This is a repository copy of Development and Validation of the OMERACT Rheumatoid Arthritis Magnetic Resonance Tenosynovitis Scoring System in a Multireader Exercise.

White Rose Research Online URL for this paper: http://eprints.whiterose.ac.uk/127116/

Version: Accepted Version

Article:

https://doi.org/10.3899/jrheum.161097


Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
**Development and Validation of the OMERACT Rheumatoid Arthritis Magnetic Resonance (RAMRIS)**

**Tenosynovitis Scoring System in a Multi-reader Exercise**


**ABSTRACT**

*Objective* To develop and validate a magnetic resonance imaging (MRI) tenosynovitis (TS) score for tendons at the wrist and metacarpophalangeal (MCP) joint levels in rheumatoid arthritis (RA) patients.

*Methods* Axial T1-weighted pre- and post-contrast fat-saturated MR image-sets of the hand of 43 RA patients initiating rituximab therapy were obtained at baseline and after 14, 26, 38 or 52 weeks. The MR-images were scored twice by four readers. Nine tendon compartments of the wrist and four flexor tendon compartments at the MCP joints were assessed. Tenosynovitis was scored as follows: 0: No; 1: <1.5mm; 2: ≥1.5mm but <3mm; 3: ≥3mm peritendinous effusion and/or post-contrast enhancement. Intra-/inter-reader intra-class correlation coefficients (ICC), smallest detectable change (SDC), percentage of exact and close agreement (PEA/PCA), and standardized response mean (SRM) were calculated.

*Results* Intra-/inter-reader ICC for status and change scores were very good (≥0.80) for total scores for all readers. Intra-reader SDC was ≤3.0 and inter-reader SDC was <2.0. The overall PEA/PCA intra-/inter-reader agreement for change scores in all tendons were 73.8%/97.6% and 47.9%/85.0% respectively. Average SRM was moderate for total scores and 60.5% of the patients had a tenosynovitis change score ≥SDC.
Conclusion The TS-score showed high intra-/inter-reader agreement for wrist and finger tendons, moderate responsiveness and the majority of the patients showed a change above the SDC. This scoring system may be included as a component of the RAMRIS.

Key indexing terms

MRI, Tenosynovitis, Rheumatoid Arthritis, OMERACT

Name of departments and institutions to which the work should be attributed

Copenhagen Center for Arthritis Research (COPECARE), Center for Rheumatology and Spine Diseases, Rigshospitalet, Glostrup, Denmark

and

Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, and NIHR Leeds Musculoskeletal Biomedical Research Unit, Leeds Teaching Hospitals NHS Trust, Leeds, UK

Sources of support in the form of grants or industrial support

This exercise was supported by the Danish Rheumatism Association. The clinical trial from which these data were acquired was supported by a research grant and study drug from Roche.
Initials, surnames, appointments and highest academic degrees of all authors

D Glinatsi, MD, research fellow, Copenhagen Center for Arthritis Research (COPECARE), Center for Rheumatology and Spine Diseases, Rigshospitalet, Glostrup, Denmark

P Bird, BMed (Hons), FRACP, PhD, Grad Dip MRI, Associate Professor, University of NSW, Sydney, Australia

F Gandjbakhch, MD, Practising Rheumatologist, Hôpital Pitié-Salpétrière, APHP, université Paris VI, Paris, France

EA Haavardsholm, MD, PhD, Postdoctoral Researcher, Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway

CG Peterfy, MD, PhD, FRCP, Chief Executive Officer, Spire Sciences inc., Boca Raton, FL, USA

EM Vital, MRCP, PhD, Associate Professor and Honorary Consultant, Leeds Institute for Rheumatic and Musculoskeletal Medicine, University of Leeds, & NIHR Leeds Musculoskeletal Biomedical Research Unit, UK

P Emery, MA, MD, FRCP, ARC Professor in Rheumatology, Leeds Institute for Rheumatic and Musculoskeletal Medicine, University of Leeds, & NIHR Leeds Musculoskeletal Biomedical Research Unit, UK

PG Conaghan, MB, BS, PhD, FRACP, FRCP, Professor of Musculoskeletal Medicine, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, & NIHR Leeds Musculoskeletal Biomedical Research Unit, UK

