAMRIS method for the quantification of findings per MRI feature or the more simplified IPSG method in a final future MRI scale. Although an extensive quantification scoring method is relatively more time-consuming, it is probably more accurate as well, especially in depicting small changes over time.

Several advanced MRI techniques are available for the evaluation of inflammatory and destructive changes in JIA, including dynamic contrast–enhanced MRI, T2-mapping, diffusion-weighted MRI, and delayed gadolinium-enhanced MRI of cartilage. Currently, these imaging techniques are used particularly in the context of research and to a lesser extent in daily practice. Moreover, specialized knowledge is needed for the interpretation and postprocessing of these images. To date, advanced imaging techniques are therefore not valuable for use in an easy-to-use scoring method such as the JAMRIS system.

Our preliminary results are promising for most scored JAMRIS and IPSG items for assessment of mild joint changes in JIA. Because interreader reliability, focusing on destructive changes, was moderate to substantial in the studied cohort, further refinement of the scoring system items and definitions are warranted for the next step of development of a single MRI scoring system to assess different stages of disease activity in patients with JIA. With this regard, the concurrent development of an atlas with representative images of different stages of disease activity in JIA can potentially improve the interpretability of the scoring system items.

ONLINE SUPPLEMENT
Supplementary material accompanies the online version of this article.

REFERENCES
1. Haverman L, Grootenhuis MA, van den Berg JM, van Veenendaal M, Dolman KM, Swart JF, et al. Predictors of health-related quality of life in children and adolescents with juvenile idiopathic arthritis: results from a web-based survey. Arthritis Care Res 2012; 64:694-703.
2. Magni-Manzoni S, Rossi F, Pistorio A, Temporini F, Viola S, Beluffi G, et al. Prognostic factors for radiographic progression, radiographic damage, and disability in juvenile idiopathic arthritis. Arthritis Rheum 2003;48:3509-17.
3. Palmisani E, Solari N, Magni-Manzoni S, Pistorio A, Labo E, Panigada S, et al. Correlation between juvenile idiopathic arthritis activity and damage measures in early, advanced, and longstanding disease. Arthritis Rheum 2006;55:843-9.
4. Oen K, Reed M, Malleson PN, Cabral DA, Petty RE, Rosenberg AM, et al. Radiologic outcome and its relationship to functional disability in juvenile rheumatoid arthritis. J Rheumatol 2003;30:832-40.
5. Miller E, Uleryk E, Doria AS. Evidence-based outcomes of studies addressing diagnostic accuracy of MRI of juvenile idiopathic arthritis. AJR Am J Roentgenol 2009;192:1209-18.
6. Colebatch-Bourn AN, Edwards CJ, Collado P, D'Agostino MA, Hemke R, Jousse-Joulin S, et al. EULAR-PReS points to consider for the use of imaging in the diagnosis and management of juvenile idiopathic arthritis in clinical practice. Ann Rheum Dis 2015;74:1946-57.
7. Malattia C, Damasio MB, Magnaguagno F, Pistorio A, Valle M, Martinoli C, et al. Magnetic resonance imaging, ultrasonography, and conventional radiography in the assessment of bone erosions in juvenile idiopathic arthritis. Arthritis Rheum 2008;59:1764-72.
8. Hemke R, Kuijpers TW, van den Berg JM, van Veenendaal M, Dolman KM, van Rossum MA, et al. The diagnostic accuracy of unenhanced MRI in the assessment of joint abnormalities in juvenile idiopathic arthritis. Eur Radiol 2013;23:1998-2004.
9. Gylys-Morin VM, Graham TB, Blebea JS, Dardzinski BJ, Laor T, Johnson ND, et al. Knee in early juvenile rheumatoid arthritis: MR imaging findings. Radiology 2001;220:696-706.

10. Hemke R, Doria AS, Tzaribachev N, Maas M, van der Heijde DM, van Rossum MA. Selecting magnetic resonance imaging (MRI) outcome measures for juvenile idiopathic arthritis (JIA) clinical trials: first report of the MRI in JIA special interest group. J Rheumatol 2014;41:354-8.

11. Feldman BM, Funk S, Lundin B, Doria AS, Ljung R, Blanchette V; International Prophylaxis Study Group (IPSG). Musculoskeletal measurement tools from the International Prophylaxis Study Group (IPSG). Haemophilia 2008;14:162-9.

12. Hemke R, van Rossum MA, van Veenendaal M, Terra MP, Deurloo EE, de Jonge MC, et al. Reliability and responsiveness of the Juvenile Arthritis MRI Scoring (JAMRIS) system for the knee. Eur Radiol 2013;23:1075-83.

13. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J, et al; International League of Associations for Rheumatology. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31:390-2.

14. Hemke R, van Veenendaal M, Kuijpers TW, van Rossum MA, Maas M. Increasing feasibility and patient comfort of MRI in children with juvenile idiopathic arthritis. Pediatr Radiol 2012;42:440-8.

15. Foppen W, van der Schaaf IC, Witkamp TD, Fischer K. Is joint effusion on MRI specific for haemophilia? Haemophilia 2014;20:582-6.

16. Doria AS, Keshava SN, Mohanta A, Jarrin J, Blanchette V, Srivastava A, et al. Diagnostic accuracy of ultrasound for assessment of hemophilic arthropathy: MRI correlation. AJR Am J Roentgenol 2015;204:W336-47.

