Development and reliability of a preliminary Foot Osteoarthritis Magnetic Resonance Imaging Score

Jill Halstead[1], **Carmen Martín-Hervás**[2,3], **Elizabeth MA Hensor**[1,4], **Dennis McGonagle**[1,4], **Anne-Maree Keenan**[4,5,7], **Anthony C Redmond**[#§,1,4,5,6], and **Philip G Conaghan**[#1,4,5,6]

[1]Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, UK

[2]Radiology Department, La Paz University Hospital, Autonomous University of Madrid, Spain

[3]Biomedical Research Networking Centre on Bioengineering, Biomaterials and Nanomedicine, Madrid, Spain

[4]Leeds NIHR Musculoskeletal Biomedical Research Unit, Leeds Teaching Hospitals Trust, Leeds, UK

[5]Arthritis Research UK Experimental Osteoarthritis Treatment Centre, Leeds, UK

[6]Arthritis Research UK Centre for Sports, Exercise and Osteoarthritis, Nottingham/Leeds, UK

[7]School of Healthcare, University of Leeds, UK

[#] These authors contributed equally to this work.

Abstract

Objective—Foot osteoarthritis (OA) is very common but under-investigated musculoskeletal condition and there is little consensus as to common MRI imaging features. The aim of this study was to develop a preliminary foot OA MRI score (FOAMRIS) and evaluate its reliability.

Methods—This preliminary semi-quantitative score included the hindfoot, midfoot and metatarsophalangeal joints. Joints were scored for joint space narrowing (JSN, 0-3), osteophytes (0-3), joint effusion-synovitis and bone cysts (present/absent). Erosions and bone marrow lesions (BMLs) were scored (0-3) and BMLs were evaluated adjacent to entheses and at sub-tendon sites (present/absent). Additionally, tenosynovitis was scored (0-3) and midfoot ligament pathology was scored (present/absent). Reliability was evaluated in 15 people with foot pain and MRI-detected OA using 3.0T MRI multi-sequence protocols and assessed using intraclass correlation coefficients (ICC) as an overall score and per anatomical site (see supplementary data).

Results—Intra-reader agreement (ICC) was generally good to excellent across the foot in joint features (JSN 0.94, osteophytes 0.94, effusion-synovitis 0.62 and cysts 0.93), bone features (BML 0.89, erosion 0.78, BML-entheses 0.79, BML sub-tendon 0.75) and soft-tissue features (tenosynovitis 0.90, ligaments 0.87). Inter-reader agreement was lower for joint features (JSN 0.60, osteophytes 0.41, effusion-synovitis 0.03) and cysts 0.65, bone features (BML 0.80, erosion

§Corresponding author and request for reprints: Prof Anthony Redmond, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, 2nd Floor, Chapel Allerton Hospital, Leeds, LS7 4SA, Tel. 0113 3924914, Fax 0113 3924991, a.redmond@leeds.ac.uk.

0.00, BML-entheses 0.49, BML sub-tendon -0.24) and soft-tissue features (tenosynovitis 0.48, ligaments 0.50).

Conclusion—This preliminary FOAMRIS demonstrated good intra-reader reliability and fair inter-reader reliability when assessing the total feature scores. Further development is required in cohorts with a range of pathologies and to assess the psychometric measurement properties.

Key Indexing Terms

osteoarthritis; foot; MRI; semi-quantitative score; reliability

Introduction

Osteoarthritis (OA) of foot is a common cause of pain and disability (1–3). Recent radiographic studies suggest OA is much more common in the foot than previously suspected (3–5). The prevalence was reported to be between 60.7% to 94.6% for the foot joints in 62 to 94 year olds (2). Magnetic resonance imaging (MRI) has been utilised in describing and defining knee OA pathology, however its use in foot OA is limited, possibly due to the complexity of foot anatomy and image acquisition. Furthermore, while semi-quantitative scores have been developed for knee, hip and hand (6–13), none exist for OA of the foot. The aim of this study was to develop a foot OA MRI score (FOAMRIS) for assessing pathological features of OA and soft-tissue features that may be commonly associated with foot pain.

Materials and Methods

Development of the FOAMRIS

Following a review of MRI scoring systems (8, 13–15), a consensus process was undertaken involving two musculoskeletal radiologists, two rheumatologists and three podiatrists. A preliminary scoring system was developed to identify and grade typical pathological features of OA in joints, bones, and soft-tissue features associated with foot pain.

The new system included 16 joints: first to fifth MTP joints and tarsometatarsal joints, navicular-medial-cuneiform, navicular-intermediate-cuneiform, navicular-lateral-cuneiform, talo-navicular, calcaneal-cuboid and subtalar. Twelve bones were included: first to fifth metatarsals (divided into the distal, central and proximal regions), lateral cuneiform, intermediate cuneiform, medial cuneiform, navicular, cuboid, calcaneus and talus. The inter-phalangeal joints and toes were not included in this assessment score as these are often not in the field of view in a foot and ankle MRI coil.

Tendon and ligaments of the foot were included: eight sites of tenosynovitis; tibialis anterior, extensor hallucis longus, extensor digitorum longus, peroneus brevis, peroneus longus, tibialis posterior, flexor hallucis longus and flexor digitorum longus. The Lisfranc ligament complex and inter-tarsal ligaments were included although not every ligament in the Lisfranc (midfoot) region was individually scored due to the large degree of anatomical variation (16). These sites were included due to the association of soft-tissues disorders in OA (17), which has been shown for Lanfranc injuries and tendon damage (18, 19).

Five sub-tendon sites of the foot (bone regions adjacent to overlying tendons) were also included: lateral calcaneus under long peroneal tendon; lateral cuboid under long peroneal tendon; medial calcaneus under posterior tibial tendon; medial navicular under posterior tibial tendon and medial cuneiform under anterior tibial tendon. These sub-tendon sites, where tendons wrap around the bones, have been described as "functional entheses" and are sites associated with pain in mechanical foot disorders (20–23). On MRI, these regions can be associated with abnormal signal in the tendon and at the adjacent bone of the ankle (23) and it is unclear whether this may be the case in the foot.