M Østergaard, MD, PhD, DMSc, Professor, Copenhagen Center for Arthritis Research (COPECARE), Center for Rheumatology and Spine Diseases, Rigshospitalet, Glostrup, Denmark and the Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

Primary e-mail address for each author

daniel.glinatsi@gmail.com
Name and address of author to whom requests for reprints should be made

Daniel Glinatsi
Copenhagen Center for Arthritis Research (COPECARE), Center for Rheumatology and Spine Diseases, Rigshospitalet, Nordre Ringvej 57, 5th entrance
2600 Glostrup
DENMARK
daniel.glinatsi@gmail.com

Name and address, telephone and fax of author responsible for correspondence about the manuscript

Daniel Glinatsi
Rigshospitalet, Center for Rheumatology and Spine Diseases, 5th entrance

2600 Glostrup

DENMARK

Phone: +45 38 63 32 49

Fax: +45 38 63 39 61

daniel.glinatsi@gmail.com

A short running footnote of no more than 4 words

MRI in tenosynovitis
What is new?

- Based on knowledge about existing MRI TS-scores, the OMERACT MRI in Arthritis Working Group developed the first consensus-based TS-score
- The TS-score was assessed for intra-/inter-reader agreement and responsiveness
- Intra-/inter-reader agreement for status and change scores was high and responsiveness was moderate
- We conclude that the TS-score may be included as a component of the RAMRIS

INTRODUCTION

Tenosynovitis (TS) in the hand of rheumatoid arthritis (RA) patients is a frequent and early occurring inflammatory feature that can cause tendon rupture and may be associated with subsequent bone erosion(1-5). Magnetic resonance imaging (MRI) provides excellent visualization of bone and soft tissues in the hand, including TS. The Outcome Measures in Rheumatology (OMERACT) MRI in Arthritis Working Group developed an RA MRI score (RAMRIS) for synovitis, osteitis, bone erosion and recently, for joint space narrowing(6, 7). Adding TS to the RAMRIS may provide improved power to study and monitor the disease course of RA.

Haavardsholm et al. designed a TS-score for the wrist that demonstrated good reliability and responsiveness(8, 9). In addition, the Working Group has suggested and validated a TS-score for the fingers as a part of the psoriatic arthritis MRI Score (PsAMRIS)(10-12). However, none of these scores covered the tendons at both wrist and metacarpophalangeal (MCP) joint level.
Our aim was to develop and validate an RA TS-score for tendons at wrist and MCP joint levels. Here we present previously unpublished results that demonstrate the reliability of this novel score, enabling it to be included in the internationally renowned RAMRIS.

METHODS

Development of scoring system

In accordance with the processes in the OMERACT Handbook for technical tool development, previous MRI TS-scores were identified through a thorough review of the literature by a fellow of the Working Group, which was constituted from 3 continents and included rheumatologists, radiologists, methodologists and industry representatives. The TS-scores suggested by Haavardsholm et al. and the PsAMRIS appeared most thoroughly assessed methodologically and were tested in a pilot phase where the two TS-scores were assessed at wrist and MCP joints by two readers at two timepoints. The methodology by Haavardsholm et al. seemed more applicable to the wrist area and was therefore chosen for subsequent intra-/inter-reader agreement analyses in a pre-exercise with 80 hands at two timepoints (baseline/1-year follow-up), in which exact thickness of TS was also measured. The results were presented and discussed at a subsequent meeting of the Working Group and, based on tendon sheath thickness distributions (supplement 1), modifications were implemented. The group then proceeded to perform the multi-reader exercise below.

Image selection
MR image-sets were selected from a study(13) where patients with active RA (disease activity score >3.2, rheumatoid factor and/or anti-cyclic citrullinated peptide positive, ≥1 previous disease-modifying antirheumatic drugs) were treated with 2-3 doses of 1000mg rituximab at baseline.

MRI of the dominant hand was performed on a Siemens 1.5 Tesla MRI-unit using a dedicated hand coil. Axial T1-weighted fat-saturated pre- and post-contrast gradient echo MR image-sets (slice thickness 0.45mm, repetition time 30ms, echo time 6.8ms, field of view 150mm) were obtained at baseline(n=43) and after 14(n=5), 26(n=8), 38(n=15) or 52(n=15) weeks.

