Novel quantitative digital image analysis methodology for assessment of inflammatory changes in MRI data in a post-hoc analysis of data acquired from a phase IIb study of baricitinib in patients with active rheumatoid arthritis

M. Boesen a, b, *, S.D. Beattie c, D.E. Schlichting c, O. Kubassova b

a Department of Radiology, Copenhagen University Hospital Bispebjerg-Frederiksberg, Copenhagen, Denmark
b Image Analysis Group, London, UK
c Eli Lilly and Company, Indianapolis, IN, USA

ARTICLE INFO

Keywords:
Rheumatoid arthritis
Baricitinib
MRI
Quantitative imaging
Biomarker
Inflammation

ABSTRACT

Purpose: To evaluate a novel quantitative methodology to assess inflammatory changes in magnetic resonance imaging (MRI) data from patients with rheumatoid arthritis (RA) and the impact of image quality on imaging outcomes compared to the RA Magnetic Resonance Imaging Score (RAMRIS).

Methods: Three-dimensional, T1-weighted, fat-suppressed MRI sequences of the hand/wrist before and after intravenous Gadolinium contrast from patients with RA in a placebo-controlled clinical trial (NCT01185353) were re-evaluated post hoc. The methodology was integrated into proprietary software (DYNAMIKA®) and assessed inflammation through pixelated measurements of the contrast-enhancing (inflammatory) volume. A semi-automatic approach outlined contrast-enhancing synovial tissue in the wrist and second to fifth metacarpophalangeal joints with a rough region of interest (ROI); quantitative imaging biomarkers were generated by means of quantitative total volume of inflammation and quantitative degree of inflammation relative to the signal in a 1 cm in diameter ROI in the center of the thenar or lumbrical muscle for internal reference. The time from Gadolinium injection to finalization of the post-contrast images was calculated from the images’ Digital Imaging and Communications in Medicine header. An experienced reader graded image quality as poor, acceptable, or good.

Results: Results from this quantitative methodology, especially when excluding images with poor quality scores (14–32%), provided a more pronounced and monotonically increasing dose-response than the original RAMRIS results on synovitis and osteitis.

Conclusions: This computer-aided quantitative scoring method provided continuous measures of inflammatory changes relative to muscle and may be more sensitive and interpretable concerning dose/response separation between RA treatment groups.

1. Introduction

Imaging techniques are integral in assessing response to treatment and disease progression in rheumatoid arthritis (RA) [1]. Reader-driven assessments, such as the Rheumatoid Arthritis Magnetic Resonance Imaging Score (RAMRIS) using categorical scores, are widely used and approved as a reference standard in many magnetic resonance imaging (MRI) trials for RA; alternative computer-aided approaches have been proposed [2–4] to provide continuous measures of the volume and degree of inflammation. However, there are several shortcomings to these...
methods which result in gaps in how inflammatory changes are measured in patients with RA.

Baricitinib is an oral, selective inhibitor of Janus kinase 1 and Janus kinase 2—protein tyrosine kinases implicated in the pathogenesis and progression of RA. It is approved for the treatment of adults with moderately-to-severely active RA in more than 70 countries including Canada, the United States, Japan, and in the European Union. The efficacy of baricitinib as a treatment for patients with moderately-to-severely active RA with an inadequate response to methotrexate therapy was demonstrated in a phase IIb study [5]. The American College of Rheumatology 20% response rates (the primary endpoint at 12 weeks) in that study were significantly higher for the 4-mg and 8-mg doses of baricitinib compared to placebo. A sub-study was conducted using MRI to assess inflammatory and structural damage (synovitis, osteitis, and bone erosion) via Outcome Measures in Rheumatology Clinical Trials (OMERACT) RAMRIS [6] and to support dose selection for the baricitinib phase III program [7]. Patients with documented radiographic erosions and no contraindications to MRI were invited to participate in this sub-study.

This research presents a novel, computer-supported, radiological decision methodology focusing on the quantitative assessment of inflammation in standard post-contrast T1-weighted fat-saturated MRI images. This method was tested post hoc on acquired imaging data from the above mentioned study (which was limited with respect to the imaging set-up used) to investigate if it could provide similar or better dose/response separation than the original RAMRIS method [5]. The primary objective of the study was to evaluate this novel quantitative analysis tool and its ability to segment synovial and inflammatory volume of the wrist and the second to fifth metacarpophalangeal (MCP) joints in patients with RA. The secondary objective was to assess how data quality and timing of contrast injection in relation to image acquisition affects these measures and, in turn, the outcomes of an analysis performed in a randomized clinical trial setting.