Enthesopathy has been shown to be somewhat associated with OA in the hands (13, 24–26). It is unclear as yet whether there may be an association in the foot, given the weight-bearing nature of the structures, therefore enthesopathy was scored at nine sites in the foot: the tibialis anterior tendon at the plantar distal medial base of the first metatarsal bone and plantar distal medial cuneiform bone; peroneus longus tendon at the plantar base of the first metatarsal bone and plantar distal base medial cuneiform bone; tibialis posterior tendon at the plantar insertion at the base of the 2nd, 3rd or 4th metatarsal bones, the plantar proximal medial cuneiform bone, the plantar medial of lateral cuneiform bone and plantar medial navicular bone, and finally the peroneus brevis tendon at the dorsal lateral base of the 5th metatarsal.

A set of MRI features were determined and semi-quantitative scores for each feature were then developed. The term bone marrow lesion (BML) was adopted in this system, rather than bone marrow oedema, as bone signal in OA may not be attributed solely to fluid (27). During the consensus process, it became apparent that the development of a cartilage score posed challenges due the small cross-sectional surfaces and complexity of the anatomy. As such, a pragmatic approach was taken and a joint space narrowing (JSN) definition was agreed. In order to provide a score that could be applied in the absence of contrast agent, we did not include multiple severity categories for scoring or differentiate between synovitis and effusion (previously adopted in rheumatoid arthritis of the foot and OA of the hand (12, 28)), but pragmatically scored for the presence or absence of joint effusion-synovitis. The final definitions of each MRI feature, anatomical locations and semi-quantitative score are summarised in Table 1.

Image acquisition

Fifteen participants were recruited as part of a larger study. In accordance with the Declaration of Helsinki, ethical approval was provided (Leeds West Ethics Committee 09/H1305/10). Participants were included if they reported foot pain on weight-bearing and the musculoskeletal radiologist judged there to be MRI features of OA, which were based on knee MRI and foot radiographic criteria in at least one foot joint (4, 29). Inclusion was based therefore on presence of osteophytes judged to be at least moderate in size (grade 2) or, where the osteophytes were graded "small" this was accompanied by JSN (partial to full thickness, grade 2) and subchondral BML with cysts.

Participants were scanned using a Siemens Magnetom Verio (3T) large-bore MRI scanner (Siemens Medical Solutions, USA). All scans were acquired using an eight-channel foot and ankle coil, with the foot placed perpendicular to the ankle and magnetic field (β0) and

centered over the navicular bone. The following protocol was employed: T2 weighted fat saturated sequence parameters were TR:3000-3600 ms, TE:69, flip angle: 155-160º, echo train length 8, 2mm slices and 0.4mm inter-slice gap, Matrix 256*256, FOV 150*150mm in three planes. Short tau inversion recovery sequence [STIR] parameters were TR:4500 ms, TE:31, NEX 2, TI 200, flip angle 150°, echo train length 11, 3mm slices and 0.6mm inter-slice gap, Matrix 320*256, FOV 150*150mm, in three planes. T1 weighted high resolution spin echo sequence parameters were TR:700 ms, TE:10, FS 3, flip angle: 90º, 1.2mm slices and 1.32mm inter-slice gap, Matrix 512*512, FOV 150*150mm, in the sagittal plane. Gradient recalled echo sequence parameters were TR:450, TE:2.5, flip angle 30°: echo train length 1, 3mm slices, 0.6 interslice gap, Matrix 336*448, FOV 250*250mm, in the sagittal plane.

FOAMRIS reliability

Anonymised scans were analyzed using OsiriX 64bit Version 5.6 (OsiriX Foundation, Geneva, Switzerland). All images were scored using the standardised score sheet (Supplementary Data 1) and the FOAMRIS system (see Table 1 and Figure 1 to 3). Intra-reader reliability was undertaken by an experienced musculoskeletal radiologist who read the same images twice in a random order more than one-week apart. Inter-reader reliability exercise was undertaken by second reader. Both readers undertook a consensus exercise together using five separate foot images prior to second reader scoring.

Features were scored for each joint, bone and soft-tissue site, with all sites grouped. Reliability scores were evaluated using descriptive statistics; percentage of exact agreement (PEA) and Chamberlain's percent positive agreement (PPA), which is the proportion of the total number of ratings made in a given category during the two readings (either intra- or inter-reader pairs) that were in agreement. Additionally intraclass correlation coefficients (ICC) were calculated using generalizability theory; the Brennan method was used to account for negative variance components (30). The individual joint or bone was considered the facet of differentiation. Joint or bones were considered to be nested within patients. Patient, occasion (for intra-reader reliability) and reader (for inter-reader reliability) were considered random facets of generalization. Occasionally a negative ICC was obtained; when this occurred we reported that the result was negative (indicating poor agreement) but did not report the actual value. ICC could not be calculated when all joints or bones scored 0.

The reliability results were evaluated according to the Cicchetti criteria as <0.4 poor, 0.4 to 0.59 fair, 0.6 to 0.74 good and 0.75 to 1 excellent (31). Analysis was undertaken using Stata 13.1 (StataCorp. 2013. Stata: 13.1, College Station, T Texas: StataCorp LP, USA) and G_STRING IV (a wrapper for urGENOVA, University of Iowa).

Results

The musculoskeletal radiologist read 61 sequential MRIs, of which 35 were classified as having foot OA and deemed eligible for the study. Fifteen participants' scans were chosen at random for the reliability study. This group were aged between 41 and 66 years (median 51 years, inter-quartile range 14), included 10 women, and had a median body mass index of

31.5 (inter-quartile range 8.2, range 23 to 40). Osteoarthritis was present in a single talo-navicular joint in five participants, in one to two joints in the tarsi in six participants, and in two joints (MTP and tarsal joints) in five participants. An experienced radiologist performed the full FOAMRIS in 30 minutes per foot and reported JSN in 12 participants (total 31 sites), osteophytes were present in all participants (total 77 sites), as was effusion-synovitis (total 182 sites), cysts were present in 13 participants (total 28 sites), BML was present in all participants (total 74 sites), erosion was present in five participants (total 10 sites), enthesopathy was present in seven participants (total nine sites), tenosynovitis was present in the entire group (total 47 sties) and ligament abnormalities were found in six participants (total seven sites).

