Seminars in Arthritis and Rheumatism 51 (2021) 929–932

Contents lists available at ScienceDirect

Seminars in Arthritis and Rheumatism

journal homepage: www.elsevier.com/locate/semarthrit

Reliability and agreement of proton density-weighted vs. gadolinium-enhanced T1-weighted MRI in hand osteoarthritis. An OMERACT MRI special interest group reliability exercise

Ø Maugesten[a,b,*], S.J. Pedersen[c], M.S. Stoenoiu[d,e], FPB Kroon[f,g], A.J. Mathew[c,h], H.K Genant[j], P.G. Conaghan[k], F. Gandjbakhch[i], M. Kloppenburg[f], C. Peterfy[l], M. Østergaard[c,h], I.K. Haugen[a]

[a] Division of Rheumatology and Research, Diakonhjemmet Hospital, Box 23 Vinderen, Oslo N-0319, Norway
[b] Faculty of Medicine, University of Oslo, Oslo, Norway
[c] Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Rigshospitalet, Copenhagen, Denmark
[d] Department of Rheumatology, Cliniques Universitaires Saint Luc, Brussels, Belgium
[e] Institut de Recherche Expérimentale et Clinique, Université catholique de Louvain, Belgium
[f] Department of Rheumatology, Leiden University Medical Center, Leiden, the Netherlands
[g] Department of Rheumatology, Zuyderland Medical Center, Heerlen, the Netherlands
[h] Department of Clinical Medicine, Faculty of Health and Medical Sciences University of Copenhagen, Copenhagen, Denmark
[i] Department of Rheumatology, Sorbonne University, APHP, Pitié-Salpêtrière Hospital, Paris, France
[j] Department of radiology and biomedical imaging, University of California, San Francisco, United States
[k] Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds and NIHR Leeds Biomedical Research Centre, Leeds, UK
[l] Spire Sciences, Inc., Boca Raton, FL 33487, United States

ARTICLE INFO

Keywords:
OMERACT
Osteoarthritis
MRI
Hand
Synovitis

ABSTRACT

Objectives: To compare reliabilities of assessing synovitis in hand osteoarthritis (OA) using Magnetic Resonance Imaging (MRI) with/without gadolinium (Gd).
Methods: Three readers scored synovitis on non-enhanced two-dimensional (2D) proton density (PD)-weighted MRI and Gd-enhanced (3D) MRI of hand joints in 20 patients. Inter-reader reliabilities were examined.
Results: Reliability was good for Gd-enhanced MRI, but poor for non-enhanced PD-weighted MRI (intraclass correlation coefficient 0.83 and 0.21, respectively). Agreement between the two sequences was poor (weighted kappa 0.18).
Conclusion: Gd-enhanced MRI was more reliable than PD-weighted MRI for assessing synovitis. Gd-enhancement, but also resolution and tissue contrast, might have contributed to this.

Introduction

Gadolinium(Gd)-enhanced magnetic resonance imaging (MRI)-detected synovitis is associated with pain and radiographic progression in hand osteoarthritis(OA) patients [1,2]. Synovitis assessment by MRI has been used as a measure of joint activity in hand OA trials for anti-inflammatory therapies [3–6]. Allergic reactions to Gd-based contrast agents(GBCA) are rare, and the risk of nephrogenic systemic fibrosis seems confined to patients with severe renal insufficiency. However, Gd deposits have been detected in basal cell ganglia,

