Consensus-driven conceptual development of a standardized whole body-MRI scoring system for assessment of disease activity in juvenile idiopathic arthritis: MRI in JIA OMERACT working group

Jyoti Panwar[a], Mirkamal Tolend[b], Bernadette Redd[c], Hemalatha Srinivasalu[d], Robert A. Colbert[e], Jonathan Akikusa[f], Simone Appenzeller[g], John A. Carrino[h], Nele Herregods[i], Lennart Jans[i], Kerri Highmore[j], Thekla von Kalle[k], Eva Kirkhus[l], Dax G. Rumsey[m], Jacob L. Jaremko[n], Inarejos Emilio J. Clemente[o], Marion A. van Rossum[p], Jennifer Stimec[b], Shirley M. Tse[q], Marinka Twilt[r], Nikolay Tzaribachev[s], Iwona Sudol-Szopinska[t], Arthur B. Meyers[u], Andrea S. Doria[b],*

[a] Department of Radiology, Christian Medical College, Vellore, India
[b] Department of Diagnostic Imaging, Research Institute, The Hospital for Sick Children, Department of Medical Imaging, University of Toronto, Toronto, ON, Canada
[c] Department of Radiology, Clinical Center, NIH, Bethesda, Maryland, United States
[d] Division of Rheumatology, Children's National Hospital and George Washington University School of Medicine, Washington, DC, United States
[e] Pediatric Translational Research Branch, Musculoskeletal and Skin Diseases, National Institute of Arthritis, NIH, Bethesda, MD, United States
[f] Rheumatology Service, Department of General Medicine, Royal Children's Hospital Melbourne, Australia
[g] Faculty of Medical Sciences, University of Campinas, Campinas, Brazil
[h] Department of Radiology, Hospital for Special Surgery, New York, United States
[i] Department of Radiology, Ghent University, Ghent, Belgium
[j] Department of Radiology, Children's Hospital of Eastern Ontario, Ottawa, ON, Canada
[k] RadiologischesInstitut, Olga Hospital Klinikum, Stuttgart, Germany
[l] Department of Radiology, Oslo University Hospital, Oslo, Norway
[m] Division of Rheumatology, Stollery Children's Hospital, University of Alberta, Canada
[n] Department of Radiology & Diagnostic Imaging, Stollery Children's Hospital, University of Alberta, Canada
[o] Department of Radiology, Hospital Sant Joan de Deu, Barcelona, Spain
[p] Amsterdam Rheumatology and Immunology Center, Reade, and Emma Children's Hospital Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands
[q] Division of Rheumatology, The Hospital for Sick Children, Toronto, ON, Canada
[r] Department of Pediatrics, Division of Rheumatology, Alberta Children's Hospital, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada
[s] Pediatric Rheumatology Research Institute, Bad Bramstedt, Germany
[t] National Institute of Geriatrics, Rheumatology and Rehabilitation, Warsaw,Poland
[u] Department of Radiology, Cincinnati Children's Hospital, Cincinnati, OH, United States

ARTICLE INFO

ABSTRACT

Objectives: Whole body-MRI is helpful in directing diagnostic and treatment approaches, and as a research outcome measure. We describe our initial consensus-driven phase towards developing a whole body-MRI scoring system for juvenile idiopathic arthritis.

Methods: An iterative approach using three rounds of anonymous Delphi surveys followed by a consensus meeting was used to draft the structure of the whole body-MRI scoring system, including the relevant anatomic joints and entheses for assessment, diagnostic item selection, definition and grading, and selection of appropriate MRI planes and sequences. The surveys were completed independently by an international expert group consisting of pediatric radiologists and rheumatologists.

Results: Twenty-two experts participated in at least one of three rounds of Delphi surveys and a concluding consensus meeting. A first iteration scoring system was developed which ultimately included the assessment of 100 peripheral, 23 chest, and 76 axial joints, and 64 entheses, with 2−4 diagnostic items graded in each of the items, using binary (presence/absence) and 2-3-level ordinal scores. Recommendations on anatomic MRI planes and sequences were specified as the minimally necessary imaging protocol for the scoring system.

 * Corresponding author.
 E-mail address: andrea.doria@sickkids.ca (A.S. Doria).

Conclusion: A novel whole body-MRI scoring system for juvenile idiopathic arthritis was developed by consensus among members of MRI in JIA OMERACT working group. Further iterative refinements, reliability testing, and responsiveness are warranted in upcoming studies.

Introduction

Juvenile idiopathic arthritis (JIA) is the most common chronic rheumatic disease of childhood and can involve peripheral and axial joints, and entheses [1]. Conventional MRI of peripheral and axial joints allows monitoring of disease activity and prediction of subsequent structural damage in JIA [2,3].

Whole-body MRI (WB-MRI) is increasingly being used for the evaluation of rheumatologic diseases assessing peripheral and axial joints, and entheses of the entire body on a single scanning session [4−12], thus overcoming the poor reliability of clinical joint examination especially the deep-seated joints like sacroiliac and temporomandibular joints (difficult to evaluate using ultrasound as well) [13] and providing a measure of the inflammatory load of the whole body. To our knowledge, no studies on WB-MRI have described a total inflammatory joint score for JIA patients [12].

The aim of this study was to develop a standardized WB-MRI scoring system to quantify the total inflammatory burden in children with JIA through formal consensus methods among an interdisciplinary group of experts.

Methods

This study was approved by the Research Ethics Board of The Hospital for Sick Children (Toronto, Canada).

Statements, definitions and items for the scoring system were discussed, reformulated, voted upon, and subsequently revised until consensus was achieved among participating members of the MRI in JIA (JAMRI) working group within the Outcome Measures in Rheumatology (OMERACT) network [14−16]. The group consisted of 13 pediatric radiologists from 13 institutions in 8 countries (Canada, n = 5; Belgium, n = 2; 1 each from Germany, Norway, United Kingdom, Spain, United States, Poland) and 5 pediatric rheumatologists from 5 institutions in 4 countries (2 from Canada, and 1 each from Australia, Germany and The Netherlands) who completed up to three rounds of anonymous, iterative Delphi surveys and attended post-survey discussion meetings over a 12-month period. All participants had at least 5 years of experience of musculoskeletal imaging after training at the time of the initial survey. The iterative surveys were used to decide the measurement scope, relevant anatomic joints and entheses for assessment, diagnostic item selection, definition and grading of items, and appropriate imaging sequences and planes, ultimately yielding a preliminary MRI scoring system. Previously published OMERACT MRI-based definitions were used as template definitions in initial surveys and were modified considering the pediatric population and the MRI sequences selected [9−11]. Agreement rates for potential choices, and additional suggested choices and questions were discussed amongst the group after each survey through video conference meetings and implemented in subsequent surveys. Results of the third-round survey were further discussed and refined during the OMERACT JAMRI group's face-to-face consensus meetings conducted at the Radiological Society of North America (RSNA) scientific assemblies held in November 2016 and November 2017 in Chicago, IL towards constructing the first draft. The stepwise development of the WB-MRI scoring system is shown in Fig. 1.

Results

Twenty-two members responded to the initial survey, 18 to the second, and 15 to the third (Fig. 1). The final consensus meeting was attended by 15 members, who voted on the final set of items and grading specifications derived from the preceding surveys and meetings, with ≥ 80% agreement being considered satisfactory. Subsequent minor revisions on wording of items and definitions were approved by all authors.