The intra-reader reliability was summarised per imaging pathology (amalgamating anatomical locations see Table 2 and 3) and the range across the anatomical locations (Supplementary Data 2, Tables S1-7). It should be noted that ICCs represent a ratio of between-object variability to total variability and can therefore be low if there is little variation in scores between different joints/bones, which was an issue when assessing agreement in specific sites.

Combining all joints, the results showed excellent agreement for the presence of JSN (ICC total = 0.94, range across joints = 0.65 to 1) and osteophytes (ICC total = 0.94, range across joints = 0.00 to 1) although there was a low proportion of severe scores in this sample and for some individual sites the ICC was poor. There were very few JSN grade 3 scores and no scores for osteophytes grade 3 (the majority were grade 1-2) therefore the reliability in this category still remains to be determined; however, for grades 0-2 the category specific agreement was generally substantial (range 60%-100%). The presence of effusion-synovitis was the least reliably scored (ICC total 0.62, range across joints = negative to 1). Lower reliability in scoring effusion-synovitis was due to poor agreement over the absence of effusion-synovitis at the MTP joints. The repeatability for the scoring of presence of cysts was excellent when all joints were combined, although ICC was low for some individual sites (ICC total = 0.93, range across joints = 0.00 to 1).

The intra-reader reliability for combined sites was excellent for BML (ICC total = 0.89, range across bones = 0.49 to 1) and erosions (ICC total = 0.78, range across bones = 0.00 to 1). As was observed for the joints, in the bony features there was a relatively low prevalence of more severe scores. Scores for severity of BML suggest similar repeatability for the range of scores one to three, although only three bones across the sample scored a grade three. While the agreement results for erosions were not equal across the severity scale, the results showed a lower level of agreement for a score of one; however, only 20 erosion scores were assigned grade one and two erosion scores assigned grade two and none were assigned grade three, therefore the reliability in this category still remains to be determined.

The intra-reader reliability of bone related and soft-tissue results, the patterns of BML associated with tendon enthesopathy (ICC total = 0.79, range across the locations 0.44 to 1) and at the sub-tendon BML regions (ICC total = 0.75, range across the locations 0.00 to 1) were similar, with excellent agreement scores when all sites were combined. Reliability of scores for tenosynovitis was also excellent (ICC total = 0.90, range = 0.43 to 1). The

repeatability of scoring tenosynovitis was stable across scores ranging from 0 to 2. Score category 3 was not assigned during either of the repeated reads in this study; therefore, the repeatability in this category remains to be determined. The agreement scores for all ligament abnormality was excellent (ICC total = 0.87 range across the 2 sites = 0.65 to 0.74), with greater scores for the Lisfranc ligament.

The inter-reader reliability scores are summarised in Supplementary Data 3 (Tables S7-12) and as might be expected, the intra-reader scoring showed greater reliability than inter-reader. The results demonstrated good agreement for the presence of JSN (ICC total = 0.60, range across joints = negative to 1) and fair agreement for osteophytes (ICC total = 0.41, range across joints = 0.00 to 1). The inter-reader reliability scores for the presence of effusion-synovitis were poor across the joints of the foot (ICC total = 0.03, range across joints = negative to 0.13). The repeatability for the scoring of presence of cysts was good (ICC total = 0.65, range across joints = negative to 1).

The inter-reader reliability was excellent for sites of BML (ICC total = 0.80, range across bones = 0.00 to 1) but was poor for erosion scores in bones with erosions present (ICC total = 0.00, values for all bones 0.00 where calculable). There were several sites for which both scorers agreed on the absence of any erosions but ICCs could not be calculated if there were no scores above zero. The inter-reader reliability of bone-related and soft-tissue scores of BML associated with tendon enthesopathy was fair (ICC total = 0.49, range across the locations 0.00 to 1) but scores were less reliable at the sub-tendon BML regions (ICC total = 0.24, range across the locations negative to 0.65). Inter-reader reliability scores for tenosynovitis were fair (ICC total = 0.48, range = 0.00 to 0.61). The inter-reader reliability scores for all ligament abnormality were fair (ICC total = 0.50, range across the 2 sites = 0.00 to 0.18), with higher scores for the Lisfranc ligaments.

Discussion

Currently there are no MRI scoring systems for OA foot pathology, although a previous study has defined some MRI features in foot OA (32). This new scoring system was deliberately inclusive of not only "traditional" OA features, but also included features that may inform studies investigating the broader construct of foot pain.

In this study, intra-reader reliability of the total MRI features was shown to be generally excellent when assessed at a whole foot level, whilst inter-reader reliability was more variable. The best intra- and inter-reader reliability was seen for joint specific features (JSN, osteophytes and cysts), and compared well to scores such as those evaluating small joints in hand OA (12). The presence of joint effusion-synovitis showed worse intra- and inter-reader reliability and was lower due to poor agreement, particularly at the MTP joints, which may or not be considered a normal finding. The reliability scores may have been affected by the size of the joint, as joint effusion-synovitis scores have been shown previously to be more variable in small joints of the hands (12). In a later reliability study of joint effusion-synovitis in the hand, the agreement improved once an atlas was developed (13). In addition, administration of a contrast agent may have aided precision in estimating the volume of joint fluid, particularly in differentiating fluid from synovial hypertrophy. Further studies with

contrast administration may be needed to potentially to refine the scoring and better characterise OA-related pathology.

Bony features demonstrated excellent intra-reader agreement across the foot as a whole. Descriptively the erosion scores were highly reliable across nearly all sites, however this may have been influenced by the low number of lesions present. ICC values (where calculable) were variable, which may reflect both limitations in agreement over the presence of erosion, and limitations in the amount of 'true' variation between bones. The BML scores were also variable, and lower agreement was shown in the cuboid and the proximal metatarsals. Where patterns of BML were associated with the tendon enthesis, intra- and inter-reader reliability was good, but at the sub-tendon region, reliability was lower. No reliability studies of these MRI features have been previously reported, and there is likely to be difficulty in scoring these regions where planar anatomy is subject to partial volume artefact; in these regions an atlas would be beneficial.

The intra- and inter-reader reliability of scoring of soft-tissue features was similar across ligament abnormalities and tenosynovitis. Similar levels of agreement have been reported for scores of hand tenosynovitis in rheumatoid arthritis (33) and hand OA (12). A limited number of ligaments of the midfoot were included in this score; which have been previously well described (16, 34). Other foot ligaments were not included due to potential issues with poor visualisation and requirement for specialist views and sequences e.g. calcaneo-cuboid and calcaneo-navicular ligaments (35).