Scoring of images

Four readers (DG;FG;MØ;PB) with previous experience in MRI-assessment of TS participated in the exercise. The readers performed a calibration session the evening before the exercise. Forty-three paired MR image-sets were blinded for patient data but not for chronology(14), and were read twice on identical 23-inch screens over two days with re-anonymizing and re-randomizing in between the two reads.

Reader rules and scoring system

At the wrist, six extensor tendon compartments and three flexor tendon compartments were assessed between the radioulnar joint and the hook of hamate. At the level of the 2nd-5th MCP joints, flexor tendons were assessed in an area from 1cm proximal to 1cm distal to each joint (figure 1A).

TS was defined as peritendinous effusion and/or post-contrast enhancement of the tendon sheath seen on axial sequences over ≥3 consecutive slices. The maximum width of the effusion/enhancing tendon sheath
was measured perpendicularly to the tendon and the TS-score was graded as follows: 0: No; 1: <1.5mm; 2: 
≥1.5mm but <3mm; 3: ≥3mm peritendinous effusion and/or post-contrast enhancement (figure 1A). Tendon 
sheaths of crossing tendons were measured proximally to the crossing point and tendon bundles within a 
common tendon sheath were assessed as one unit (figure 1B-C). See supplement 2 for examples of TS.

**Statistical analysis**

Descriptive statistics and the Wilcoxon signed-rank test were used to assess change in score over time.

Intra-/inter-reader agreement was assessed using single and average measure intra-class correlation 
coefficient (smICC/avmICC) respectively. The smallest detectable change (SDC) was calculated for intra-/inter-
reader change scores(15) and was also expressed as the percentage of maximum observed change (%MOC). 
The percentage of exact agreement (PEA, scores equal) and the percentage of close agreement (PCA, scores ≤1 
different) between the 2 reads and the 4 readers were calculated for intra-/inter-reader agreement, 
respectively.

Responsiveness was assessed by the standardized response mean (SRM). Ability to show change was also 
assessed by the percentage of patients with change score ≥SDC.

**RESULTS**

Baseline and follow-up characteristics of the patients are presented in supplement 3. Median (range) 
change in total TS score was -1.0(-1.0;-2.5) (p<0.01).
Intra-reader smICCs for baseline scores were very good for all parameters in all readers, except for MCP-scores in one reader (smICC 0.79). Intra-reader smICCs for change scores were good to very good for all parameters. All readers demonstrated very good intra-reader smICCs for total scores. Baseline/change inter-reader avmICCs were >0.90 (i.e. very good) for all parameters. Median (range) intra-reader SDC was 2.8(2.1;3.0) for total scores. Inter-reader SDCs for scores averaged over four readers were <2.0 for wrist, MCP (<1.0) and total scores. The percentage of patients with a total change score ≥SDC was 39.5% to 54.7% for intra-reader SDCs. For inter-reader SDCs, this percentage was 60.5%. The %MOC was below 20% for intra-/inter-reader total scores (table 1).

The overall PEA/PCA intra-/inter-reader agreement for change scores in all tendons were 73.8%/97.6% and 47.9%/85.0% respectively (table 1). PEA/PCA for individual tendons is presented in supplement 4. The inter-reader PEA/PCA for pairs of 2 readers were 72.4%/99.6% (baseline) and 73.3%/98.8% (change) for the best matched pair and 64.3%/98.3% (baseline) and 67.5%/97.0% (change) for the average pair when all tendons were considered.

The average SRM for total scores was moderate (table 2).

Overall, there were no differences in ICC, %MOC, PEA/PCA and SRM between tendon sheaths at wrist versus MCP joint levels.

DISCUSSION

This longitudinal multi-reader exercise of active RA patients showed that the TS-score had high intra-/inter-reader agreement and moderate responsiveness. The results were similar at wrist and MCP joint level.
Intra-/inter-reader ICCs were very good, both for total baseline and change scores indicating that the TS-score is reliable and can monitor change over time.

The average intra-reader PEA was high and the PCA was close to 100%. Inter-reader PEA was acceptable considering the increased difficulty of 4 assessors needing to reach exact agreement, as compared to 2 assessors. Since most clinical trials have 2 assessors we also analyzed the PEA/PCA for paired readers, and found the average percentages close to the intra-reader agreements and the percentages of the best paired assessors in level with the intra-reader agreements.