Scope of the outcome measure

Since active inflammation determines the need for treatment and frequency of follow up in JIA, the working group decided to limit the scope of the scoring system to synovial and entheseal inflammation in peripheral and axial joints. It was recognized that assessment of chronic osteochondral changes and estimation of total damage using a scoring system would be highly challenging and unreliable considering the low spatial resolution and large field-of-view of WB-MRI. Hence, our group decided to keep osteochondral damage as an ancillary assessment and prioritize inflammatory joint and entheseal changes for item and protocol selection.

Selection of anatomic sites for assessment

All peripheral joints including joints of upper and lower limbs, and chest outside spine (Fig. 2 A) were included in the MRI scoring system. Axial joints included sacroiliac joints (SIJs) and joints of the spine represented by all disco-vertebral units (DVUs) from C2-3 to L5-S1, atlanto-dental, pairs of lateral atlanto-occipital and atlanto-axial joints, and pairs of facets joints from C2-3 to L5-S1 (Fig. 2 B) [17]. Costovertebral, costotransverse, and temporomandibular joints were excluded from the scoring system due to the wide field-of-view and out of plane imaging of these articulations on WB-MRI. If any abnormality was identified in any of these joints as they were not part of the MRI scoring system, this would be recorded as an ancillary finding. The members agreed to include most of the entheseal sites of the body (Fig. 2 C).

First Iteration of the WB-MRI scoring system

The proposed scoring system is organized into 3 parts, one part dedicated to imaging of peripheral and chest joints (Fig. 2 A), one to axial joints including SIJ and spine (Fig. 2 B), and one to entheses (Fig. 2 C).

(1) Peripheral and chest joints

Effusion/synovial thickening, bone marrow edema (BME), and pericapsular soft tissue edema were selected as key findings for scoring peripheral and chest joints (Table 1A, Fig. 3 A). The single slice with most extensive inflammation was selected for scoring all items. Based on size or volume, the joints were divided into: (i) large and medium joints: glenohumeral, hip, knee, elbow; and (ii) small joints: acromioclavicular, sternoclavicular, manubriosternal, costochondral, wrist, ankle and small joints of hands and feet.

(a) Effusion/synovial thickening

Effusion/synovial thickening was considered altogether as a single item as it is often difficult to differentiate these two findings by MRI without intravenous administration of contrast. Whereas grading levels of large and medium joints ranged from 0 to 2 (ordinal data: 0-absent or normal amount of intraarticular fluid, 1-mild, 2-moderate/severe pathology), grading levels of small volume joints ranged from 0,1 (binary data: 0-absent or minimal trace amount of physiologic intraarticular fluid appearing as "pencil thin linear intra-articular
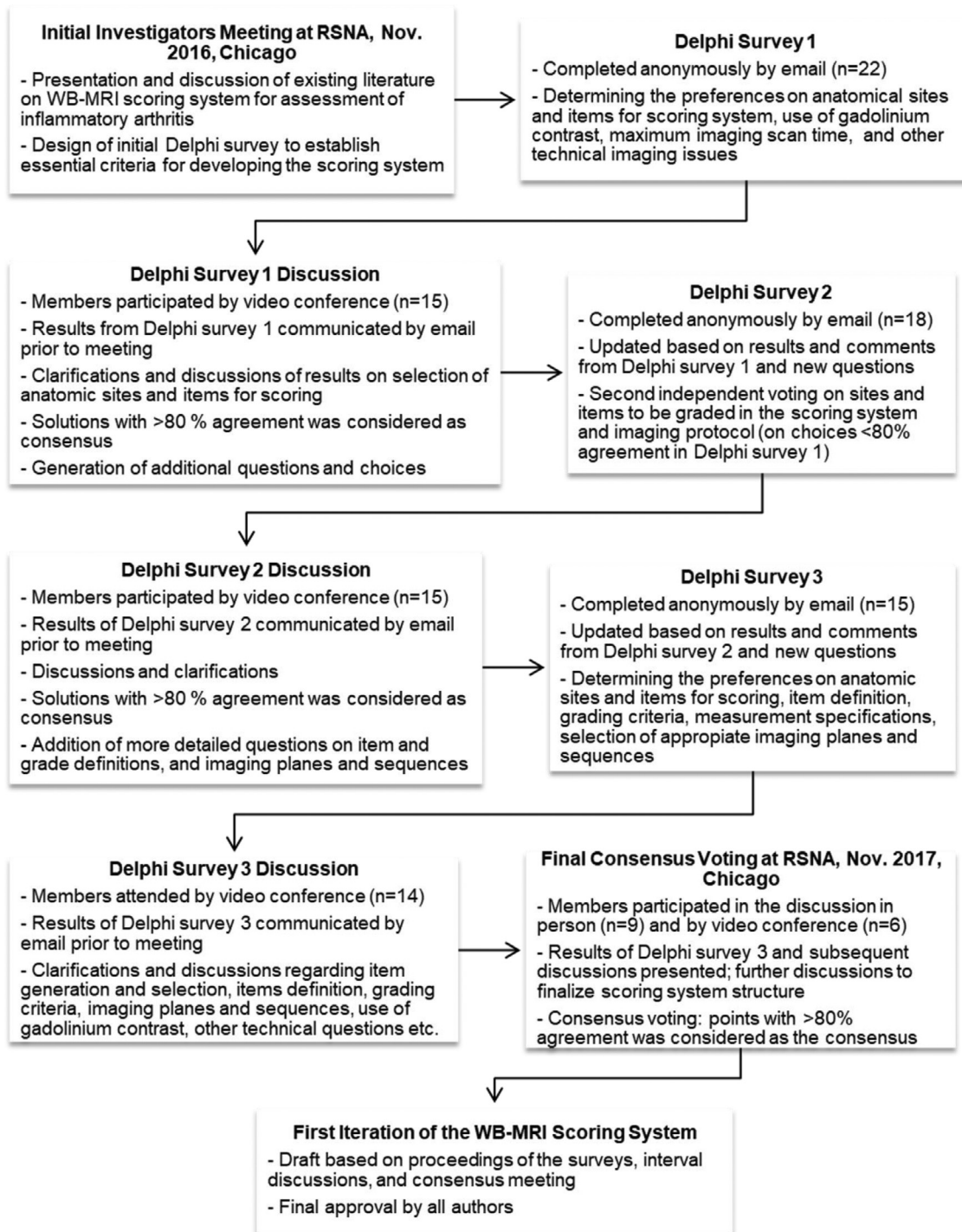
Initial Investigators Meeting at RSNA, Nov. 2016, Chicago

- Presentation and discussion of existing literature on WB-MRI scoring system for assessment of inflammatory arthritis

- Design of initial Delphi survey to establish essential criteria for developing the scoring system

Delphi Survey 1

- Completed anonymously by email (n=22)

- Determining the preferences on anatomical sites and items for scoring system, use of gadolinium contrast, maximum imaging scan time, and other technical imaging issues

Delphi Survey 1 Discussion

- Members participated by video conference (n=15)

- Results from Delphi survey 1 communicated by email prior to meeting

- Clarifications and discussions of results on selection of anatomic sites and items for scoring

- Solutions with >80 % agreement was considered as consensus

- Generation of additional questions and choices

Delphi Survey 2

- Completed anonymously by email (n=18)

- Updated based on results and comments from Delphi survey 1 and new questions

- Second independent voting on sites and items to be graded in the scoring system and imaging protocol (on choices <80% agreement in Delphi survey 1)

Delphi Survey 2 Discussion

- Members participated by video conference (n=15)

- Results of Delphi survey 2 communicated by email prior to meeting

- Discussions and clarifications

- Solutions with >80 % agreement was considered as consensus

- Addition of more detailed questions on item and grade definitions, and imaging planes and sequences

Delphi Survey 3

- Completed anonymously by email (n=15)

- Updated based on results and comments from Delphi survey 2 and new questions

- Determining the preferences on anatomic sites and items for scoring, item definition, grading criteria, measurement specifications, selection of appropiate imaging planes and sequences

Delphi Survey 3 Discussion

- Members attended by video conference (n=14)

- Results of Delphi survey 3 communicated by email prior to meeting

- Clarifications and discussions regarding item generation and selection, items definition, grading criteria, imaging planes and sequences, use of gadolinium contrast, other technical questions etc.