The results of this study should be considered in light of the following limitations. The sample for this preliminary study included a group with relatively mild structural OA, and more severe damage was limited to few joint regions. In addition, the definition of OA on MRI as applied in this study, while based on consensus approaches developed for other joints, requires further work and validation and this has implications for the results presented. Definitions of the individual features are difficult due to the variation in presentation in the various anatomical sites and the technical aspects of acquiring MRIs. For example we did not use contrast enhanced imaging in this study and so have not differentiated between synovitis and effusion. A detailed definition of osteophyte grading was not provided in the this score, given the widely varying presentation of peri-articular bone change in sites such as the 1st MTP joint versus the small joints of te hindfoot or midfoot. Future work is required to refine the FOAMRIS approach and explore validity in larger and more diverse samples. In addition, eight participants were obese (30 BMI), which may influence the frequency of the tendon and ligament pathology, as greater occurrence has been shown in obese people at the ankle (36).

The foot poses unique challenges when using MRI due to the complexity of the anatomy and inherent variability of shape and size. This manifests as problems with coil positioning, homogenous fat saturation, imaging wrap and magic-angle effect (37). In this study a foot and ankle coil was used, which was beneficial for maintaining a consistent position within the magnet; however this can be limited with larger and longer feet. Using a larger coil may allow for imaging of the entire foot, although the positioning might be difficult due to flexibility and foot type. In future studies it may be appropriate to reposition the target the

hindfoot, midfoot and forefoot, although this will increase acquisition times and may not be desirable.

The issue of how many planes and sequences to acquire is a complex one. In this study, both T2 weighted water-sensitive and STIR sequences in three planes were included to account for possible failure of the fat saturation. T1 weighted sequences included high-resolution spin echo and gradient recalled echo in a single plane, which may have affected the scoring of erosions and osteophytosis. In practice, where acquisition time is of primary importance, a T2 fat saturated sequence may suffice.

A minimum of two planes for each T1 weighted sequence could improve scoring however, defining the optimum plane for each foot joint requires further work and a 3D sequence may provide a compromise. Gradient recalled echo sequences are sensitive in delineating subchondral cysts and were helpful in the verification in this study. These sequences however, are insensitive to diffuse marrow abnormalities, because of trabecular magnetic susceptibility and will not show the full extent of these lesions, so in this study spin echo sequences were also employed for better BML detection (8). Further consensus regarding sequence choice is recommended.

Across most scores, inter-reader reliability scores were lower than intra-reader. We have identified training (the second reader was less experienced) and case definitions as likely contributors to these findings. Improved description of certain scoring features, accompanied by an atlas would be a natural next step as this process has improved inter-reader reliability in other MRI scores (13).

Finally it is recognised that ICCs can be affected by the degree of 'true' variability in the sample, which is this relatively mild group was limited for some features. Further validation in more diverse samples should give a more accurate assessment of inter- and intra-reader reliability.

In summary, we have proposed a set of definitions and scoring criteria for a semi-quantitative MRI investigation of multiple foot pathology; FOAMRIS. This preliminary scoring system generally showed acceptable reliability for a broad range of pathologies with the exception of effusion-synovitis, and for some features at anatomical sites where visualisation may be particularly influenced by acquisition plane. Iterative development is now needed, and will include application in other cohorts, expert consensus on acquisition protocol, use of contrast and the development of an atlas to aid scoring.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

J Halstead PhD, Visiting Research Fellow, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, UK and Principal Podiatrist Salford Royal Hospital (NHS) Foundation Trust, Manchester, UK.

C Martín-Hervás PhD, MD, Consultant Radiologist, Department of Musculoskeletal Radiology, La Paz University Hospital, Spain. Associate Professor of Radiology, School of Medicine, Autonomous University of Madrid, Spain and Biomedical Research Networking Centre on Bioengineering, Biomaterials and Nanomedicine, Madrid, Spain

EMA Hensor PhD, Biostatistician, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, UK.

D McGonagle PhD, Professor of Investigative Rheumatology, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, UK.

AM Keenan PhD Professor of Allied Health Research, School of Health Care, University of Leeds, UK.

AC Redmond PhD, Professor of Clinical Biomechanics, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, UK.

PG Conaghan PhD, Professor of Musculoskeletal Medicine, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, UK.

References

1. Roddy E, Zhang W, Doherty M. Prevalence and Associations of Hallux Valgus in a Primary Care Population. Arthritis Rheum. 2008; 59:857–862. [PubMed: 18512715]

2. Menz HB, Munteanu SE, Landorf KB, Zammit GV, Cicuttini FM. Radiographic evaluation of foot osteoarthritis: sensitivity of radiographic variables and relationship to symptoms. Osteoarthritis Cartilage. 2009; 17:298–303. [PubMed: 18789728]

3. Roddy E, Thomas MJ, Marshall M, Rathod T, Myers H, Menz HB, et al. The population prevalence of symptomatic radiographic foot osteoarthritis in community-dwelling older adults: cross-sectional findings from the Clinical Assessment Study of the Foot. Ann Rheum Dis. 2015; 74:156–163. [PubMed: 24255544]

4. Menz HB, Munteanu SE, Landorf KB, Zammit GV, Cicuttini FM. Radiographic classification of osteoarthritis in commonly affected joints of the foot. Osteoarthritis Cartilage. 2007; 15:1333–1338. [PubMed: 17625925]

5. Van Saase J, Van Romunde L, Cats A, Vandenbroucke J, Valkenburg H. Epidemiology of osteoarthritis: Zoetermeer survey. Comparison of radiological osteoarthritis in a Dutch population with that in 10 other populations. Ann Rheum Dis. 1989; 48:271–280. [PubMed: 2712610]

6. Peterfy CG, Guermazi A, Zaim S, Tirman PFJ, Miaux Y, White D, et al. Whole-Organ Magnetic Resonance Imaging Score (WORMS) of the knee in osteoarthritis. Osteoarthritis Cartilage. 2004; 12:177–190. [PubMed: 14972335]