Dr. Peterfy reports contract research from AbbVie, Centrexion, Flexion Therapeutics, Paradigm Biopharmaceuticals, Regeneron. Dr. Kloppenburg reports consultancy (Abbvie, Pfizer, Levicept, GlaxoSmithKline, Merck-Serono, Kiniksa, Flexion, Galapagos, Jansen, CHDR) and local investigator of industry-driven trial (Abbvie), from Wolters Kluwer (UptoDate), Springer Verlag (Reumatologie en klinische immunologie), grants from Pfizer, grants from IMI-APPROACH. Dr. Østergaard reports grants, personal fees and non-financial support from AbbVie, grants, personal fees and non-financial support from BMS, personal fees from Boehringer-Ingelheim, personal fees from Eli Lilly, personal fees and non-financial support from Janssen, grants, personal fees and non-financial support from Merck, personal fees and non-financial support from Pfizer, personal fees and non-financial support from Roche, grants, personal fees and non-financial support from UCB, grants and personal fees from Celgene, personal fees from Sanofi, personal fees from Regeneron, grants, personal fees and non-financial support from Novartis, personal fees from Gilead. Dr. Haugen has received grants from Pfizer (Advance) and has been in advisory board for Abbvie; All outside the submitted work. There are no relevant disclosures for the remaining authors. The Nor-Hand study was supported by funds from the South-Eastern Norway Regional Health Authority, the Pahles Foundation/the Norwegian Rheumatism Association, the Simon Fougner Hartmanns Family Foundation, and the Trygve Gythfeldt and Wife's Research Foundation. The ExtraFoundation for Health and Rehabilitation through EXTRA funds has financed the PhD position of Øystein Maugesten.
* Corresponding author at: Division of Rheumatology and Research, Diakonhjemmet Hospital, Box 23 Vinderen, Oslo N-0319, Norway.
E-mail address: oystein.maugesten@gmail.com (Ø. Maugesten).

https://doi.org/10.1016/j.semarthrit.2021.05.007
0049-0172/

muscle, liver and skin tissue, questioning whether repeated use of GBCA might be harmful [7–10]. Synovitis and effusion are visible on proton density (PD)-weighted MRI without the need of a radiocontrast agent and could thus be a safer alternative. Our main objective in this study was to compare the reliability of non-enhanced two-dimensional (2D) PD-weighted MRI to that of Gd-enhanced three-dimensional (3D) T1-weighted MRI for assessing synovitis in hand OA patients. Furthermore, we wanted to assess the agreement in detecting synovitis between PD-weighted vs. Gd-enhanced MRI

Methods

Members of the Outcome Measures in Rheumatology(OMERACT) MRI Working Group prepared an atlas containing examples of synovitis in distal interphalangeal (DIP) and proximal interphalangeal (PIP) joints on non-enhanced 2D PD-weighted MRIs in the axial plane and Gd-enhanced 3D T1-weighted MRIs in the axial and sagittal planes (Fig. 1, Supplementary file 1). The grading was based on the Hand Osteoarthritis Magnetic Resonance Imaging Scoring System (HOAMRIS), ranging from 0 to 3 based on thirds of the estimated maximum volume of enhancing tissue in the synovial compartment (0=normal; 1=mild; 2=moderate; 3=severe) [12]. Additionally, enhancement had to be present in 3 consecutive slices in all planes assessed. Example images of joints in which the severity of synovitis (grade 0–3) was regarded the same on both non-enhanced 2D PD-weighted and Gd-enhanced 3D T1-weighted images were included. The atlas was presented during a webinar and key images were selected through consensus. The OMERACT Thumb Base Osteoarthritis Magnetic Resonance Imaging Scoring System (TOMS) was used for scoring of the thumb base joints and the readers were aided by an atlas including MRIs of the 1st carpometacarpal (CMC-1) and scaphotrapeziotrapezoid (STT) joints, with similar grading of synovitis as in HOAMRIS [13]. According to the OMERACT Filter Instrument Selection Algorithm we defined synovitis by PD-weighted and GD-enhanced MRI as core set measure “disease” [11,14].