Final Consensus Voting at RSNA, Nov. 2017, Chicago

- Members participated in the discussion in person (n=9) and by video conference (n=6)

- Results of Delphi survey 3 and subsequent discussions presented; further discussions to finalize scoring system structure

- Consensus voting: points with >80% agreement was considered as the consensus

First Iteration of the WB-MRI Scoring System

- Draft based on proceedings of the surveys, interval discussions, and consensus meeting

- Final approval by all authors

Fig. 1. Stepwise development of the whole-body MRI scoring system.

hyperintense signal", 1-more than minimal trace amount of physiologic fluid) considering differences in size of the joint.

(b) Bone marrow edema (BME)

BME was graded according to both extent and intensity of edema. For assessment of extent of BME in each selected medium and large joint, the subchondral surface was evaluated in halves (2/2 for the entire joint) provided equal weight to the proximal and distal articular surfaces of the joint, *e.g.* acetabular and femoral articular surfaces (one score was given to signal abnormality noted in each half of the articular surface, binary data: 0-absent; 1-present, as shown in Fig. 3 A1). For the small joints, BME was recorded on either side of the

articular surface without segmentation and received a score of one for signal changes observed on each side of the articular surface (as shown in Fig. 3 A2). An additional score of 1 was given to indicate presence of intense edema (as shown in Fig. 3 A1). The reader would give a score for intensity of BME in each joint based on the anatomic area with most severe signal changes noted in this joint in any slice (not on a per bone basis).

BME was also recorded in carpals, tarsals, metacarpals, metatarsals and phalanges after excluding BME that was confined to the articular surface which had already been scored as part of joint involvement. A score of one was given to each involved bone.
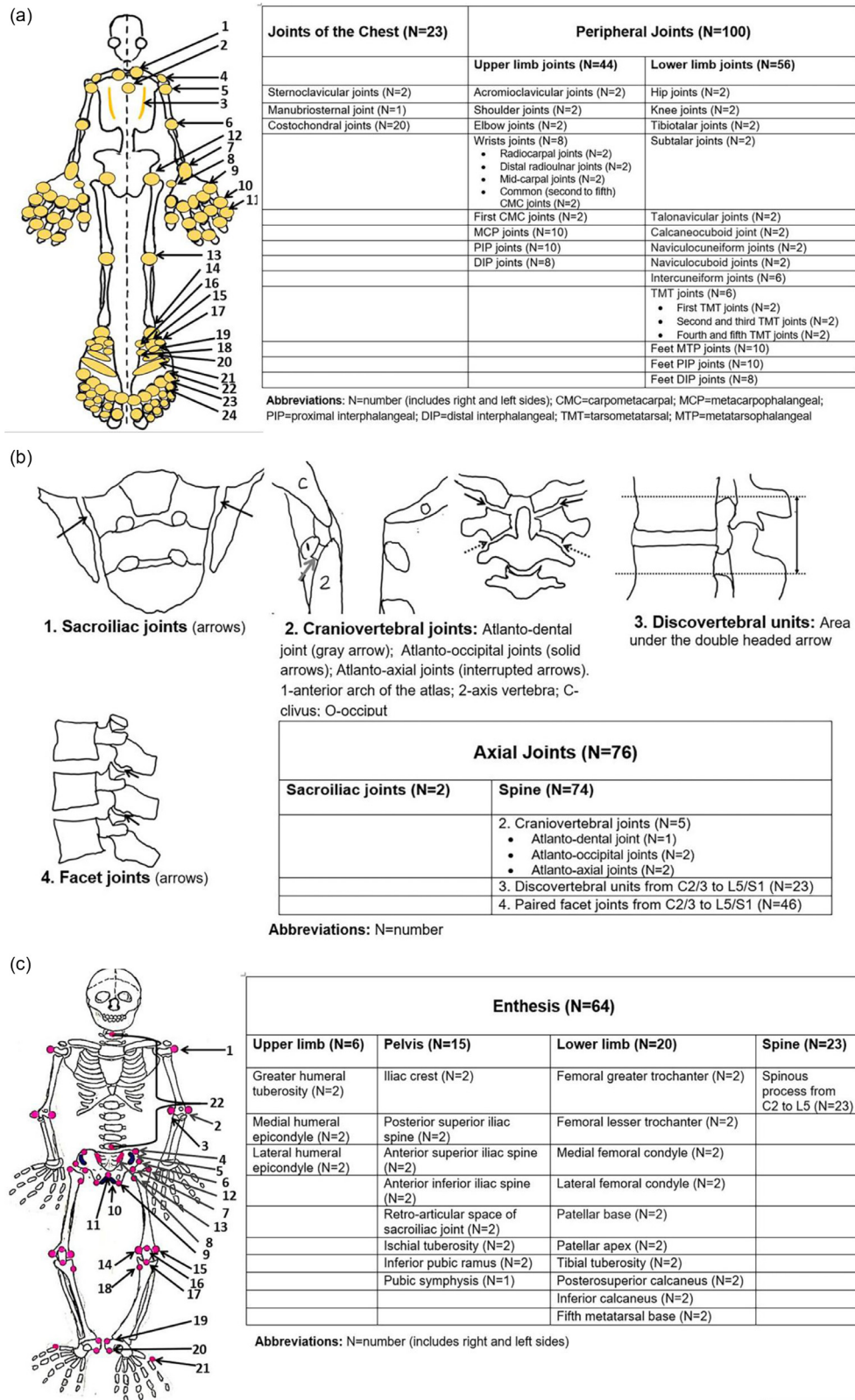
(a)

Joints of the Chest (N=23)	Peripheral Joints (N=100)	
	Upper limb joints (N=44)	Lower limb joints (N=56)
Sternoclavicular joints (N=2)	Acromioclavicular joints (N=2)	Hip joints (N=2)
Manubriosternal joint (N=1)	Shoulder joints (N=2)	Knee joints (N=2)
Costochondral joints (N=20)	Elbow joints (N=2)	Tibiotalar joints (N=2)
	Wrists joints (N=8) • Radiocarpal joints (N=2) • Distal radioulnar joints (N=2) • Mid-carpal joints (N=2) • Common (second to fifth) CMC joints (N=2)	Subtalar joints (N=2)
	First CMC joints (N=2)	Talonavicular joints (N=2)
	MCP joints (N=10)	Calcaneocuboid joint (N=2)
	PIP joints (N=10)	Naviculocuneiform joints (N=2)
	DIP joints (N=8)	Naviculocuboid joints (N=2)
		Intercuneiform joints (N=6)
		TMT joints (N=6) • First TMT joints (N=2) • Second and third TMT joints (N=2) • Fourth and fifth TMT joints (N=2)
		Feet MTP joints (N=10)
		Feet PIP joints (N=10)
		Feet DIP joints (N=8)