7. Hunter DJ, Lo GH, Gale D, Grainger AJ, Guermazi A, Conaghan PG. The reliability of a new scoring system for knee osteoarthritis MRI and the validity of bone marrow lesion assessment:

BLOKS (Boston Leeds Osteoarthritis Knee Score). Ann Rheum Dis. 2008; 67:206–211. [PubMed: 17472995]

8. Hunter DJ, Guermazi A, Lo GH, Grainger AJ, Conaghan PG, Boudreau RM, et al. Evolution of semi-quantitative whole joint assessment of knee OA: MOAKS (MRI Osteoarthritis Knee Score). Osteoarthritis Cartilage. 2011; 19:990–1002. [PubMed: 21645627]

9. Kornaat P, Ceulemans R, Kroon H, Riyazi N, Kloppenburg M, Carter W, et al. MRI assessment of knee osteoarthritis: Knee Osteoarthritis Scoring System (KOSS)-inter-observer and intra-observer reproducibility of a compartment-based scoring system. Skeletal Radiol. 2005; 34:95–102. [PubMed: 15480649]

10. 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; 19:946–962. [PubMed: 21550411]

11. Lee S, Nardo L, Kumar D, Wyatt CR, Souza RB, Lynch J, et al. Scoring hip osteoarthritis with MRI (SHOMRI): a whole joint osteoarthritis evaluation system. J Magn Resonan Imaging. 2015; 41:1549–1557.

12. Haugen IK, Lillegraven S, Slatkowsky-Christensen B, Haavardsholm EA, Sesseng S, Kvien TK, et al. Hand osteoarthritis and MRI: development and first validation step of the proposed Oslo Hand Osteoarthritis MRI score. Ann Rheum Dis. 2011; 70:1033–1038. [PubMed: 21436160]

13. Haugen IK, Østergaard M, Eshed I, McQueen FM, Bird P, Gandjbakhch F, et al. Iterative development and reliability of the OMERACT hand osteoarthritis MRI scoring system. J Rheumatol. 2014; 41:386–391. [PubMed: 24293570]

14. Østergaard M, McQueen F, Wiell C, Bird P, BÃ¸yesen P, Ejbjerg B, et al. The OMERACT psoriatic arthritis magnetic resonance imaging scoring system (PsAMRIS): definitions of key pathologies, suggested MRI sequences, and preliminary scoring system for PsA Hands. Journal Rheumatol. 2009; 36:1816–1824.

15. Østergaard M, Edmonds J, McQueen F, Peterfy C, Lassere M, Ejbjerg B, et al. An introduction to the EULAR-OMERACT rheumatoid arthritis MRI reference image atlas. Ann Rheum Disease. 2005; 64:i3–i7. [PubMed: 15647420]

16. Castro M, Melão L, Canella C, Weber M, Negrão P, Trudell D, et al. Lisfranc joint ligamentous complex: MRI with anatomic correlation in cadavers. AJR Am J Roentgenol. 2010; 195:W447–W455. [PubMed: 21098178]

17. Øiestad BE, Engebretsen L, Storheim K, Risberg MA. Knee osteoarthritis after anterior cruciate ligament injury a systematic review. Am J Sports Med. 2009; 37:1434–1443. [PubMed: 19567666]

18. Hardcastle PH, Reschauer R, Kutscha-Lissberg E, Schoffmann W. Injuries to the tarsometatarsal joint. Incidence, classification and treatment. J Bone Joint Surg Bri. 1982; 64:349–356.

19. Bluman EM, Title CI, Myerson MS. Posterior tibial tendon rupture: a refined classification system. Foot Ankle Clin. 2007; 12:233–249. [PubMed: 17561198]

20. Benjamin M, McGonagle D. The anatomical basis for disease localisation in seronegative spondyloarthropathy at entheses and related sites. J Anat. 2001; 199:503–526. [PubMed: 11760883]

21. O'Donnell P, Saifuddin A. Cuboid oedema due to peroneus longus tendinopathy: a report of four cases. Skeletal Radiol. 2005; 34:381–388. [PubMed: 15937710]

22. Lo LD, Schweitzer ME, Fan JK, Wapner KL, Hecht PJ. MR imaging findings of entrapment of the flexor hallucis longus tendon. AJR Am J Roentgenol. 2001; 176:1145–1148. [PubMed: 11312169]

23. Morrison WB, Carrino JA, Schweitzer ME, Sanders TG, Raiken DP, Johnson CE. Subtendinous bone marrow edema patterns on MR images of the ankle: association with symptoms and tendinopathy. AJR Am J Roentgenol. 2001; 176:1149–1154. [PubMed: 11312170]

24. McGonagle D, Tan AL, Carey J, Benjamin M. The anatomical basis for a novel classification of osteoarthritis and allied disorders. J Anat. 2010; 216:279–291. [PubMed: 20070426]

25. Tan AL, Grainger AJ, Tanner SF, Shelley DM, Pease C, Emery P, et al. High-resolution magnetic resonance imaging for the assessment of hand osteoarthritis. Arthritis Rheum. 2005; 52:2355–2365. [PubMed: 16052535]

26. Kortekaas MC, Kwok W-Y, Reijnierse M, Wolterbeek R, Bøyesen P, van der Heijde D, et al. Magnetic resonance imaging in hand osteoarthritis: intraobserver reliability and criterion validity for clinical and structural characteristics. J Rheumatol. 2015; 42:1224–1230. [PubMed: 25979715]

27. Zanetti M, Bruder E, Romero J, Hodler J. Bone marrow edema pattern in osteoarthritic knees: correlation between MR imaging and histologic findings. Radiology. 2000; 215:835–840. [PubMed: 10831707]

28. Baan H, Bezooijen R, Avenarius JKA, Dubbeldam R, Drossaers-Bakker WK, van de Laar MAFJ. Magnetic resonance imaging of the rheumatic foot according to the RAMRIS system is reliable. J Rheumatol. 2011; 38:1003–1008. [PubMed: 21362763]

29. Hunter DJ, Arden N, Conaghan PG, Eckstein F, Gold G, Grainger A, et al. Definition of osteoarthritis on MRI: results of a Delphi exercise. Osteoarthritis Cartilage. 2011; 19:963–969. [PubMed: 21620986]

30. Brennan, RL. Generalizability Theory. New York: Springer-Verlag; 2001.

31. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess. 1994; 6:284–290.