The inter-reader SDC was ≤2.0 and the %MOC <11%. Although the average SRM was only 0.50 (moderate), >60% of the patients showed a change ≥SDC. Variable treatment regimens (related to different timing of infusions) and different timing of follow-up MRI may have affected the responsiveness but would not be expected to alter the intra-/inter-reader agreement. Future studies should assess the responsiveness in early RA cohorts and in placebo-controlled studies, where responsiveness can be compared between groups.

This TS-score was developed as a modification from the score by Haavardsholm et al. after assessing this and the PsAMRIS TS-score in a pilot phase (8-11). A key modification of the methodology by Haavardsholm et al. was to narrow the intervals of tendon sheath thickness within each increment to potentially increase the ability to detect change. As proven by this exercise, the reliability of the current TS-score remained high. The responsiveness was not increased compared to the results reported by Haavardsholm et al.(9), but since therapies in the two cohorts were different, the results are not fully comparable. Other advances, compared to the score by Haavardsholm et al. included clarifications on how to score tendon sheaths of crossing tendons and common tendon sheaths of tendon bundles (figure 1B-C).

Other MRI TS-scores have previously been used. Several have not assessed reliability and responsiveness(1, 16) or have scored TS qualitatively as absent/present(2, 4, 17-22). In terms of monitoring
change, qualitative scores may have less power to detect changes in TS. Semiquantitative TS-scores have been suggested by McQueen et al.(23) and Schirmer et al.(24), but did not assess the performance in longitudinal settings nor the inter/intra-reader agreement. Lisbona et al.(25) suggested a TS-score for the hand, based on incomplete and complete halos of enhancing tendon sheath, which showed high intra-/inter-reader ICCs, but small SRM.

This TS-score was developed as a potential addendum to the existing RAMRIS(6, 26). Therefore, we strived to design the score so that it covered the tendons at the joint regions included in the RAMRIS core set. Since TS is scored on the same MRI-sequences and projections as synovitis, these pathologies may be scored simultaneously and therefore the addition of TS-score will only add a small amount of time to assessing a hand when using the RAMRIS method. Based on this and the reliability and responsiveness data, we conclude that this TS-score fulfills the OMERACT filter criteria concerning truth, discrimination and feasibility(27) and may be included as a component of the RAMRIS for assessing TS of the hand in RA clinical trials.

Acknowledgement

We thank illustrator Axel Norén for help with figure 1 and the Danish Rheumatism Association. The clinical trial from which these data were acquired was supported by a research grant and study drug from Roche. The research is supported by the National Institute for Health Research (NIHR) Leeds Musculoskeletal Biomedical Research Unit. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Conflicts of interest
DG:

PB: None

FG: Consultation fees from: Pfizer, AbbVie, Roche, Janssen; Research Support and Grants from Pfizer, Roche

EAH: Consultation fees from: Pfizer, MSD; Research Support and Grants from AbbVie, UCB, MSD, Pfizer, Roche

CGP: Shareholder in Spire Sciences inc.

EMV: Research support and grants from Roche and AstraZeneca

PE: None

PGC: Consultant or speakers bureau for Abbvie, Lilly, Novartis, Pfizer, Roche; research grant from BMS

MØ: Consultation fees from: Abbvie, BMS, Boehringer-Ingelheim, Celgene, Eli-Lilly, Centocor, GSK, Hospira, Janssen, Merck, Mundipharma, Novartis, Novo, Orion, Pfizer, Regeneron, Schering-Plough, Roche, Takeda, UCB, Wyeth; Research support and grants from: Abbvie, BMS, Janssen and Merck


At the wrist, six extensor tendon compartments and three flexor tendon compartments are assessed and at the level of the 2nd-5th metacarpophalangeal (MCP) joints, the flexor tendons are assessed. The range of the scores is 0-27, 0-12 and 0-39 for the tendons at the wrist, MCP joints and total score respectively (A). Measuring the common tendon sheath perpendicularly from an individual tendon (B) will result in overestimation of the thickness of the tendon sheath. Therefore, the tendons within a common tendon sheath are assessed as one unit (here, illustrated by the dashed line) and the enhancing tendon sheath is measured perpendicularly to the unit, at the thickest point (C).