Abbreviations: N=number (includes right and left sides); CMC=carpometacarpal; MCP=metacarpophalangeal; PIP=proximal interphalangeal; DIP=distal interphalangeal; TMT=tarsometatarsal; MTP=metatarsophalangeal

(b)

1. **Sacroiliac joints** (arrows)

2. **Craniovertebral joints:** Atlanto-dental joint (gray arrow); Atlanto-occipital joints (solid arrows); Atlanto-axial joints (interrupted arrows). 1-anterior arch of the atlas; 2-axis vertebra; C-clivus; O-occiput

3. **Discovertebral units:** Area under the double headed arrow

4. **Facet joints** (arrows)

Axial Joints (N=76)	
Sacroiliac joints (N=2)	Spine (N=74)
	2. Craniovertebral joints (N=5) • Atlanto-dental joint (N=1) • Atlanto-occipital joints (N=2) • Atlanto-axial joints (N=2)
	3. Discovertebral units from C2/3 to L5/S1 (N=23)
	4. Paired facet joints from C2/3 to L5/S1 (N=46)

Abbreviations: N=number

(c)

Enthesis (N=64)			
Upper limb (N=6)	Pelvis (N=15)	Lower limb (N=20)	Spine (N=23)
Greater humeral tuberosity (N=2)	Iliac crest (N=2)	Femoral greater trochanter (N=2)	Spinous process from C2 to L5 (N=23)
Medial humeral epicondyle (N=2)	Posterior superior iliac spine (N=2)	Femoral lesser trochanter (N=2)	
Lateral humeral epicondyle (N=2)	Anterior superior iliac spine (N=2)	Medial femoral condyle (N=2)	
	Anterior inferior iliac spine (N=2)	Lateral femoral condyle (N=2)	
	Retro-articular space of sacroiliac joint (N=2)	Patellar base (N=2)	
	Ischial tuberosity (N=2)	Patellar apex (N=2)	
	Inferior pubic ramus (N=2)	Tibial tuberosity (N=2)	
	Pubic symphysis (N=1)	Posterosuperior calcaneus (N=2)	
		Inferior calcaneus (N=2)	
		Fifth metatarsal base (N=2)	

Abbreviations: N=number (includes right and left sides)

Figure 2. Selection of anatomic sites for assessment of inflammation in juvenile idiopathic arthritis using whole body (WB)-MRI.

(c) Pericapsular soft tissue edema

For the pericapsular soft tissue inflammation item, a 0,1 score (binary data) was used to grade this finding.

Definitions and item scoring for (a) effusion/synovial thickening, (b) BME and (c) pericapsular soft tissue edema are shown in Table 1A.

Ancillary items

Items that were not scored as a part of the scoring system but were recorded if present, such as chronic nonbacterial osteomyelitis (CNO)-like abnormalities in periphyseal regions, involvement of cost-overtebral and costotransverse joints, tenosynovitis, tendinosis and

Table 1
Structure of whole body (WB)-MRI scoring system.

A	Peripheral and Chest Joints		
Item	(a) Effusion/synovial thickening	(b) Bone marrow edema (BME)	(c) Pericapsular soft tissue inflammation
Definition	Hyperintense signal intensity (isointense to that of cerebrospinal fluid) within the joint space distending the joint capsule on T2-weighted fat-saturated or STIR or other fluid sensitive sequence. Note: Synovial thickening is not properly assessed with fluid sensitive sequences solely and on the given large field-of-view of WB-MRI. It may show low or intermediate signal on T2- or STIR images and requires the use of gadolinium to differentiate from effusion which is not part of this protocol.	Edema-like marrow signal changes within the sub-chondral bone on T2-weighted fat-saturated or STIR or other fluid sensitive sequence that is not compatible with normal hematopoietic marrow signal (Fig.2A).	Hyperintense signal intensity within the extra-capsular soft tissue around the joint on T2-weighted fat-saturated or STIR or other fluid sensitive sequence, which does not involve the tendons.
Grading	**For large and medium volume joints:** **0**-Absent or physiologic amount of intraarticular fluid, **1**-Mild: Amount of fluid mildly distending one or more recesses and / or involving the entire joint compartment, **2**-Moderate/severe: Amount of fluid moderately to markedly distending one or more joint recesses and / or involving the entire joint compartment. **No recording** of the extent of effusion/synovial thickening. More precise differentiation between 1 and 2 severity categories should be displayed on representative images of a future atlas. **For small volume joints:** **0**-Absent or trace amount of physiologic intraarticular fluid appearing as "pencil thin linear intra-articular hyperintense signal" **1**-More than trace amount of physiologic fluid (appearing as "pencil thin linear high signal") in one or more joint recesses and / or involving the entire joint compartment.	**(i) Extent of BME:** **For large and medium volume joints** Segmentation of articular surfaces into halves. One score is given to signal abnormality noted in each half of the articular surface (binary data, Fig. 2A): **0**-Absent; **1**-Present: maximum total score of 4. **(i) Extent of BME (Fig. 2A):** **For small volume joints:** **0**-Absent, **1**-Present on one side of the articular surface **2**-Present on both sides of the articular surface. **(ii) Intensity of BME (Fig. 2B):** **0**-Signal intensity less than that of an adjacent vessel (vein with slow flow) or fluid (joint fluid or fluid in the urinary bladder), **1**-Signal intensity equal to that of an adjacent vessel /body fluid.	**0**-Absent, **1**-Present.
Item Score	/2 OR /1 per joint: 2- for large and medium volume joints and 1- for small volume joints respectively **Total peripheral joint inflammation index:**	/4 + /1 per joint for the large and medium volume joints /2 + /1 per joint for the small volume joints. Sum of (a + b + c) for all assessed joints.	/1 per joint.

B1 Item	Axial joints: sacroiliac joint (SIJ) (d) Bone Marrow Edema (BME)	(e) Effusion/synovial thickening	(f) Capsulitis
Definition	Edema-like marrow signal changes within the sub-chondral bone of sacral and iliac sides of the SIJ on T2-weighted fat-saturated or STIR or other fluid sensitive sequence, that is not compatible with physiologic hematopoietic marrow signal [14]. Bone marrow signal in the centre of sacrum constitutes the reference normal signal (Fig. 2B).	Areas of hyperintense signal within the synovial and or cartilaginous portion of the SIJ which is equivalent to that of cerebrospinal fluid and more than that of a thin, regular line of physiologic high signal on T2-weighted fat-saturated or STIR or other fluid sensitive coronal oblique sequence (Fig. 2B).	Areas of hyperintense signal along the superior portion of the SIJ capsule on T2-weighted fat-saturated or STIR or other fluid sensitive coronal oblique sequence (Fig. 2B).
Grading	**0**-Absent **i) Presence of edema:** The coronal oblique slice with most extensive inflammation through the synovial and /cartilaginous portion(s) of the SIJ should be scored. Each SIJ is divided into 4 quadrants, upper iliac, lower iliac, upper sacral and lower sacral (Fig. 2B). **1**-Presence of edema-like marrow signal, 1 score given for each quadrant of the SIJ. **ii) Presence of intense edema:** An extra score for "intense" signal should be assigned to each SIJ on each slice if applicable. Hyperintense signal from slow flowing venous flow within presacral veins acts as a reference for assigning an "intense" reading score to a bone lesion (Fig. 2B). **1-**Presence of "intense" signal, given for up to 4 quadrants of each SIJ.	**0**-Absent **1**-Presense of effusion/synovial thickening anywhere within the synovial and/or cartilaginous portion of the SIJ on the slice with most extensive inflammation, given a score of 1 for each involved joint.	**0**-Absent **1**-Presense of capsulitis along the superior joint capsule on the slice with most extensive inflammation, given a score of 1 for each involved joint.
Item Score	/8 + /8 Total SI inflammation index	/2 d + e + f	/2