2. Materials and methods

2.1. Study design and patient population

Detailed methods of the primary study were described previously [5] (ClinicalTrials.gov Identifier: NCT01185353) and are summarized here. Individuals ≥ 18 and ≤ 75 years of age (n = 301) with active disease despite methotrexate treatment were enrolled between October of 2010 and February of 2012. Active RA was defined as having ≥ 8 of 66 swollen joints and ≥ 8 of 68 tender joints, and a C-reactive protein measurement ≥ 1.2 times the upper limit of normal (ULN) or erythrocyte sedimentation rate > ULN (28 mm/h). Patients must have received methotrexate for ≥ 12 weeks, with a stable dose of 10 to 25 mg per week for ≥ 8 weeks, prior to baseline. Key exclusion criteria included history of biologic disease-modifying antirheumatic drug therapy, recent or concurrent infection including active or latent tuberculosis, an estimated glomerular filtration rate from serum creatinine of < 50 mL/min and any history of chronic liver disease or current serum aspartate aminotransferase or alanine aminotransferase concentration > 3 × ULN or total bilirubin ≥ 1.5 × ULN.

Patients were randomized 2:1:1:1:1 to receive placebo or baricitinib 1, 2, 4, or 8 mg orally once daily for 12 weeks. At week 12, patients previously assigned to placebo or baricitinib 1 mg were re-randomized (1:1) to baricitinib 2 mg twice daily or baricitinib 4 mg once daily through week 24. Patients initially assigned to the 2-, 4-, or 8-mg doses of baricitinib continued the same treatment through week 24.

The study was conducted in accordance with the ethical principles of the Declaration of Helsinki and Good Clinical Practice Guidelines. All patients provided written informed consent.

2.2. Imaging

MRIs of the most severely involved hand/wrist, determined by the highest swollen/tender joint count at baseline, were obtained from patients with ≥ 1 definitive radiographic erosion at baseline and repeated at week 12 and week 24. MCP and proximal interphalangeal joints were scanned separately from the wrist joints (carpometacarpal, intercarpal, and distal radioulnar) using surface coils and 1.5-Tesla MRI scanners before and after an intravenous bolus infusion of Gadolinium contrast (Fig. 1A). Both an acrylic hand frame and planar alignment of coronal slices were utilized to ensure reproducibility of cross-sectional anatomy on serially acquired MRI scans. Three MRI sequences were acquired for the hand and wrist, coronal two-dimensional Short Tau Inversion Recovery and three-dimensional pre- and post-contrast T1-weighted, fat-saturated, gradient echo (GRE), for a total of six series per study. A study was defined as the group of MRI series acquired per patient per time point. Owing to the limited field of view, the patient’s hand was moved, and the coil was repositioned during the image acquisition between the wrist and the MCP and the finger joints following the Gadolinium injection (Fig. 1B). This introduced a systematic time delay of the post-contrast sequences in the two anatomies in all patients. In the original published study [5], all images were read independently in two sessions by two expert radiologists who were blinded to both treatment assignment and visit order for synovitis, osteitis, and bone erosion using RAMRIS [5], a process that usually is completed in 10 to 15 min. In addition, cartilage loss was scored with the validated 9-point Cartilage Loss Scale [8], and scale-adjusted scores for combined inflammation (osteitis + 3 × synovitis) and total joint damage (erosion + 2.5 × cartilage loss) were calculated.