17. Keshava SN, Gibikote SV, Mohanta A, Poonnoose P, Rayner T, Hilliard P, et al. Ultrasound and magnetic resonance imaging of healthy paediatric ankles and knees: a baseline for comparison with haemophilic joints. Haemophilia 2015;21:e210-22.

18. Damasio MB, Malattia C, Tanturri de Horatio L, Mattiuz C, Pistorio A, Bracaglia C, et al. MRI of the wrist in juvenile idiopathic arthritis: proposal of a paediatric synovitis score by a consensus of an international working group. Results of a multicentre reliability study. Pediatr Radiol 2012;42:1047-55.

19. Norman GR, Wyrwich KW, Patrick DL. The mathematical relationship among different forms of responsiveness coefficients. Qual Life Res 2007;16:815-22.

20. Lassere MN, van der Heijde D, Johnson KR, Boers M, Edmonds J. Reliability of measures of disease activity and disease damage in rheumatoid arthritis: implications for smallest detectable difference, minimal clinically important difference, and analysis of treatment effects in randomized controlled trials. J Rheumatol 2001;28:892-903.

21. Lassere M, Boers M, van der Heijde D, Boonen A, Edmonds J, Saudan A, et al. Smallest detectable difference in radiological progression. J Rheumatol 1999;26:731-9.

22. Spannow AH, Stenboeg E, Pfeiffer-Jensen M, Fiirgaard B, Haislund M, Ostergaard M, et al. Ultrasound and MRI measurements of joint cartilage in healthy children: a validation study. Ultraschall Med 2011;32 Suppl 1:S110-6.

23. Müller LS, Avenarius D, Damasio B, Eldevik OP, Malattia C, Lambot-Juhan K, et al. The paediatric wrist revisited: redefining MR findings in healthy children. Ann Rheum Dis 2011;70:605-10.

24. Doria AS, Babyn PS, Feldman B. A critical appraisal of radiographic scoring systems for assessment of juvenile idiopathic arthritis. Pediatr Radiol 2006;36:759-72.

25. Crema MD, Roemer FW, Marra MD, Burstein D, Gold GE, Eckstein F, et al. Articular cartilage in the knee: current MR imaging techniques and applications in clinical practice and research. Radiographics 2011;31:37-61.

26. Link TM. MR imaging in osteoarthritis: hardware, coils, and sequences. Radiol Clin North Am 2009;47:617-32.

27. Schaefer FK, Kurz B, Schaefer PJ, Fuerst M, Hedderich J, Graessner J, et al. Accuracy and precision in the detection of articular cartilage lesions using magnetic resonance imaging at 1.5 Tesla in an in vitro study with orthopedic and histopathologic correlation. Acta Radiol 2007;48:1131-7.

28. Mohr A. The value of water-excitation 3D FLASH and fat-saturated PDw TSE MR imaging for detecting and grading articular cartilage lesions of the knee. Skeletal Radiol 2003;32:396-402.

29. Hayashi D, Roemer FW, Guermazi A. Choice of pulse sequences for magnetic resonance imaging-based semiquantitative assessment of cartilage defects in osteoarthritis research: comment on the article by Dore et al. Arthritis Rheum 2010;62:3830-1.

30. Malattia C, Damasio MB, Pistorio A, Ioseliani M, Vilca I, Valle M, et al. Development and preliminary validation of a paediatric-targeted MRI scoring system for the assessment of disease activity and damage in juvenile idiopathic arthritis. Ann Rheum Dis 2010;70:440-6.

31. McQueen F, Lassere M, Duer-Jensen A, Wiell C, Conaghan PG, Gandjbakhch F, et al. Testing an OMERACT MRI scoring system for peripheral psoriatic arthritis in cross-sectional and longitudinal settings. J Rheumatol 2009;36:1811-5.

32. Tanturri de Horatio L, Damasio MB, Barbuti D, Bracaglia C, Lambot-Juhan K, Boavida P, et al. MRI assessment of bone marrow in children with juvenile idiopathic arthritis: intra- and inter-observer variability. Pediatr Radiol 2012;42:714-20.

33. Boavida P, Lambot-Juhan K, Müller LS, Damasio B, de Horatio LT, Malattia C, et al. Carpal erosions in children with juvenile idiopathic arthritis: repeatability of a newly devised MR-scoring system. Pediatr Radiol 2015;45:1972-80.

34. Lambot K, Boavida P, Damasio MB, Tanturri de Horatio L, Desgranges M, Malattia C, et al. MRI assessment of tenosynovitis in children with juvenile idiopathic arthritis: inter- and intra-observer variability. Pediatr Radiol 2013;43:796-802.

35. Conaghan P, Lassere M, Ostergaard M, Peterfy C, McQueen F, O'Connor P, et al. OMERACT Rheumatoid Arthritis Magnetic Resonance Imaging Studies. Exercise 4: an international multicenter longitudinal study using the RA-MRI Score. J Rheumatol 2003;30:1376-9.

36. Bird P, Ejbjerg B, McQueen F, Ostergaard M, Lassere M, Edmonds J. OMERACT Rheumatoid Arthritis Magnetic Resonance Imaging Studies. Exercise 5: an international multicenter reliability study using computerized MRI erosion volume measurements. J Rheumatol 2003;30:1380-4.