B2	Axial joints: spine			
Item	**(g) Bone marrow edema (BME)** – Disco vertebral Units (DVU)	**(h) Corner inflammatory lesion (CIL)**	**(i) BME and/or effusion/synovial thickening** – Craniovertebral junction joints: -Atlanto-dental,	**(j) BME and/or effusion/synovial thickening** – Facet Joints

Table 1 (*Continued*)

B2	Axial joints: spine			
			-Paired atlanto-occipital & -Paired atlanto-axial joints	
Definition	(g, h, i, j) Edema-like marrow signal changes within the vertebral body at or in continuation with the vertebral endplate or within the bones forming the joints of the craniovertebral junction or in the facets on T2-weighted fat-saturated or STIR or other fluid sensitive sequence, that is not compatible with physiologic hematopoietic marrow signal [15]; (i, j) and/or an area of hyperintense signal within the craniovertebral junction and facet joint space on fluid sensitive sequence. Bone marrow signal in the center of each vertebra constitutes the reference signal. If the entire vertebra has abnormal signal, the signal intensity that is closest to the physiologic level for the patient's age is used for reference. Disc lesions are not scored.			
Grading	**0-**Absent **i) Presence of edema:** Each DVU is divided into four quadrants: upper anterior endplate, upper posterior endplate, lower anterior endplate, and lower posterior endplate (Fig. 2C). **1-**Presence of edema-like marrow changes, given for up to 4 quadrants on the sagittal slice with most extensive inflammation. **ii) Presence of intense edema:** Hyperintense signal isointense to that of cerebrospinal fluid acts as a reference for assigning one "intense" score to a bone lesion (Fig. 1C). **1-**Presence of "intense" signal for bone marrow edema in any quadrant on the sagittal slice with most extensive inflammation.	**0-**Absent **1-**Presence of edema-like marrow changes within the corner of superior and inferior endplates of vertebrae. Each lesion gets a score of 1 on the sagittal slice with most extensive inflammation (Supplementary-Fig. 2).	**0-**Absent **1-**Presence of edema-like marrow changes and/or effusion/synovial thickening on any slice, given for up to 5 craniovertebral junction joints.	**0-**Absent **1-**Presence of BME and/or effusion/synovial thickening on any slice, given for up to 46 facet joints.
Item Score	/4 + /1 per DVU **Total spine inflammation index**	/1 per lesion	/5 g + h + i + j	/46

Abbreviations: STIR, Short Tau Inversion Recovery.

bursitis. CNO-like abnormality is defined as presence of edema-like marrow signal changes around the growth plate within the epimetaphysis or in the diaphysis of long bones on fluid sensitive sequences, that is not compatible with normal hematopoietic marrow signal (18). Locations for assessment included areas around the long bones (humerus, radius, ulna, femur, tibia, fibula).

(2) Axial joints

Findings were scored according to individual joints:

Sacroiliac joints

BME, effusion/synovial thickening, and capsulitis were selected as key pathologic items for assessing SIJs (Table 1B1). All items were scored by using single slice with most extensive inflammation. The items were scored on the same or different slices depending on which slice demonstrates the worst finding for each item. Different slices can be selected for the right and left side of joint based on the slice with worst finding. SIJ was graded only on T2-weighted fat-saturated or short tau inversion recovery (STIR) or other fluid sensitive sequences as shown in Fig. 3 B.

(d) BME

In order to maintain uniformity, BME was graded both for extent and intensity of edema as it was the approach for peripheral joints. For the scoring of extent of BME, each SIJ was divided into quadrants (first into halves and each half was further divided into upper and lower parts) as shown in Fig. 3 B1. Thus, a score of 1 was considered for the presence of BME into each quadrant with a maximum score of 8 (range 0−8). An additional score of 1 per quadrant was given for each joint for intense edema when the signal was equal or greater than the signal of presacral veins resulting in a range of 0−8 (Fig. 3 B1) score. So, a total score of 0-16 could be given for BME as shown in Fig. 3 B1.

(e) Effusion/synovial thickening

Effusion/synovial thickening was considered altogether as a single item. Each SIJ received a 0,1 score (binary data) for pathology, with a maximum score of 2. It was scored on a single coronal slice with most extensive inflammation, separately for the left and right joints as shown in Fig. 3 B2 (utilizing all slices of that joint to check for presence (score of 1) or absence (score of 0) of the item.

(f) Capsulitis

Capsulitis refers to inflammatory changes in the joint capsule, and is most commonly evident at the superior margin of the SIJ. Only the superior portion of the SIJ capsule is scored, with a 0,1 score (binary data). Once again, the single slice with the worst finding was used for scoring purposes and received a score of 1 per joint with a maximum score of 2 as shown in Fig. 3 B2.