2.3. Normalized intensity methodology

The post-contrast MRI sequences were utilized to measure the volume of synovial and bony inflammation in the wrist and second to fifth MCP joints and compute the degree (intensity) of contrast enhancement at the time of acquisition. Contrast enhancement in the image was normalized against a standardized, 1 cm in diameter baseline region of interest (ROI) placed in the center of an included muscle of the hand (thenar muscle for wrist and lumbrical muscle for MCP joints) within the same sequence using a custom, computer-aided algorithm in the software DYNAMIKA® (image Analysis, LTD, London, England, UK). An experienced radiologist outlined the second to fifth MCP joints and the entire wrist joints by a rough ROI, including all visible, contrast-enhancing synovitis and bone, while excluding enhanced larger blood vessels in all available slices displaying the anatomy. The process of outlining ROIs took approximately 20 min. From the average signal intensity of the muscle ROI, , the normalized intensity (NormI) parametric map was calculated as . The aforementioned average was treated as a fixed constant and a 95% prediction interval for a new normalized observation consistent with muscle signal intensity was calculated as . The aforementioned average was treated as a fixed constant and a 95% prediction interval for a new normalized observation consistent with muscle signal intensity was calculated as , where corresponds to the standard deviation of signal intensities within a ROI. Only voxels with a NormI value greater than this prediction interval threshold were classified, highlighted, and used for the volume and inflammation calculation inside the rough ROIs, as they were deemed statistically incompatible with muscle signal intensities. Semi-automatic extracted imaging biomarkers, quantitative total volume of inflammation (QTVI), and quantitative inflammation (signal intensity) were generated from each ROI to provide a continuous volume and degree of contrast enhancement reflecting the inflammatory severity and allowing depiction of areas of both high and low contrast uptake relative to muscle. A test–retest analysis of a random sample was performed, yielding an intraclass correlation coefficient (ICC) for intra-reader reproducibility, to assess the reliability of the estimates of intensity and the extracted biomarkers from the placement of the ROIs.

Methodological details regarding the development of a NormI map
and imaging biomarkers obtained from it are outlined in Supplemental Fig. 1 and Supplemental Table 1. A NormI map superimposed onto a sample T1-weighted post-contrast image of the MCPs and wrist is shown in Supplemental Fig. 2.

2.4. Image quality assessment

Experience readers who assessed MRIs for quality were blinded to treatment regimen, patient visits, and sequence acquisition time after Gadolinium injection; they assigned a visual image quality score to each image—1 = poor, 2 = acceptable, and 3 = good quality—focusing on image artefacts and representative anatomy inside the image field of view. Images were further categorized by quality based on a combination of the visual quality score and a quantitative Digital Imaging and Communications in Medicine [DICOM] header assessment-based score measuring the time in seconds of image acquisition following intravenous Gadolinium injection. Table 1 and Fig. 2 present image quality score assignment and a key to interpretation; the higher the score, the better the quality of an MRI scan. This information was used to exclude data based on poor quality resulting from acquisition issues and image artefacts that are known to affect computer-aided analysis, including missing or incomplete fat saturation, severe movement, pulsation, ring and/or aliasing artefacts, missing image anatomy in the field of view, as well as imaging protocol violations like inconsistency in the duration between injection of Gadolinium contrast injection and post-contrast scan of wrist and/or MCP joints of more than 10 min.

2.5. Quality assessment according to gadolinium injection timing delays

The delay between Gadolinium injection and image acquisition (Gadolinium delay) was estimated from the acquisition time (as stored in the DICOM header) and the duration of individual sequences. This estimated Gadolinium delay relied on two assumptions in its calculation: 1) the T1-weighted GRE sequence took 05:30 (mm:ss) to acquire (based
on the imaging manual) and 2) the injection of the Gadolinium-based contrast agent occurred immediately after the acquisition of the last pre-contrast T1-weighted GRE fat-saturated sequence.

2.6. Statistical analysis

Time points from 12 patients (approximately 8% of the total scans) were randomly selected for repeat analysis to examine the intra-reader reproducibility of individual biomarkers (expressed as ICC).

Any images that were obtained from patients at the time of early study discontinuation were considered as having been obtained at the scheduled study visit (week 12 or week 24) for analysis purposes. This form of last-observation-carried-forward imputation was applied because there was only one scheduled observation to be obtained in each period of the study design relative to randomized treatment groups.

Statistical analyses were performed using two sets of image data. In one, all available images were included in the analysis set. In the second, only images that were categorized as “good” or “acceptable” were retained. For the latter purpose, the applied quality criteria included the image quality (see Image Quality Assessment) and the timing consistency in the delay in Gadolinium injection between baseline and post-contrast images (see Quality Assessment According to Gadolinium Injection Timing Delays). A composite score of “good” was defined as good for both image quality and timing consistency; “acceptable” was defined as good or acceptable for both, but not both “good” scores; “poor” was defined as poor for either image quality or timing consistency. After excluding images (if any), all patients included in the analyses had to have a baseline and ≥ 1 post-baseline observation. Data were analyzed by mixed model for repeated measures analysis on the change from baseline values with treatment group, time point (week 12 or week 24), and treatment group by time point interaction as fixed factors and baseline score and baseline by time point interaction as fixed covariates in the model. The week-24 data for patients randomized to the placebo and baricitinib 1 mg treatment groups were included in the unified mixed model for repeated measures but were withheld for reporting, as these data do not represent expected effects at 24 weeks corresponding to initially randomized therapy. Each baricitinib dose group was compared to placebo via least-squares means and p-values unadjusted for multiplicity.