32. Halstead J, Bergin D, Keenan A-M, Madden J, McGonagle D. Ligament and bone pathologic abnormalities more frequent in neuropathic joint disease in comparison with degenerative arthritis of the foot and ankle: implications for understanding rapidly progressive joint degeneration. Arthritis Rheum. 2010; 62:2353–2358. [PubMed: 20506318]

33. Haavardsholm EA, Østergaard M, Ejbjerg BJ, Kvan NP, Kvien TK. Introduction of a novel magnetic resonance imaging tenosynovitis score for rheumatoid arthritis: reliability in a multireader longitudinal study. Ann Rheum Dis. 2007; 66:1216–1220. [PubMed: 17392347]

34. Raikin SM, Elias I, Dheer S, Besser MP, Morrison WB, Zoga AC. Prediction of midfoot instability in the subtle Lisfranc injury. J Bone Joint Surg AM. 2009; 91:892–899. [PubMed: 19339574]

35. Melão L, Canella C, Weber M, Negrão P, Trudell D, Resnick D. Ligaments of the Transverse Tarsal Joint Complex: MRI-Anatomic Correlation in Cadavers. AJR Am J of Roentgenol. 2009; 193:662–671. [PubMed: 19696279]

36. Galli MM, Protzman NM, Mandelker EM, Malhotra A, Schwartz E, Brigido SA. Comparing Tendinous and Ligamentous Ankle Pathology in Atraumatic Overweight and Nonoverweight Patients A Comprehensive MRI Review. Foot Ankle Spec. 7:449–456. [PubMed: 25005703]

37. Rosenberg ZS, Bencardino J, Mellado JM. Normal variants and pitfalls in magnetic resonance imaging of the ankle and foot. Top Magn Reson Imaging. 1998; 9:262–272. [PubMed: 9773967]
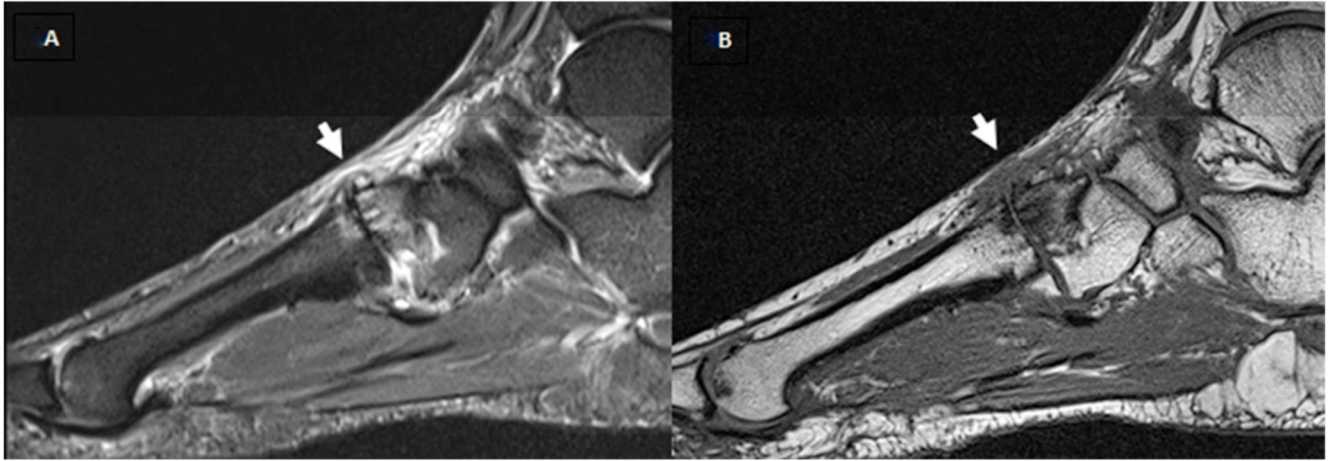
Figure 1.
A and B show T1w and fat saturated T2w sagittal images. Short white arrow shows the second tarsometatarsal joint, scored as a grade 2 (partial thickness or focal loss) for joint space narrowing (JSN), grade 2 (moderate) osteophyte, the presence of effusion, the presence of cysts, grade 2 (34%-66%) bone marrow lesion (BML) at the intermediate cuneiform and grade 1 (1%-33%) in the base of the second metatarsal. The second metatarsophalangeal joint was scored as a grade 2 for JSN, grade 2 osteophytes, in addition to grade 1 BML at the head of the second metatarsal.
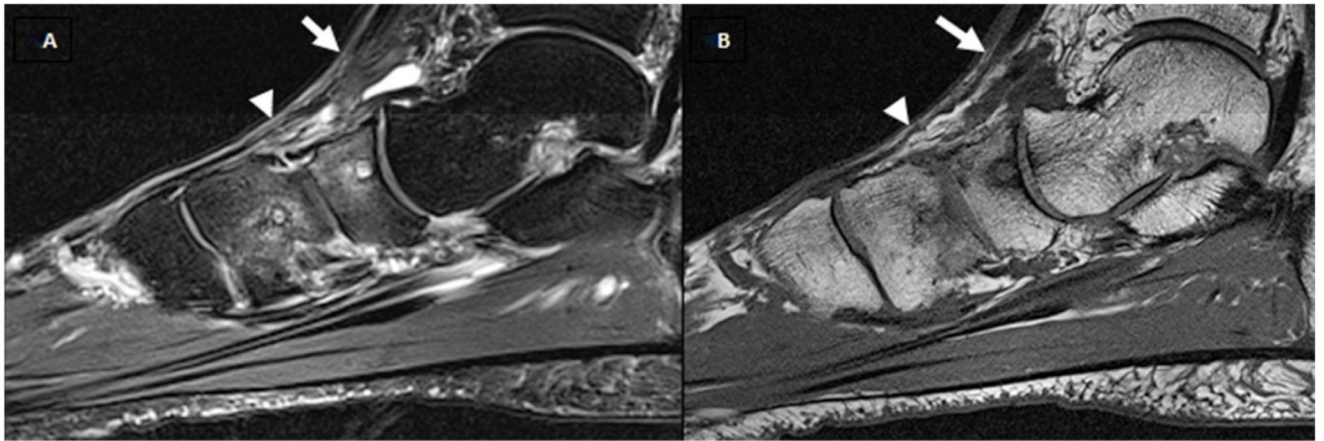
Figure 2.
A and B show T1w and fat saturated T2w sagittal images. White arrow shows the talo-navicular joint was scored as a grade 0 (normal joint space and signal) joint space narrowing (JSN), grade 2 (moderate) osteophyte, the presence of effusion, the presence of cysts, navicular bone were scored as grade 2 (34%-66%) bone marrow lesion (BML).
White arrow head shows the navicular-medial cuneiform was scored as grade 1 (increased signal in the joint space) JSN, grade 1 (mild) osteophytes, the presence of effusion, the presence of cysts, and medial cuneiform bone were scored as grade 2 (34%-66%) bone marrow lesion (BML).