Definitions and item scoring for (d) BME, (e) effusion/synovial thickening, and (f) capsulitis for sacroiliac joints (SIJs) are shown in Table 1B1.
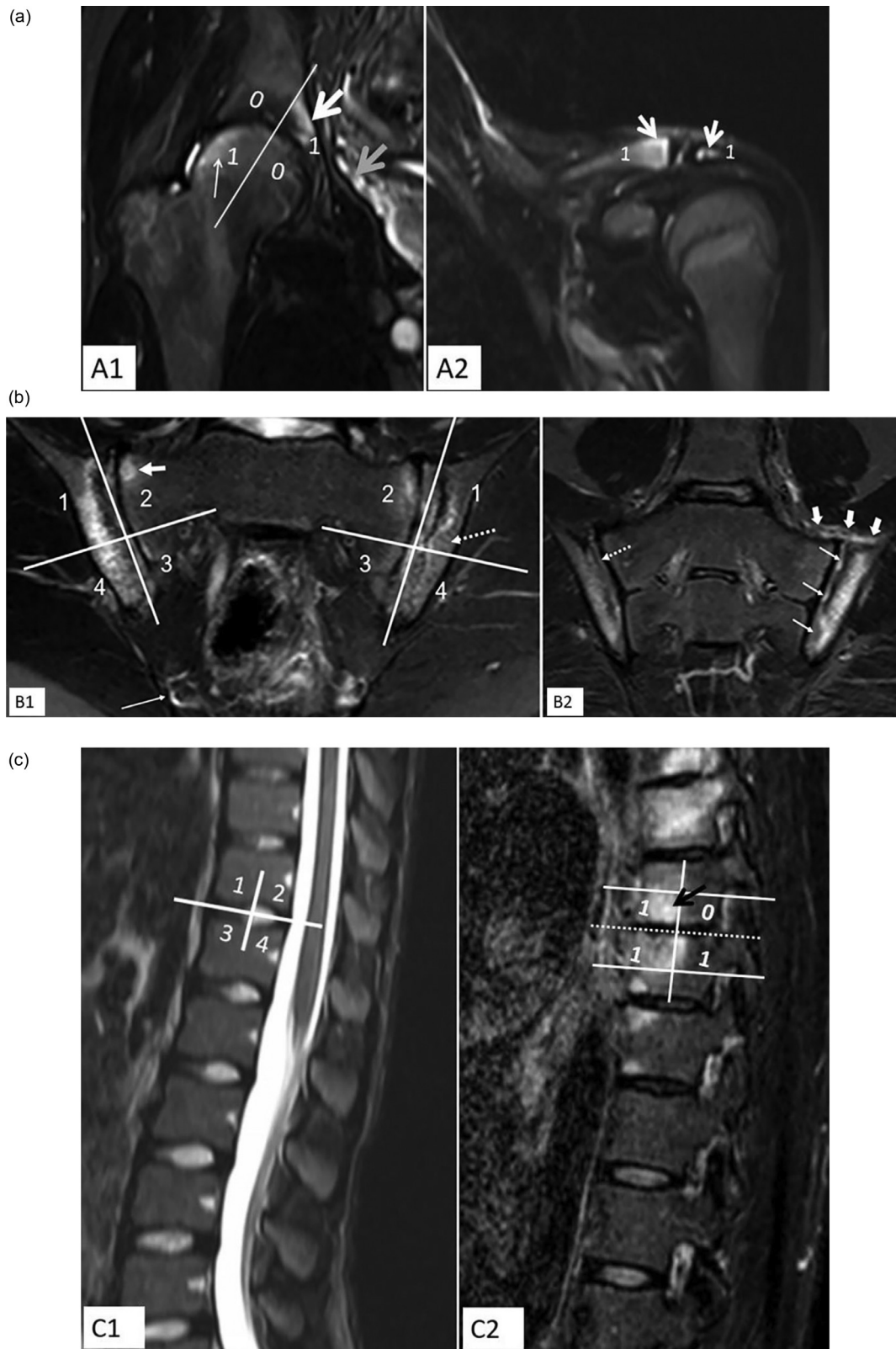
Figure 3. A. Examples of use of scoring system in peripheral (hip and acromioclavicular) joints. Grading of bone marrow edema (BME): **A1.** Extent and intensity of BME for large and medium joints: Coronal short tau inversion recovery (STIR) MR image of the right hip of a 15-year-old boy with enthesitis-related arthritis (ERA) shows segmentation of articular surfaces into halves. A score of 1 is given to each half for the presence of BME (thin arrow). An additional score of 1 is given per joint as the signal in the area of most severe abnormal signal (thick arrow) is similar to that of the signal of the adjacent vessels (gray arrow). In this case the final BME score would be 3 which equals extent (0+1+1+0)) + intensity (1). **A2.** Extent of BME for small joints: Coronal STIR MR image through the acromioclavicular joint in a 13-year-old boy with ERA shows a score of one for each articular surface. In this case the BME score would be 2 which equals extent (1+1) + intensity (0). **B.** Examples of use of scoring system in axial (sacroiliac) joints. **B1.** Grading of BME of sacroiliac joint (SIJ) in a 13-year-old boy with psoriatic arthritis presenting with axial involvement. Coronal STIR MR image of the SIJs shows segmentation of each joint into quadrants, right (R): upper iliac=1, upper sacral=2, lower sacral=3, lower iliac=4; left (L): upper iliac=1, upper sacral=2, lower sacral=3, lower iliac=4. An ill-defined hyperintense STIR signal is noted within the subchondral bone (dashed arrow,) of SIJ suggests BME. A focus of hyperintense signal (thick arrow) equals that signal of presacral vessels (thin arrow) considered to represent an

Table 2
Selection of anatomic sites for assessment of inflammation in juvenile idiopathic arthritis using whole body (WB)-MRI.

C	Entheses		
Item	(k) Bone marrow edema (BME)	(l) Perientheseal soft tissue high signal	(m) Tendon/ligament high signal
Definition	Edema-like marrow signal changes within the bone at and around the attachment of tendon/ ligament/muscle on T2-weighted fat-saturated or STIR or other fluid sensitive sequence, measured perpendicular to the long axis of entheses (Supplementary Fig. 3).	An area of abnormal increased signal intensity in the surrounding bursa/soft tissue apart from the concerned tendon/ ligament/ muscle attached at the entheses on fluid sensitive sequences (Supplementary Fig. 3).	Abnormal increased signal within the inserted structure: Intrasubstance increased signal at or within 1,2 cm of tendon / ligament / muscle attachment after excluding magic angle effect on fluid sensitive sequences Supplementary Fig. 3).
Grading	**(i) Extent of BME:** **0-**Absent, **1-**Mild: Confluent marrow hyperintense signal measures < 1 cm from the entheseal surface **2-**Moderate to severe: Confluent marrow hyperintense signal measures ≥ 1 cm from the entheseal surface. **(ii) Intensity of BME:** Graded on the worst area of signal abnormality per entheseal site, **0-**Signal intensity less than that of an adjacent vessel (vein with slow flow) or body fluid area (joint fluid/fluid in the urinary bladder), **1-**Signal equal to that of an adjacent vessel /fluid.	**0-**Absent **1-**Present	**0-**Absent **1-**Present
Item Score	/2 + /1 per entheseal site	/1 per entheseal site	/1 per entheseal site
	Total entheseal inflammation index	k + l + m	

Spine

In this scale, BME is the most important item for evaluating spine, although with different grading schemes for different types of spinal joints (Table 1B2). Additionally, an item named "joint effusion/synovial thickening" was used for assessing craniovertebral and facet joints, although its scoring was combined with BME for final scoring.

(g) BME−disco-vertebral units (DVUs)

The DVUs from C2/C3 to L5/S1 were graded for BME. After scanning the entire spine, all the abnormal disco-vertebral levels were selected for scoring. After selecting DVUs, a single sagittal slice representing the areas with most abnormal signal for each DVU level were selected for scoring. Similar to the SIJ evaluation, edema within the DVUs was assessed for "extent" and "intensity". Each DVU was divided into four parts as shown in Fig. 3 C1. Each part received a score of 1 for the presence of BME (Fig. 3 C2). An additional score of 1 was assigned for intense edema (Fig. 3 C2). So, the maximum score per DVU was 5. BME in the center of each vertebra constituted the reference normal signal. Edema was considered "intense" if the marrow high signal was equal or greater to that of the adjacent cerebrospinal fluid.

(h) Corner inflammatory lesion (CIL)

Usually it is not very easy to separate a CIL from a DVU lesion. It was suggested that as of now we should record and score both lesions. A CIL is defined as focal edema-like marrow signal changes at the corner of superior and inferior endplates of vertebrae (Data-in-brief Fig. 1) and should be scored if seen separately from adjacent DVU lesion. Each lesion gets a score of 1.

(i) BME and/or effusion/synovial thickening-Craniovertebral junction joints (atlanto-dental, atlanto-occipital, and atlanto-axial joints)

This includes atlanto-dental and paired lateral atlanto-occipital and atlanto-axial joints. Binary scoring was deemed sufficient for the five craniovertebral junction joints. Presence of BME and/or joint effusion/synovial thickening in any number of slices per joint received a score of 1.

(j) BME and/or effusion/synovial thickening-facet joints

Likewise, disease involvement in the 46 (23 paired joints) facet joints was scored as 0-absent or 1-present per joint, regardless of the number of slices involved per joint. Each of the items, BME and joint effusion/synovial thickening together, received a score of 1 per joint.

Definitions and item scoring for (g) BME - DVU, (h) CIL, BME and/ or effusion/synovial thickening-(i) craniovertebral junction (atlanto-dental, atlanto-occipital and atlanto-axial joints), and (j) facet joints are shown in Table 1B2.