### 3. Results

A total of 154 (of 301) patients randomized in the primary study with definitive radiographic erosion had an MRI of the hand and wrist at baseline, week 12, and/or week 24, and were subsequently scored for synovitis and osteitis by experienced radiologists using RAMRIS (data previously published [3]). Of these patients, 144 had both baseline and post-baseline imaging completed and their images sent for re-analysis (Supplemental Fig. 3a). Of these, 142 had an MRI obtained at week 12 and 130 had an MRI obtained at week 24 (Supplemental Fig. 3b). Of the missing data, six were the result of early study discontinuation, all for the images at week 24, while the others were the result of missed MRI scans. Baseline demographic and disease characteristics were generally well balanced and RAMRIS-based scores were reported previously [3]. Overall, data were well distributed between the baricitinib treatment groups and placebo: placebo (n = 42), baricitinib 1 mg (n = 25), baricitinib 2 mg (n = 29), baricitinib 4 mg (n = 24), and baricitinib 8 mg (n = 24).

All available time points from 12 randomly selected patients (n = 34 in total) were chosen for repeat analysis to examine intra-reader reproducibility. A total of 154 (of 301) patients randomized in the primary study with definitive radiographic erosion had an MRI of the hand and wrist at baseline, week 12, and/or week 24, and were subsequently scored for synovitis and osteitis by experienced radiologists using RAMRIS (data previously published [3]). Of these patients, 144 had both baseline and post-baseline imaging completed and their images sent for re-analysis (Supplemental Fig. 3a). Of these, 142 had an MRI obtained at week 12 and 130 had an MRI obtained at week 24 (Supplemental Fig. 3b). Of the missing data, six were the result of early study discontinuation, all for the images at week 24, while the others were the result of missed MRI scans. Baseline demographic and disease characteristics were generally well balanced and RAMRIS-based scores were reported previously [3]. Overall, data were well distributed between the baricitinib treatment groups and placebo: placebo (n = 42), baricitinib 1 mg (n = 25), baricitinib 2 mg (n = 29), baricitinib 4 mg (n = 24), and baricitinib 8 mg (n = 24).

All available time points from 12 randomly selected patients (n = 34 in total) were chosen for repeat analysis to examine intra-reader reproducibility.

### Abbreviations: Δ, difference from baseline; IQR, interquartile range; Q, quartile

---

Table 1

<table>
<thead>
<tr>
<th>Score/Measure</th>
<th>Description</th>
<th>Output Key Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Image quality</td>
<td>Visual assessment of image suitability, scored by experienced radiologist</td>
<td>(1) Poor</td>
</tr>
<tr>
<td>2. Scan timing delay: Length of estimated time in minutes from Gd injection to performance of post-contrast T1-W sequences</td>
<td></td>
<td>(1) Gd delay &gt; Q3 + 1.5 IQR</td>
</tr>
<tr>
<td>3. Scan timing consistency: Consistency of estimated Gd delay from baseline in minutes</td>
<td></td>
<td>(1) Gd delay Δ &lt; Q1 ≤ 1.5 IQR or &gt; Q3 + 1.5 IQR</td>
</tr>
</tbody>
</table>

Abbreviations: Δ = difference from baseline; Gd = Gadolinium; IQR = inter-quartile range; Q = quartile; T1-W = T1-weighted.
reproducibility. Intra-reader agreement (expressed as ICC) for the mean intensity values of ROIs was 0.987 (95% confidence interval [CI] 0.974, 0.993, \( P < 0.001 \)) for the MCP area and 0.988 (95% CI 0.977, 0.994, \( P < 0.001 \)) for the wrist area. Intra-reader agreement for the test–retest repeatability of the individual quantitative biomarkers was \( \geq 0.849 \) (Table 2).

Fig. 3 displays the wrist series of a single patient in which there was comparable longitudinal placement of ROIs and highly reproducible placement between the test and re-test analysis. Fig. 4 shows a sample MRI series of the MCP joints and wrist of a patient treated with 4 mg baricitinib at baseline, week 12, and week 24. Inflammation was resolved over time as shown by the reduced numbers of color-coded enhanced pixels, encircled by ROIs.