Figure 3.
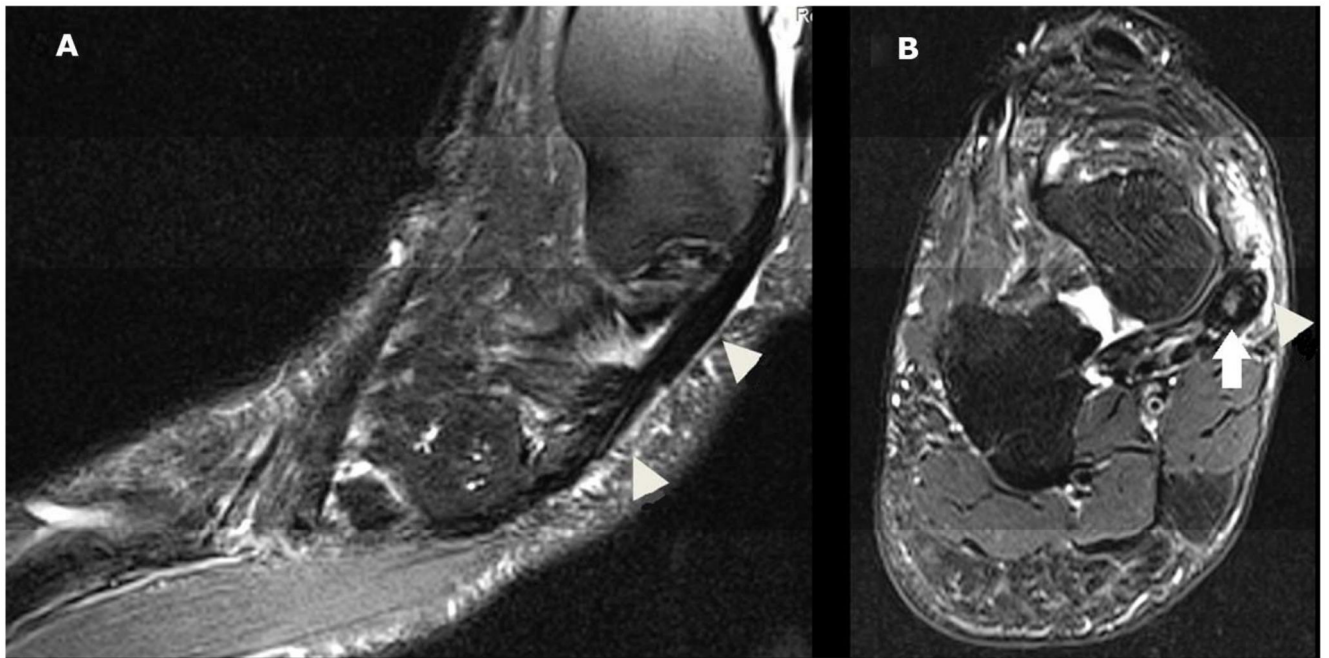A and B show fat saturated T2w images in the sagittal and axial plane respectively. Tenosynovitis of tibialis posterior defined high signal on T2-weighted in a region of the tendon (white arrow) and the enclosing tendon sheath (white arrow head) in two planes was shown. This was scored as grade 2 (>2 and <5 mm peritendinous effusion and/or thickening and high intra-tendinous signal intensity).