(3) Entheses

For assessing entheses, "BME", "abnormal intra-tendinous/ligamentous high signal" and "perientheseal soft tissue inflammation" were considered key items.

(k) BME

For consistency, this item was also separately graded for the i) extent and ii) intensity of edema: (i) 0-2 score for "extent" of BME (0 = absent; 1 = present and < 1 cm depth; 2 = present and > 1 cm depth), and (ii) an additional 0,1 score (0 = absent; 1 = present) for presence of "intense" edema. The area with most severe signal abnormality per entheseal site was graded, regardless of number of slices (Data-in-brief Fig. 2).

(l) Perientheseal soft tissue abnormal signal

Binary scoring was adopted for this item. A score of 1 was given if a signal abnormality was noted in each involved perientheseal site, regardless of number of slices (Data-in-brief Fig. 2).

area of "intense edema" and receives an additional score of one if present in each quadrant for the right and left joints. In this case the final BME score would be 13 which encompasses extent (R=1+1+1; L=1+1+1+1) and intensity (R=1+1+0+1; L=1+1+0+0). **B2.** Effusion/synovial thickening and capsulitis in a 12-year-old boy with ERA. Coronal oblique STIR MR image of the SIJs shows hyperintense signal within the joint space (thin arrows) of the left joint which is equivalent to that of the cerebrospinal fluid and minimally wider than that of a thin, regular line of physiologic signal within the right joint space (dashed arrow). This suggests left-sided effusion/synovial thickening. Hyperintense signal is noted along the superior joint capsule of the left joint (thick arrows) in keeping with capsulitis. A 0-1 score (binary data) is used to grade both findings. In this case, the score for effusion/synovial thickening and capsulitis would be 1 for each. **C.** Examples of use of scoring system in axial (spine) joints. **C1**. Sagittal STIR MR image in a16-year-old girl with ERA shows segmentation of DVU. Each DVU is divided into four quadrants: 1. upper anterior endplate, 2. upper posterior endplate, 3. lower anterior endplate, 4. lower posterior endplate. BME in each quadrant is recorded on sagittal slice with the most severe finding. Bone marrow signal in the centre of each vertebra constitutes the reference normal signal. If the entire vertebra is abnormal, the signal intensity closest to that of physiologic level is used for reference. **C2.** Parasagittal STIR MR image in a 17-year-old boy with psoriatic arthritis with axial involvement demonstrates "extent" and "intensity" of BME. A score of 1 is given if marrow edema-like signal seen in any quadrant on the single slice with most severe findings for each selected level. Marrow edema is considered "intense" if the intensity of the area of hyperintense signal within the bone marrow (black arrow) is equal to the signal noted in the adjacent cerebrospinal fluid. A score of 1 is assigned if "intense" signal is noted in any quadrant of a chosen single slice for each selected level. Hence, the maximum score is 5 per DVU. In this case, the total BME score at the marked level would be 4 which encompasses extent (1+0+1+1) and intensity (1).

(m) Tendon/ligament abnormal signal

Binary scoring was adopted for this item. A score of 1 was given if an intrasubstance increased signal abnormality was noted in each involved tendon/ligament, regardless of number of slices (Data-in-brief Fig. 2).

Definitions and item scoring for (k) BME; (l) perientheseal soft tissue abnormal signal and (m) tendon/ligament abnormal signal are shown in Table 1B2. An example of the use of the scoring system in enthenses is shown in Data-in-brief Fig. 2.

MRI sequences and planes

Concepts and gradings for inflammation were based on considerations on a protocol that would have a fluid sensitive sequence, either a STIR or a fat-suppressed T2-weighted sequence. A T1-weighted sequence without contrast would not be needed for assessing inflammation as marrow edema may not be conspicuous on non-contrast T1-weighted sequence. The group felt that the use of intravenous contrast could be informative but not feasible in the screening setting considering the added scan time and potential long-term effects related to retention of contrast agent [19]. Recommendations for a preliminary core unenhanced WB-MRI protocol that evaluated multiple joints and enthenses affected in JIA on coronal STIR images with additional images for specific parts of the body is shown in Table 2, Data-in-brief Fig. 3.

Discussion

Our WB-MRI scoring system for JIA focused on the assessment of the inflammation in the joints and enthenses of the body. Future validation studies of this scale are required which may modify this initial iteration.

Periphyseal regions demonstrating CNO-like lesions [18] are not uncommon on WB-MRI due to overlap between imaging findings of enthesitis-related arthritis (ERA) subtype of JIA and CRMO. Nevertheless, the working group members voted to not include these regions in the scoring system at this initial iteration.

Rationale for excluded items

Based on the experience of the expert group, the low spatial resolution limitations of WB-MRI for assessing components of small joints make it more challenging for readers to interpret abnormalities in small joints. Hence, detailed evaluation of the small joints of hands and feet should be conducted by dedicated regional imaging if clinically warranted. The costovertebral and costotransverse joints were excluded from the scoring system based on consensus opinion, since the group felt that it is not usual that these joints are not covered on the sagittal plane of WB-MRI, and it is difficult to assess these joints on coronal planes. Their detail assessment requires a dedicated axial plane. Thickening of attached tendons/ligaments were not considered in the scoring system as our members felt that no existing definitions or standardized normal MRI values' data are currently available for tendon thickness in different locations.

For assessment of effusion/synovial thickening of peripheral joints, these joints were divided into three categories based on their size. The experts decided not to use normative measurements for grading purposes, instead to apply a gestalt pattern recognition criterion-based assessment upon comparison with reference images for grading (atlas under development).

Most members considered that diffusion-weighted and contrast-enhanced imaging were techniques that were not ready to be used as outcome measures for a WB-MRI scoring system at the present. Use of contrast in pediatric population will remain an issue of concern particularly given that the effects of retention of gadolinium within the brain, remains uncertain [20]. With the long scan times in WB-MRI, the post-injection delay varies widely across body parts resulting in differential enhancement of structures at varying times after injection, which may lead to incorrect interpretation of findings [21−22].

The total scan time of the proposed protocol is approximately 40−45 min per patient, depending of the body size, which is challenging for young patients. The group felt that dedicated coronal images of hands would add time for little benefit, as the field-of-view would be too large to visualize subtle changes. Instead, positioning the patients' hands in supine position on their thighs or buttocks enables acquisition of coronal views of hands without the need for additional sequences.

Conclusion

WB-MRI is a promising tool with great potential in determining the total inflammatory burden and assessing treatment response in JIA. Our structured consensus efforts within the OMERACT MRI in JIA working group have initiated the development of a pediatric WB-MRI scoring system for JIA. Iterative refinements to the scoring system are warranted in response to subsequent feasibility, reliability and responsiveness testing in upcoming studies.