For the image-quality analysis, 416 time points from 144 patients were examined (Table 3). Across the three evaluated time points, approximately 93% of the MRIs were considered of good or acceptable quality (Fig. 5). The frequency distribution of the estimated imaging delay in minutes after Gadolinium injection for the analyzed series and the resulting quality scores are displayed in Supplemental Fig. 4. The number of series with Gadolinium delays greater than quartile 3 \( + 1.5 \times \text{interquartile range (IQR)} \) range (outliers with a significantly longer than expected Gadolinium delay) was 11 (3%) for the MCP series, corresponding to a delay longer than 24:20 (mm:ss), and 17 (4%) for the wrist series, corresponding to a delay longer than 24:20 (mm:ss). Consistency between baseline and post-baseline visits also varied significantly, as shown in Supplemental Fig. 5. The change in Gadolinium delay from baseline acquisition varied by \( > 1.5 \times \) IQR for 21 patient visits for the hand acquisition, corresponding to differences \( < -10:45 \) (mm:ss) or \( > 7:48 \) (mm:ss), and for 48 patient visits for the wrist sequences, corresponding to differences \( < -8:34 \) (mm:ss) or \( > 6:40 \) (mm:ss). The data used when the analysis was limited to good and acceptable quality images ranged from 68% to 86% (Supplemental Table 2).

Adjusted (least-squares) mean changes from baseline to weeks 12 and 24 in QTVI using images of good and acceptable quality and all qualities are shown in Fig. 6A and B, respectively. At week 12, when limiting to images of good and acceptable quality, statistically significant mean decreases from baseline in QTVI were observed for the 8-mg baricitinib dose (\( P < 0.01 \) relative to placebo for images of wrist and \( P < 0.05 \) relative to placebo for images of MCP and wrist combined) and 4-mg (\( P < 0.05 \) relative to placebo for images of wrist and MCP and wrist combined) baricitinib doses.

4. Discussion

This post-hoc study re-evaluated the MRI sequences of the hand/wrist before and after intravenous Gadolinium contrast in patients with RA in a randomized placebo-controlled phase IIb study of baricitinib using a new, computer-aided method. Findings suggested that this method provided a more pronounced and monotonically increasing dose-response than the original RAMRIS results on synovitis and osteitis [5], especially when excluding images with poor quality scores.

This analysis method provided excellent intra-reader agreement for both MCP and wrist ROIs (0.987 and 0.988) as well as individual quantitative biomarkers ICC > 0.85. Reader variability for synovitis and osteitis has been observed periodically with the RAMRIS method, especially for scores representing changes over time [9]. Periodic systematic evaluation and standardization have been proposed [10] to ensure optimal RAMRIS performance and to optimize image quality and scoring precision. In this context, a computer-aided, quantitative methodology, such as that described here, may supplement or enhance RAMRIS reading of inflammatory changes—utilization of a Norml map that provides at least similar separation of dose response in patients with RA produces better reproducibility of the results, and requires less training for the readers (although it takes approximately twice as long to perform (20 vs. 10 min)). This method also provides continuous, reproducible, and accurate scoring of inflammatory changes that might be used efficiently even in smaller patient cohorts. Quantifying the volume and/or degree of inflammation in patients with RA using T1-weighted MRI images after Gadolinium contrast is an established concept. In 1996 and 1999, Østergaard and colleagues reported that the volume of synovial enhancement was higher in patients with swollen and tender wrist joints [11] and that the baseline volume predicted future erosive progression but also decreased significantly after conventional disease-modifying treatment [12]. More recently, several methods were tested for computer-aided synovitis and inflammation quantification in the hand and wrist of patients with RA using static postcontrast T1-weighted images [13–15]. Chand and colleagues showed that a manual computer-aided segmentation of the synovial volume in the three wrist compartments (which took 19–21 min to outline) resulted in high interobserver reliability (mean ICC 0.87) in patients with early and established RA and demonstrated scores that correlated highly to excellently with RAMRIS synovitis scores in the same wrist compartments (Spearman rho = 0.86–0.96). To our knowledge, no other studies have used the proposed contrast enhancement characterization method relative to muscle enhancement inside a rough ROI that outlines the wrist and second to fifth MCP joints. However, Boesen et al. used a similar rough ROI method to outline the same hand anatomicies using dynamic contrast enhanced MRI (DCE-MRI) data without the need of an internal muscle ROI reference and demonstrated that the volume and intensity/degree of contrast enhancement had a high correlation to RAMRIS synovitis and bone marrow edema (Spearman rho > 0.74) [16].

