Update on magnetic resonance imaging and ultrasound in rheumatoid arthritis

J.F. Baker¹, P.G. Conaghan², F. Gandjbakhch³

¹Philadelphia Veterans’ Affairs Medical Center, School of Medicine, Division of Rheumatology, University of Pennsylvania Perelman School of Medicine, and Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania, Philadelphia, PA, USA; ²Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, and NIHR Leeds Biomedical Research Centre, Leeds UK; ³Universités Pierre et Marie Curie, Paris 6, Sorbonne Universités, GRC-08 (EEMOIS); APHP, Rheumatology Department, Pitié Salpêtrière University Hospital, Paris, France.

Address correspondence to: Dr Joshua F. Baker, Division of Rheumatology, University of Pennsylvania, 504 Maloney Building, 3600 Spruce Street, Philadelphia, PA, USA. E-mail: bakerj@uphs.upenn.edu

© Copyright CLINICAL AND EXPERIMENTAL RHEUMATOLOGY 2018.

Key words: rheumatoid arthritis, magnetic resonance imaging, ultrasound

Funding: J.F. Baker is supported by a Veterans Affairs Clinical Science Research and Development Career Development Award (IK2 CX000955). The contents of this work do not represent the views of the Department of the Veterans Affairs or the United States Government. P.G. Conaghan is supported in part by the National Institute for Health Research (NIHR) Leeds Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Competing interests: none declared.

ABSTRACT

Rheumatoid arthritis (RA) disease activity often remains difficult to assess and quantify accurately. As a result, numerous measures using various techniques to estimate clinical activity have been developed for clinical research and care. More objective imaging biomarkers for early detection and accurate, quantitative measurement of the disease burden are therefore of interest both for clinical use and for investigational studies. Two widely studied imaging biomarkers are magnetic resonance imaging (MRI) and ultrasound (US), imaging tests that are increasingly available to clinicians. While substantial and increasing evidence has been reported that these tools are valid and provide advantages in both clinical trials and clinical assessments, more information is needed to inform their appropriate use in routine clinical care. The goals of this review are to outline the current literature regarding each of these objective imaging tools, assess their strengths and limitations, and to clarify knowledge gaps to be filled before these techniques may be more optimally utilised.

Introduction

The assessment of disease activity has become an important part of clinical research and management of rheumatoid arthritis (RA). Measuring disease activity and targeting therapies to the achievement of low disease activity and remission? results in better outcomes (1). Clinical disease activity measures have been defined and validated for use in research and clinical practice. However, these measures are imperfect and there is no “gold standard” for assessment of RA disease activity. Clinical disease activity measures also are sensitive to joint damage and suffer from bias in the context of comorbid conditions including fibromyalgia.

These phenomena can result in inappropriate use of very expensive treatments or, conversely, a lack of recognition of sub-clinical disease. Objective tools could enhance our conceptual understanding of the disease process and, if used judiciously in clinical practice, could facilitate the quantification of the disease in ways that may add importantly to clinical assessments, reduce costs, and improve outcomes. As sensitive imaging modalities such as magnetic resonance imaging (MRI) and ultrasound (US) become increasingly available for research and clinical purposes, clinicians and researchers should understand the potential benefits and limitations of these tools.

Magnetic resonance imaging

The major strength of MRI is the precise assessment of the bony and surrounding soft tissue structures of the joint. The direct visualisation of synovial and bone involvement provide excellent face-validity. In recent years, MRI has been used with increasing frequency in clinical research and clinical trials. A validated scoring system [RA MRI Scoring (RAMRIS)] has been applied successfully to assess efficacy of therapies in a number of clinical trials (2-4). Furthermore, MRI features have been shown to have diagnostic and prognostic value, suggesting strong construct validity (5, 6). In 2013, a Task Force of the American College of Rheumatology examined the reported data and concluded that MRI measures met the conditions of the Outcomes Measures in Rheumatology (OMERACT) filter for truth, discrimination, and feasibility in measuring relevant structural outcomes for randomised clinical trials of RA (7).