CRediT authorship contribution statement

Jyoti Panwar: Visualization, Investigation, Writing − original draft, Writing − review & editing. **Mirkamal Tolend:** Investigation, Writing − review & editing. **Bernadette Redd:** Investigation, Writing − review & editing. **Hemalatha Srinivasalu:** Investigation, Writing − review & editing. **Robert A. Colbert:** Investigation, Writing − review & editing. **Jonathan Akikusa:** Investigation, Writing − review & editing. **Simone Appenzeller:** Investigation, Writing − review & editing. **John A. Carrino:** Investigation, Writing − review & editing. **Nele Herregods:** Investigation, Writing − review & editing. **Lennart Jans:** Investigation, Writing − review & editing. **Kerri Highmore:** Investigation, Writing − review & editing. **Thekla von Kalle:** Investigation, Writing − review & editing. **Eva Kirkhus:** Investigation, Writing − review & editing. **Dax G. Rumsey:** Investigation, Writing − review & editing. **Jacob L. Jaremko:** Investigation, Writing − review & editing. **Inarejos Emilio J. Clemente:** Investigation, Writing − review & editing. **Marion A. van Rossum:** Investigation, Writing − review & editing. **Jennifer Stimec:** Investigation, Writing − review & editing. **Shirley M. Tse:** Investigation, Writing − review & editing. **Marinka Twilt:** Investigation, Writing − review & editing. **Nikolay Tzaribachev:** Investigation, Writing − review & editing. **Iwona Sudol-Szopinska:** Investigation, Writing − review & editing. **Arthur B. Meyers:** Investigation, Writing − review & editing. **Andrea S. Doria:** Visualization, Investigation, Writing − original draft, Writing − review & editing.

References

1 Ravelli A, Martini A. Juvenile idiopathic arthritis. Lancet 2007;369(9563):767−78.

2 Breton S, Jousse-Joulin S, Finel E, et al. Imaging approaches for evaluating peripheral joint abnormalities in juvenile idiopathic arthritis. Semin Arthritis Rheum 2012;41(5):698–711.

3 Damasio MB, Malattia C, Tanturri de Horatio L, 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 multi centre reliability study. Pediatr Radiol 2012;42:1047–55.

4 Weckbach S. Whole-body MR imaging for patients with rheumatism. Eur J Radiol 2009;70(3):431–41.

5 Panwar J, Tolend M, Lim L, Tse SM, Doria AS, Laxer RM, Stimec J. Whole-body MRI quantification for assessment of bone lesions in CNO patients treated with pamidronate: a prevalence, reproducibility, and responsiveness study. J Rheumatol 2020 Sep 15:jrheum.200329Epub ahead of print. PMID: 32934131. doi: 10.3899/jrheum.200329.

6 Zhao Y, Sato TS, Nielsen SM, Beer M, Huang M, Iyer RS, McGuire M, Ngo AV, Otjen JP, Panwar J, Stimec J, Thapa M, Toma P, Taneja A, Gove NE, Ferguson PJ. Development of a scoring tool for chronic nonbacterial osteomyelitis magnetic resonance imaging and evaluation of its interrater reliability. J Rheumatol 2020 May 1;47 (5):739–47 Epub 2019 Oct 1. PMID: 31575701. doi: 10.3899/jrheum.190186.

7 Malattia C, Tolend M, Mazzoni M, Panwar J, Zlotnik M, Otobo T, Vidarsson L, Doria AS. Current status of MR imaging of juvenile idiopathic arthritis. Best Pract Res Clin Rheumatol 2020;34(6):101629 Epub 2020 Dec 3. PMID: 33281052. doi: 10.1016/j.berh.2020.101629.

8 Elessawy SS, Abdelsalam EM, Abdel Razek E, Tharwat S. Whole-body MRI for full assessment and characterization of diffuse inflammatory myopathy. Acta Radiol Open 2016;5(9):2058460116668216 Published 2016 Sep 21. doi: 10.1177/2058460116668216.

9 Poggenborg RP, Pedersen SJ, Eshed I, et al. Head-to-toe whole-body MRI in psoriatic arthritis, axial spondyloarthritis and healthy subjects: first steps towards global inflammation and damage scores of peripheral and axial joints. Rheumatology 2015;54(6):1039–49 (Oxford).

10 Axelsen MB, Eshed I, Duer-Jensen A, et al. Whole-body MRI assessment of disease activity and structural damage in rheumatoid arthritis: first step towards an MRI joint count. Rheumatology 2014;53(5):845–53 (Oxford).

11 Østergaard M, Eshed I, Althoff CE, et al. Whole-body magnetic resonance imaging in inflammatory arthritis: systematic literature review and first steps toward standardization and an OMERACT scoring system. J Rheumatol 2017;44:1699–705.

12 Panwar J, Patel H, Tolend M, Akikusa J, Herregods N, Highmore K, Inarejos Clemente EJ, Jans L, Jaremko JL, von Kalle T, Kirkhus E, Meyers AB, van Rossum MA, Rumsey DG, Stimec J, Tse SM, Twilt M, Tzaribachev N, Doria AS. Toward developing a semiquantitative whole body-MRI scoring for juvenile idiopathic arthritis: critical appraisal of the state of the art, challenges, and opportunities. Acad Radiol 2020 Mar 2S1076-6332(20)30050-7Epub ahead of print. PMID: 32139304. doi: 10.1016/j.acra.2020.01.022.

13 Guzman J, Burgos-Vargas R, Duarte-Salazar C, Gomez-Mora P. Reliability of the articular examination in children with juvenile rheumatoid arthritis: interobserver agreement and sources of disagreement. J Rheumatol 1995;22:2331–6.

14 Working Group website (https://urldefense.com/v3/__https://omeract.org/working-groups/jam/__;!!D0zGoin7BXfl!sI3suGssGykkpUMAG-yu8Qm8G2k0AZsIDf-WIAGfMJ6OoRq8sDm1V3svSXlFutxFhWIxd$)

15 Boers M, Beaton DE, Shea BJ, Maxwell LJ, Bartlett SJ, Bingham CO, et al. OMERACT filter 2.1: elaboration of the conceptual framework for outcome measurement in health intervention studies. J Rheumatol 2019;46(8):1021. Aug 1.

16 The OMERACT Handbook (2019): (https://urldefense.com/v3/__https://omeracthandbook.org/handbook__;!!D0zGoin7BXfl!sI3suGssGykkpUMAG-yu8Qm8G2-k0AZsIDfWIAGfMJ6OoRq8sDm1V3svSXlFut5_Nkvsb$)

17 Chauvin NA, Xiao R, Brandon TG, Biko DM, Francavilla M, Khrichenko D, Weiss PF. MRI of the sacroiliac joint in healthy children. AJR Am J Roentgenol 2019 Apr 11:1–7 Epub ahead of print. PMID: 30973768. doi: 10.2214/AJR.18.20708.

18 Fritz J, Tzaribatchev N, Claussen CD, Carrino JA, Horger MS. Chronic recurrent multifocal osteomyelitis: comparison of whole-body MR imaging with radiography and correlation with clinical and laboratory data. Radiology 2009;252(3):842–51.

19 Garcia J, Liu SZ, Louie AY. Biological effects of MRI contrast agents: gadolinium retention, potential mechanisms and a role for phosphorus. Philos Trans A Math Phys Eng Sci 2017;375(2107):20170180. doi: 10.1098/rsta.2017.0180.

20 Tedeschi E, Caranci F, Giordano F, et al. Gadolinium retention in the body: what we know and what we can do. Radiol Med 2017;122:589–600 (Torino).

21 Rieter J, de Horatio LT, Nusman CM, et al. The many shades of enhancement: timing of post-gadolinium images strongly influences the scoring of juvenile idiopathic arthritis wrist involvement on MRI. Pediatr Radiol 2016;46:1562–7.

22 Østergaard M, Klarlund M. Importance of timing of post-contrast MRI in rheumatoid arthritis: what happens during the first 60 minutes after IV gadolinium-DTPA? Ann Rheum Dis 2001;60:1050–4.