The MRIs for the calibration and reliability exercise were selected with a random number generator from the Nor-Hand study, which is a cohort including 300 patients aged from 40 to 70 years with confirmed hand OA in at least one joint on clinical examination and/or ultrasound [15]. Participants were imaged with a 1.5 tesla MRI device (Siemens Aera, Germany) with a 16-channel hand/wrist coil covering the fingers and thumb base of the dominant hand. Sequences included 2D PD-weighted Turbo Spin Echo MRI with 3.2 mm axial slices of the fingers along with axial and coronal slices of the thumb base, followed by intravenous GBCA and T1-weighted 3D gradient-echo MRI with 0.4 mm coronal slices and sagittal and axial reformations [15].

A calibration exercise with 7 readers (SVB,FK,SJP,AM,MS,FG,ØM) from 5 centers was arranged. Eleven joints, including DIP 2–5, (P) IP 1–5, CMC-1 and STT joints of the dominant hand of 10 patients were graded from 0 to 3 according to the atlas. After calibration, 2 rheumatologists (SJP, MS) with experience assessing MRIs and 1 PhD-student (ØM) trained in assessing MRI-defined synovitis in hand joints conducted a reliability exercise with 20 patients. Results of the first reliability exercise were discussed in a webinar. Disagreements on the presence or absence of synovitis, or score differences of ≥2 grades were discussed in an online meeting. After re-calibration, a final reliability exercise with 20 new cases was performed by the same 3 readers.

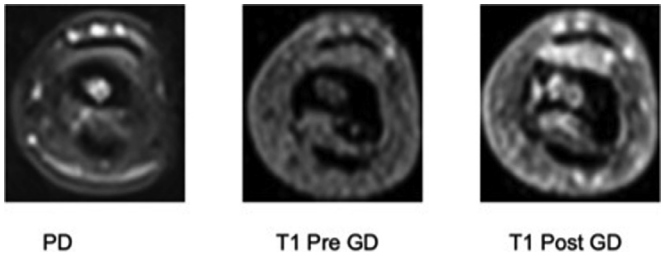

Fig. 1. Example image from the atlas. DIP joints in the axial plane in proton density-weighted MR images and pre/post gadolinium T1-weighted images with grade 3 enhancement.

Mean (SD) sum score of synovitis for all joints together and for different joint groups for 3 readers was calculated. The sum score was based on 11 joints graded from 0 to 3 with a maximum possible sum score of 33. Inter-reader reliability between the two different MRI techniques was calculated with the average (range) intraclass correlation coefficients (ICC) by two-way mixed-effects model with absolute (individual) measure for 3 reader pairs. At joint level, percent exact agreement (PEA) and percent close agreement (PCA) between the 3 readers were calculated. PCA was defined as same grade or one grade difference across the 3 readers. Agreement at joint level between each MRI technique was assessed by linearly weighted kappa values for each reader and presented as a mean (range) weighted kappa value for the 3 readers. Kappa values for dichotomized scores (Grade 0–1 vs. Grade 2–3) between the two MRI techniques were also calculated. All results were presented for all joints together and for separate joint groups. Both ICC, weighted kappa and kappa values were interpreted as poor (0.00–0.19), fair (0.20–0.39), moderate (0.40–0.59), good (0.60–0.79) or very good (0.80–1.00). Stata version 15.0 was used for all analyses. The Nor-Hand study was approved by the regional ethics committee (Ref: 2014/2057).

Results

The study participants were predominantly women (90%) and had a mean (SD) age of 60.5 (6.5) years. Forty percent of the participants had erosive hand OA, defined as having at least one finger joint in the erosive or remodeling phase of the Verbruggen-Veys anatomical phase scoring system, and the mean (SD) Kellgren Lawrence sum score of bilateral DIP, (P)IP, metacarpophalangeal and CMC-1 (range: 0–120) was 27.6 (16.0). Eighty percent fulfilled the American College of Rheumatology hand OA criteria and mean (SD) body mass index was 25.7 (4.5) kg/m^2.