Validity of MRI

Over the past decade, scoring systems have been developed and validated to quantify both inflammatory features
MRI and ultrasound in RA / J.F. Baker et al.

of the disease as well as damage. In terms of inflammatory manifestations of the disease, MRI can be used to quantify synovitis, tenosynovitis, and bone marrow edema (BME) at the hand and wrist. While most studies quantify these features by review of an expert reader, using these validated scoring systems (i.e. RAMRIS), automated quantitative approaches using Dynamic Contrast-Enhanced MRI (for synovitis) and statistical shape modelling have also been studied and performed similarly to RAMRIS measures (8-12). To date, all of these quantification systems have been confined to use in clinical research studies. It is important to note that MRIs performed in clinical care typically would provide only a local radiologist’s description of these abnormalities, not a quantitative score. Techniques that utilise automated quantitative approaches may therefore eventually be of value but are not currently clinically available.

Synovial hypertrophy detected on MRI has excellent construct validity. Synovitis on MRI is typically defined as hypertrophy, effusion and/or synovial tissue enhancement post intravenous (IV) contrast (e.g. gadolinium), suggesting increased vascularity (13, 14). There is a clear advantage the use of gadolinium in the visualisation of synovitis, however, it is not clear whether it should be recommended that gadolinium be used for clinical use given the increase in cost, invasiveness, and potential for side effects. MRI synovitis is correlated significantly with synovial inflammatory activity on histological assessment, macroscopic findings on mini-arthroscopy, clinical disease activity measures, and patient-reported symptoms (15-19).

Tendon sheath inflammation also is common in RA, and an MRI scoring system for tenosynovitis with good reliability has been introduced and validated (20, 21). As with synovitis, estimates of tenosynovitis on MRI are correlated significantly with relevant patient-reported outcomes (22) (further information concerning tenosynovitis is found in an article by van der Helm in this supplement). Bone marrow oedema (also called osteitis) is a finding that is more unique to MRI and is observed when a lesion in the trabecular bone has signal characteristics consistent with increased water content (23). BME lesions in RA are comprised of cellular infiltrates in the subchondral bone, suggesting that they represent inflammatory lesions (24). These lesions likely represent precursor lesions to subsequent bone erosion (6, 25, 26). MRI also can be used to quantify damage to the joints including bone erosion and joint space narrowing (JSN) in the wrist and metacarpophalangeal joints (MCPs) (27). The RAMRIS methods of quantification have been validated for scoring joint abnormalities and have both good inter-/intra-reader reliability and excellent construct validity, demonstrating excellent correlation with traditional radiographic methods (27-30). MRI is a more sensitive method than conventional radiography, due to the three-dimensional acquisition, and has been shown to have a high concordance rate when compared with computerised tomography (31).

Clinical and research application of MRI measures of disease activity
Several studies have demonstrated that measurement of synovitis and BME can provide prognostic value and that these measures can predict radiographic changes on x-ray. In particular, previous studies have demonstrated MRI BME to be the strongest predictor of subsequent erosions on conventional x-ray (32, 33). Thus, MRI provides prognostic information that may help to inform risk stratification and early treatment decisions. It has also been shown that early changes in measures of synovitis and BME at 12 and 24 weeks were seen in patients who received rituximab compared to patients who received placebo. In this small study, the change in DAS28 was not significantly between treatment groups at these time-points, suggesting greater sensitivity of MRI to recognise treatment effects. Tenosynovitis is also a responsive outcome in the context of active treatment. For example, the OPERA study (OPtimised treat-ment algorithm in Early Rheumatoid Arthritis) indicated significant improvements in tenosynovitis from baseline in both treatment groups (4). These studies, among others, suggest that MRI measures of inflammation could be used in Phase 2 clinical trials, and potentially in routine clinical care, to identify individuals with early responses.