Additional analyses were conducted on a series of images scored as “good” and “acceptable” quality. The methodology was assessed for the sensitivity of detecting efficacy differences between the treatment arms. Images of “poor” quality were excluded to examine how image quality can impact mean change from baseline in QTVI and treatment efficacy assessment. Overall, 93% of the MRIs were considered of “good” or “acceptable” image quality. When “poor” images were excluded from the analysis, a clearer visual dose response was observed, although not a statistically significant one (Fig. 5A), illustrating the importance of achieving high-quality images at the point of image acquisition. The lack of statistically significant changes for the 4-mg dose might have been the

### Table 2

Test–retest repeatability of the quantitative biomarkers calculated.

<table>
<thead>
<tr>
<th>Region of Interest</th>
<th>Biomarker</th>
<th>Intra-reader Agreement</th>
<th>95% CI</th>
<th>( P )-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCP joints individually (MCP2–5)</td>
<td>QTVI</td>
<td>0.953</td>
<td>[0.935, 0.966]</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>QNI</td>
<td>0.853</td>
<td>[0.800, 0.893]</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>QI</td>
<td>0.956</td>
<td>[0.939, 0.968]</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>MCP total</td>
<td>QTVI-MCP</td>
<td>0.921</td>
<td>[0.847, 0.959]</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>QI-MCP</td>
<td>0.931</td>
<td>[0.866, 0.965]</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Wrist</td>
<td>QTVI</td>
<td>0.950</td>
<td>[0.901, 0.974]</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Wrist</td>
<td>QNI-Wrist</td>
<td>0.849</td>
<td>[0.719, 0.922]</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>QI-Wrist</td>
<td>0.930</td>
<td>[0.864, 0.964]</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Total (MCP + wrist)</td>
<td>QTVI</td>
<td>0.952</td>
<td>[0.906, 0.976]</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>QI</td>
<td>0.940</td>
<td>[0.884, 0.970]</td>
<td>&lt;0.001</td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** CI = confidence interval; MCP = metacarpophalangeal; QI = quantitative inflammation; QNI = quantitative mean of normalized intensity; QTVI = quantitative total volume of inflammation.
result of limiting the analysis to only “good” and “acceptable” images (i.e., only 68% to 86% of the data); however, these results could be better interpreted graphically. Results at week 24 for the 2-mg baricitinib treatment group were driven primarily by three outlier patients who appeared to have genuine degradation of their inflammatory condition over the second 12 weeks of the trial. Further analysis may be warranted with a larger group of patients to fully examine the impact of how the quality of acquired post-contrast images affects the mean change from baseline in QTVI.

This analysis was limited by several confounding factors. First, this methodology only works for post-contrast T1-weighted images and imaging artefacts affecting individual post-contrast series, such as gradient and aliasing artefacts shown in Supplemental Figures 6 and 7; this can impact inflammation quantification results. In turn, these series would be excluded from the analysis. The proposed method uses a signal intensity two standard deviations above muscle for segmentation of
imaging quality was the delay between Gadolinium injection and image acquisition, although this delay could have occurred when repositioning the hand over the coils to switch from imaging the MCP joints to the wrist or vice versa. This study protocol did not use a dynamic contrast-enhanced-MRI sequence between the pre- and post-contrast static scans and the sites did not consistently use a power injector, which is recommended to time contrast injections correctly; therefore it was difficult to control the timing of the post-contrast image. Any series without enough contrast enhancement (due to scan time being too early or very late) was excluded if the difference in the delay was noteworthy between baseline and the time point for analysis. The set-up used in this study was not ideal since it was conducted when imaging technology was less advanced; it may be necessary to focus on feasibility during a large, multicenter study, further emphasizing the importance of data quality and timing during imaging. Incorrect timing and inconsistency in image quality can be prevented by focusing on pre-training and site qualifications, using the same technicians throughout the study, using coils covering the entire hand, and including a dynamic contrast-enhanced-MRI sequence immediately following intravenous Gadolinium injection by a power injector to ensure the correct timing of the post-contrast MRI sequence [17].