Table 1

Definitions of each MRI feature and the related semi-quantitative scores

MRI Feature and Anatomical Location	Definition	Score
Joint space narrowing (JSN) All joints of the hindfoot, tarsus, midfoot and metatarsophalangeal joints.	Increased signal in T2-weighted sequences (fat suppressed or inversion recovery sequences) and or loss of joint space as a partial or complete loss on T1-weighted images and or gradient echo sequence. Visible in two planes.	JSN was scored as: 0 to 3: 0 = Normal thickness and signal1 = Increased signal 2 = Partial-thickness focal loss 3 = Full thickness loss of joint space (>=75% of the region).
Osteophytes All joints of the hindfoot, tarsus, midfoot and metatarsophalangeal joints.	Abnormal bone formation in the periarticular region on T1-weighted images.	Osteophytes were scored as 0 to 3:0 = None, 1 = Mild, 2 = Moderate, 3 = Large.
Effusion-synovitis All joints of the hindfoot, tarsus, midfoot and metatarsophalangeal joints	The presence of increased intra articular fluid, demonstrated as high signal intensity on T2-weighted sequences (fat suppressed or inversion recovery sequences). Visible in two planes: coronal and sagittal.	Effusion-synovitis was scored as 0 to 1: 0 = Absent and 1 = Present
Subchondral cyst All joints of the hindfoot, tarsus, midfoot and metatarsophalangeal joints	A sharply marginated subchondral bone lesion that showed increased signal intensity on T2-weighted images (fat suppressed or inversion recovery sequences). Visible in two planes, without a cortical break.	Cysts were scored as 0 to1: 0 = Absent and 1 = Present.
Bone marrow lesion (BML) All bones of the hindfoot, tarsus, midfoot and metatarsals.	An area of ill delineated signal within the trabecular bone that shows decreased signal intensity on T1-weighted images and increased signal intensity onT2-weighted images (fat suppressed or inversion recovery sequences). Visible in at least two planes.	BML was scored as 0 to 3, according to the proportion of bone with abnormal signal: 0 = None, 1 = 1%–33%, 2 = 34%–66%, 3 = 67%–100%. Except in the long bones of the metatarsals, where the BML was scored in three regions per bone: (i) At the proximal joint, the base (up to the epiphysis and metaphysis) was included in the tarsometatarsal joint and scored 0-3. (ii) At the central region, metatarsal shaft (diaphysis) was divided into proximal, central and distal in one third increments (33%) on a bone level and scored 0-3. (iii) At the distal region the head (to the epiphysis and metaphysis) was included in the metatarsophalangeal joint and scored 0-3.
Bone erosion All bones of the hindfoot, tarsus, midfoot and metatarsals.	A bone defect in the cortical and juxta-cortical region, with sharp margins visible on T1-weighted images and with a loss of normal low signal intensity of cortical bone and loss of normal high signal intensity of marrow fat. Visible in two planes with a cortical break seen in at least one plane.	Erosions were scored from 0 to 3: according to the volume of the erosion as a proportion of the joint margin. 0 = No erosion; 1 = 1%–33%, 2 = 34%–66%, 3 = 67%–100%
Enthesopathy Locations at the insertion of the tendons: posterior tibial, anterior tibial, flexor digitorum, flexor hallucis, extensor digitorum, extensor hallucis, peroneus longus and peroneus brevis. Location at sites at the attachments of the Lisfranc and intertarsal ligament complex.	A BML pattern, where altered signal intensity within the bone was adjacent to insertions of anatomically defined ligaments and/or tendons of the foot. Visible in at least two planes.	Enthesopathy scored as 0 to 1: 0 = Absent and 1 = Present.
Sub-tendon BML (functional enthesopathy). All bones of the hindfoot, tarsus, midfoot and metatarsals adjacent to the course of a tendon at the medial (posterior tibial, and anterior tibial tendons), lateral (peroneus longus and peroneus brevis), plantar (flexor digitorum and flexor hallucis longus tendons) and dorsal (extensor digitorum and extensor hallucis longus tendons) regions.	A BML pattern where increased signal intensity within the bone was adjacent to the course of a tendon and away from an articular surface. Shown as hyperintensity on T2 weighted sequences and decreased signal intensity on T1-weighted images. Visible in at least two planes.	Sub-tendon BML was scored as 0 to 1: 0 = Absent and 1 = Present.

MRI Feature and Anatomical Location	Definition	Score
Tenosynovitis Tendon locations: Posterior tibial, anterior tibial, flexor digitorum, flexor hallucis, extensor digitorum, extensor hallucis, peroneus longus and peroneus brevis	Decreased signal intensity on T1-weighted images and increased signal on T2-weighted (fat supressed or inversion recovery sequences) in a region of the tendon with an enclosing tendon sheath. Visible in at least two planes.	Tenosynovitis was scored 0 to 3: 0 = Normal, 1 = <2 mm peritendinous effusion. 2 = > 2 and <5 mm peritendinous effusion and/or thickening and high intra-tendinous signal intensity on T2-weighted sequences. 3 = > 5 mm peritendinous effusion and/or higher thickening and high intra-tendinous signal intensity.
Ligament abnormality Ligament locations: Lisfranc and inter-tarsal ligament complex.	Thickening and high signal intensity seen on T2-weighted images (fat supressed or inversion recovery sequences) with or without disruption. Visible in at least two planes: axial and coronal.	Ligament abnormality was scored 0 to 1: 0 = Absent and 1 = Present.

*
Joint space narrowing (JSN), Bone marrow lesion (BML)

Table 2

Intra-rater repeatability results of scoring all joints and bones for joint space narrowing, osteophytes, effusion, cysts, bone erosions and marrow lesions.

Statistic	JSN 0-3	Osteophytes 0-3	Effusion-Synovitis 0-1	Cysts 0-1	Erosion 0-3	BML 0-3
PEA	96% (238/248)	93% (231/248)	81% (202/248)	98% (242/248)	97% (242/250)	91% (225/248)
PEA range	80%-100%	80%-100%	50%-100%	71%-100%	86%-100%	73%-100%
% ± 1 category	100% (248/248)	100% (248/248)	n/a	n/a	100% (250)	>99% (247/248)
PPA score=0	98% (422/429)	96% (322/335)	62% (74/120)	99% (432/438)	98% (470/478)	95% (338/354)
PPA score=1	85% (44/52)	86% (100/117)	88% (330/376)	90% (52/58)	60% (12/20)	79% (88/111)
PPA score=2	77% (10/13)	91% (40/44)	n/a	n/a	100% (2/2)	79% (22/28)
PPA score=3	0% (0/2)	- (none scored)	n/a	n/a	- (none scored)	67% (2/3)
ICC	0.90	0.90	0.46	0.87	0.66	0.83
ICC range $	0.65-1	0.00-1	0.00-1	0.00-1	0.00-1	0.49-1

Joint space narrowing=JSN; bone marrow lesion=BML; ICC=Intraclass correlation coefficient from generalizability theory; PEA=Percent exact agreement; PPA=Percent positive agreement.

 PEA per location (see supplementary data tables).

$ ICC per location (see supplementary data tables).

Table 3

Intra-rater repeatability results for scoring sites of enthesopathy, sub-tendon bone marrow lesion (BML), tenosynovitis and ligament abnormalities.

Statistic	Enthesopathy 0-1	Sub-tendon BML 0-1	Tenosynovitis 0-3	Ligament 0-1
PEA	96% (130/135)	91% (68/75)	88% (106/120)	90% (27/30)
PEA range	80%-100%	73%-100%	73%-100%	87%-93%
PEA ± 1 category	n/a	n/a	100% (120/120)	n/a
PPA score=0	98% (248/253)	94% (120/127)	92% (134/146)	100% (30/30)
PPA score=1	71% (12/17)	70% (16/23)	83% (68/82)	93% (42/45)
PPA score=2	n/a	n/a	83% (10/12)	n/a
PPA score=3	n/a	n/a	- (none scored)	n/a
ICC	0.66	0.60	0.83	0.77
ICC range $	0.44-1	0.00-1	0.43-1	0.65-0.74

Bone marrow lesion=BML; ICC=Intraclass correlation coefficient from generalizability theory; PEA=Percent exact agreement; PPA=Percent positive agreement.

PEA per location (see supplementary data tables),

$ ICC per location (see supplementary data tables).