The mean (SD) sum score for synovitis in all hand joints was numerically higher for the non-enhanced PD-weighted images (13.5 (3.6)) than the Gd-enhanced 3D images (10.2 (6.8)). These results were driven by higher scores on the PD-weighted images in the DIP and PIP joints, whereas more synovitis was scored in the thumb base joints on the Gd-enhanced 3D MRIs. Patients with erosive hand OA demonstrated higher scores in both MRI sequences compared to non-erosive hand OA patients (Supplementary Table 1).

For the non-enhanced PD-weighted images, inter-reader reliability for all joint areas was poor or fair. For the Gd-enhanced 3D images, the inter-reader reliability was very good for all joints together and the DIP joints, good for the PIP joints and moderate for the thumb base joints (Table 1).

The agreement between the two MRI techniques was poor to fair (Table 2). Dichotomizing the scores and assessing grade 0–1 versus grade 2–3, slightly improved values, however the agreement remained poor (data not shown). The strongest agreement was found in the PIP joints with fair kappa (range) values of 0.3 (0.3, 0.4). The agreement between PD-weighted and Gd-enhanced MRI remained poor to fair when analyzing erosive hand OA patients only (data not shown).

Discussion

The inter-reader reliability for synovitis detection was poor for the non-enhanced PD-weighted images, regardless of the joint evaluated. Scoring of the Gd-enhanced 3D images demonstrated good reliability

Table 1
Inter-reader reliability across 3 readers for the assessment of synovitis in hand joints of 20 patients with proton density-weighted and gadolinium-enhanced T1-weighted MR images.

	All joints (n = 220)	DIP (n = 80)	PIP (n = 100)	CMC-1, STT (n = 40)
Intraclass correlation coefficients*				
PD	0.20 (0.16, 0.27)	0.14 (−0.03, 0.42)	0.40 (0.26, 0.50)	0.35 (0.27, 0.51)
GD	0.83 (0.78, 0.90)	0.81 (0.76, 0.87)	0.75 (0.66, 0.85)	0.56 (0.44, 0.78)
Percentage exact agreement				
PD	18%	15%	20%	20%
GD	42%	48%	43%	30%
Percentage close agreement				
PD	71%	64%	73%	83%
GD	89%	88%	92%	83%

CMC-1 = 1st carpometacarpal, DIP =distal interphalangeal joints, Gd =Gadolinium-enhanced MR images, PD=Proton Density weighted MR images, PIP = proximal interphlangeal joints, STT = scaphotrapeziotrapezoidal.
*two-way mixed-effects model, absolute agreement, individual measure, average of 3 reader pairs.

for all joints collectively as well as the DIP and PIP joints in particular. Scoring of the thumb base had poorer reliability (ICC=0.6) compared with findings in a previous MRI study by Kroon et al. (ICC=0.8) [13]. This discrepancy might be explained by different level of experience among readers, as Kroon et al. included a radiologist and 2 rheumatologists with extensive MRI experience.

While no previous studies have examined the reliability of the two MRI modalities in hand OA, Hagiwara et al. found moderate to good reliability for both PD-weighted images and Gd-enhanced images when assessing synovitis in Hoffa's fat pad of the knee. The divergent results may suggest the PD-weighted MRI is more suitable for assessment of synovitis in larger rather than smaller joints [16].

Several factors might explain the discrepancy in reliability between non-enhanced PD-weighted and the Gd-enhanced 3D MRIs. Firstly, the slice-thickness of the two techniques differed substantially and might have resulted in loss of information in the small finger joints on the PD-weighted MRIs. Slice thickness on 2D PD-weighted MRI can be reduced slightly with longer scanning times, but cannot feasibly approach that of 3D gradient-echo techniques. Secondly, finger joints were assessed in both axial and sagittal planes with the Gd-enhanced 3D MRIs, but only in the axial plane on PD-weighted images, as the coronal plane with this sequence lacked sufficient resolution. Nevertheless, assessments of the thumb base included coronal and axial planes, and reliability still remained lower for PD-weighted MRIs. Furthermore, poor reliability might be due to the lack of previous experience in assessing PD-weighted MR images among the readers. Finally, the atlas applied has not been externally validated. Poor agreement was detected between Gd-enhanced and PD-weighted MRIs. Although we assume that Gd-enhanced MRI is more accurate than PD-weighted MRI, we cannot make a firm conclusion due to the lack of a true gold standard in our study.