5. Conclusion

The novel quantitative analysis methodology was evaluated and applied post hoc to randomized clinical trial data in patients with RA. Data quality and timing of contrast injection in relation to image acquisition was assessed and results confirmed that the computer-supported methodology allows for at least similar, if not more sensitive, quantification of inflammation and interpretable dose/response separation between treatment groups in patients with RA compared to results using the reference RAMRIS method.

CRediT authorship contribution statement

M. Boesen: Conceptualization, Methodology, Investigation, Formal analysis, Writing – review & editing. S.D. Beattie: Formal analysis, Writing – review & editing. D.E. Schlichting: Formal analysis, Writing – review & editing. O. Kubassova: Conceptualization, Methodology.

---

**Table 3**

<table>
<thead>
<tr>
<th>Score/Purpose</th>
<th>Image Category</th>
<th>Image Category Count, n (%)</th>
<th>MCP</th>
<th>Wrist</th>
<th>MCP and Wrista</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image quality</td>
<td>Number of images</td>
<td></td>
<td>416</td>
<td>415</td>
<td>415</td>
</tr>
<tr>
<td></td>
<td>Images of good (3) and</td>
<td></td>
<td>373</td>
<td>398</td>
<td>357 (86)</td>
</tr>
<tr>
<td></td>
<td>acceptable (2) quality</td>
<td></td>
<td>(90)</td>
<td>(96)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scores</td>
<td></td>
<td>43</td>
<td>17</td>
<td>58 (14)</td>
</tr>
<tr>
<td></td>
<td>Images of poor (1) quality scores</td>
<td></td>
<td>(10)</td>
<td>(4.1)</td>
<td></td>
</tr>
<tr>
<td>Scan timing delay</td>
<td>Number of images</td>
<td></td>
<td>416</td>
<td>416</td>
<td>416</td>
</tr>
<tr>
<td></td>
<td>Images of good (3) and</td>
<td></td>
<td>405</td>
<td>399</td>
<td>395 (95)</td>
</tr>
<tr>
<td></td>
<td>acceptable (2) quality</td>
<td></td>
<td>(97)</td>
<td>(96)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scores</td>
<td></td>
<td>11</td>
<td>17</td>
<td>21 (5.0)</td>
</tr>
<tr>
<td></td>
<td>Images of poor (1) quality scores</td>
<td></td>
<td>(2.6)</td>
<td>(4.1)</td>
<td></td>
</tr>
<tr>
<td>Scan timing consistency</td>
<td>Number of images</td>
<td></td>
<td>272</td>
<td>272</td>
<td>272</td>
</tr>
<tr>
<td></td>
<td>Images of good (3) and</td>
<td></td>
<td>251</td>
<td>224</td>
<td>220 (81)</td>
</tr>
<tr>
<td></td>
<td>acceptable (2) quality</td>
<td></td>
<td>(92)</td>
<td>(82)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scores</td>
<td></td>
<td>21</td>
<td>48</td>
<td>52 (19)</td>
</tr>
<tr>
<td></td>
<td>Images of poor (1) quality scores</td>
<td></td>
<td>(7.7)</td>
<td>(18)</td>
<td></td>
</tr>
</tbody>
</table>

*Abbreviations: MCP, metacarpophalangeal; MRI, magnetic resonance imaging.

a For the MCP and wrist together, good (3) quality was considered if the MCP and wrist had individual good scores, acceptable (2) quality was considered if the MCP and wrist had individual acceptable or good scores but not both good scores; and poor (1) quality was considered if the MCP or wrist had individual poor scores.

---

**Fig. 5.** Abbreviation: MCP, metacarpophalangeal. Image quality scores are assigned by experienced radiologist during reading. Labels indicate the number and percentage of a given quality score. **Left** – image quality for MCP; **Right** – image quality for wrist.
Eli Lilly provides access to all individual participant data collected during the trial, after anonymization, except for pharmacokinetic or genetic data. Data are available upon request six months after the indication studied has been approved in the United States and European Union and after primary publication acceptance, whichever is later. No expiration date of data requests is currently set once data are made available. Access is provided after a proposal has been approved by an independent review committee identified for this purpose and after receipt of a signed data sharing agreement. Data and documents, including the study protocol, statistical analysis plan, clinical study report, and blank or annotated case report forms, will be provided in a secure data-sharing environment. For details on submitting a request, see the instructions provided at www.vivli.org.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejrad.2021.109877.

References