In conclusion we found very good reliability between 3 readers for measuring synovitis in hand OA with Gd-enhanced 3D T1-weighted MRI, but not for non-enhanced 2D PD-weighted MRI. We also found poor agreement between the two MRI modalities. Despite the possible risks and extra costs related to intravenous contrast, our results might suggest that assessment of synovitis in clinical hand OA trials should be done with Gd-enhanced MRIs. However, while Gd-enhancement was probably an important driver of performance, slice thickness, plane of section and reader experience likely contributed significantly as well. Future studies exploring the validity of the two MRI modalities in hand OA are needed.

Table 2
Agreement of synovitis measured by proton density-weighted MRI and gadolinium-enhanced T1-weighted MRI.

	Mean weighted kappa of 3 individual readers (range)
All joints	0.18 (0.18, 0.19)
DIP	0.09 (0.04, 0.16)
PIP	0.27 (0.23, 0.31)
CMC-1, STT	0.17 (0.12, 0.24)

CMC-1 = 1st carpometacarpal joint, DIP =distal interphalangeal joints, PIP = proximal interphalangeal joints, STT = scaphotrapeziotrapezoid joint, w.kappa= weighted kappa.

Acknowledgments

We would like to thank all participants in the OMERACT MRI special interest group (SIG) for input in the project. Thanks to Sjoerd van Beest for participating in the first reliability exercise. We would also like to thank all participants in the Nor-Hand study. PGC is supported in part by the National Institute for Health Research (NIHR) Leeds Biomedical Research center, United Kingdom. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Declaration of Competing Interest

None.

CRediT authorship contribution statement

Ø Maugesten: Conceptualization, Funding acquisition, Formal analysis, Data curation, Writing - original draft, Writing - review editing. **S.J. Pedersen:** Conceptualization, Funding acquisition, Formal analysis, Data curation, Writing - original draft, Writing - review editing. **M.S. Stoenoiu:** Conceptualization, Funding acquisition, Formal analysis, Data curation, Writing - original draft, Writing - review editing. **FPB Kroon:** Conceptualization, Funding acquisition, Formal analysis, Data curation, Writing - original draft, Writing - review editing. **A.J. Mathew:** Conceptualization, Funding acquisition, Formal analysis, Data curation, Writing - original draft, Writing - review editing. **H.K Genant:** Conceptualization, Writing - original draft. **P.G. Conaghan:** Conceptualization, Writing - original draft, Formal analysis, Data curation, Writing - review editing. **F. Gandjbakhch:** Conceptualization, Funding acquisition, Formal analysis, Data curation, Writing - original draft, Writing - review editing. **M. Kloppenburg:** Conceptualization, Writing - original draft, Formal analysis, Data curation, Writing - review editing. **C. Peterfy:** Conceptualization, Writing - original draft, Formal analysis, Data curation, Writing - review editing. **M. Østergaard:** Conceptualization, Writing - original draft, Formal analysis, Data curation, Writing - review editing. **I.K. Haugen:** Conceptualization, Writing - original draft, Formal analysis, Data curation, Writing - review editing.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.semarthrit.2021.05.007.

References

[1] Haugen IK, Boyesen P, Slatkowsky-Christensen B, Sesseng S, van der Heijde D, Kvien TK. Associations between MRI-defined synovitis, bone marrow lesions and structural features and measures of pain and physical function in hand osteoarthritis. Ann Rheum Dis 2012;71(6):899–904.

[2] Haugen IK, Slatkowsky-Christensen B, Boyesen P, Sesseng S, van der Heijde D, Kvien TK. MRI findings predict radiographic progression and development of erosions in hand osteoarthritis. Ann Rheum Dis 2016;75(1):117–23.

[3] Kloppenburg M, Bøyesen P, Visser AW, Haugen IK, Boers M, Boonen A, et al. Report from the OMERACT hand osteoarthritis working group: set of core domains and preliminary set of instruments for use in clinical trials and observational studies. J Rheumatol 2015;42(11):2190–7.

[4] Kloppenburg M, Ramonda R, Bobacz K, Kwok WY, Elewaut D, Huizinga TWJ, et al. Etanercept in patients with inflammatory hand osteoarthritis (EHOA): a multicentre, randomised, double-blind, placebo-controlled trial. Ann Rheum Dis 2018;77(12):1757–64.

[5] Aitken D, Laslett LL, Pan F, Haugen IK, Otahal P, Bellamy N, et al. A randomised double-blind placebo-controlled crossover trial of HUMira (adalimumab) for erosive hand OsteoaRthritis - the HUMOR trial. Osteoarthr Cartil 2018;26(7):880–7.

[6] Kroon FPB, Kortekaas MC, Boonen A, Bohringer S, Reijnierse M, Rosendaal FR, et al. Results of a 6-week treatment with 10mg prednisolone in patients with hand osteoarthritis (HOPE): a double-blind, randomised, placebo-controlled trial. Lancet 2019;394:1993–2001.

[7] Roberts DR, Lindhorst SM, Welsh CT, Maravilla KR, Herring MN, Braun KA, et al. High levels of gadolinium deposition in the skin of a patient with normal renal function. Invest Radiol 2016;51(5):280–9.

[8] Maximova N, Gregori M, Zennaro F, Sonzogni A, Simeone R, Zanon D. Hepatic gadolinium deposition and reversibility after contrast agent-enhanced MR imaging of pediatric hematopoietic stem cell transplant recipients. Radiology 2016;281(2):418–26.

[9] Guo BJ, Yang ZL, Zhang LJ. Gadolinium deposition in brain: current scientific evidence and future perspectives. Front Mol Neurosci 2018;11:335.

[10] Schieda N, Blaichman JI, Costa AF, Glikstein R, Hurrell C, James M, et al. Gadolinium-based contrast agents in kidney disease: a comprehensive review and clinical practice guideline issued by the canadian association of radiologists. Can J Kidney Health Dis 2018;5:2054358118778573.

[11] Maria Antonietta D'Agostino DEB, Maxwell Lara J, Cembalo Sam Michel, Hoens Alison Maria, Hofstetter Catherine, Zabalan Codruta, Bird Paul, Christensen Robin, Wit Maarten de, Doria Andrea S, Maksymowych Walter, Oo Win Min, Østergaard Mikkel, Serban Teodora, Terslev VSSL, Rossum Marion A van, Conaghan Philip G, Boers Maarten. Improving domain definition and outcome instrument selection: lessons learned for OMERACT from imaging. Semin Arthritis Rheum 2021.

[12] Haugen IK, Ostergaard 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(2):386–91.

[13] Kroon FPB, Conaghan PG, Foltz V, Gandjbakhch F, Peterfy C, Eshed I, et al. Development and reliability of the OMERACT thumb base osteoarthritis magnetic resonance imaging scoring system. J Rheumatol 2017;44(11):1694–8.

[14] 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):1021–7.

[15] Gloersen M, Mulrooney E, Mathiessen A, Hammer HB, Slatkowsky-Christensen B, Faraj K, et al. A hospital-based observational cohort study exploring pain and biomarkers in patients with hand osteoarthritis in Norway: the nor-hand protocol. BMJ Open 2017;7(9):e016938.

[16] Hagiwara S, Yang A, Takao S, Kaneko Y, Nozaki T, Yoshioka H. New scoring system in assessment of Hoffa's fat pad synovitis: a comparative study with established scoring systems. World J Radiol 2018;10(11):162–71.