MRI inflammatory features also are observed in patients with RA in clinical remission and low disease activity (37, 38). This subclinical inflammation is predictive of progressive x-ray damage (38). Thus, in certain clinical contexts, MRI may provide important information that could suggest the need for more aggressive management. However, the benefits of escalation of treatment for subclinical disease are not yet clear. Conversely, the capacity to define “low MRI activity” may help to identify patients who are at low-risk of structural damage progression. Since many patients are unable to achieve low clinical disease activity due to comorbidity (39), it is critical to identify those individuals where escalation of therapy is not needed despite moderately elevated clinical disease activity scores. Low imaging scores might potentially be used in the future to help to reassure providers to monitor patients and hold off on escalation of therapies in certain circumstances. Two prior studies have defined RAMRIS thresholds for low MRI activity for synovitis and osteitis (40, 41). However, much work is needed to extrapo-
ulate these research observations to use in clinical decision-making. While extensive evidence supports MRI as a powerful research tool, it remains unclear how best to utilise MRI in clinical care, either alone or in combination with clinical disease activity measures. Data regarding MRI in real-world settings and its impact on clinical decision-making are limited. A limitation of non-standard MRI is that studies have shown that low grade abnormalities are commonly observed among non-RA individuals, including among those with osteoarthritis, suggesting that clinicians may need to take great care to not over-diagnose active RA inflammatory lesions (42). It is notable that quantification of low-grade abnormalities would likely fall below previously proposed thresholds for “low MRI activity”, suggesting that accurate quantification may overcome this problem. Some studies have shown that use of a reference group to identify “abnormal” can also reduce false-positive results (43). To better deal with the problem of overly sensitive techniques, it will be critical to have standardisation and/or automation of reading techniques, and dissemination of such techniques with adequate characterisation of normal variation in non-RA populations.

Clinical and research application of MRI joint damage
Assessments of erosion and JSN on MRI using the RAMRIS method are also highly sensitive and sensitive to change (27, 44-47). For example, in one study, significant differences were seen between the rituximab active treatment group and placebo group at 24 weeks in measures of bone erosion, while no significant differences were seen using conventional x-rays (36). Other studies have demonstrated that MRI measures of bone erosion may identify differences between treatment groups in earlier time points and in fewer subjects compared to x-ray (48). Utilisation of MRI measures of bone erosion in clinical trials would reduce calculated sample sizes and follow-up times if used as the primary outcome (45). In the Impact of Rituximab on Magnetic Resonance Imaging Evidence of Synovitis and Bone Lesions in Patients with Moderate or Severe Rheumatoid Arthritis (IMPRESS) study, significant changes also were also seen in JSN as measured by MRI over 24 weeks, while no significant change was demonstrated measuring JSN using x-ray. Other studies have also demonstrated highly significant differences in a measure of cartilage loss between treatment groups, while no significant differences in JSN were seen by x-ray, suggesting greater discrimination using MRI (36). In addition to being sensitive to change, early MRI progression is associated with later x-ray progression, suggesting that these changes are important and have prognostic value for structural damage (34). The sensitivity of MRI for RA joint damage means careful attention is required if utilised as a diagnostic test. It has been noted that erosions noted on MRI can be observed in non-RA controls and that there is significant overlap in the setting of early RA (42, 49). As a result, erosions noted on MRI in clinical evaluations may lack specificity. Evaluation for certain types and locations of erosions may help overcome this limitation. One study found that grade ≥2 erosions and erosions at the 5th MTP remained specific for RA (specificity >89%) (49). Erosions at the 1st MTP also were specific among patients under the age of 40 (specificity 93%).

Interestingly, the presence of erosions combined with inflammation were not specific, suggesting that evidence of inflammation in this context is not specific to the cause of the erosion. Clinicians should carefully consider these issues when interpreting results of MRIs that are performed in clinical care to prevent over-diagnosis. For example, a patient with hand osteoarthritis may have significant synovitis and bone erosions in the affected joints. Modern imaging does not replace the need for careful consideration of the clinical context based on the patient history and physical examination. Overall, while the use of MRI to assess damage in RA is promising, more work is needed to better characterise appropriate use of the technology in clinical practice to ensure the judicious use of MRI and to limit the over-emphasis on the importance of minor abnormalities and minor changes that may occur over time. As MRI becomes increasingly available to clinicians, guidelines for interpretation of these imaging studies will be paramount to ensure thoughtful and accurate interpretation of this powerful tool. Further research to inform development of existing management guidelines, such as those developed through EULAR, will help to inform clinical use (50).

Ultrasound
Interest in the use of ultrasound in RA has increased considerably in recent years. Thanks to improvements in technology, musculoskeletal US provides real-time, high resolution images for assessment of superficial structures (where it has an acoustic window) and allows detection of both inflammatory activity (synovitis, tenosynovitis) and structural damage (erosion and joint space narrowing). Ultrasound presents a lot of advantages: it is patient-friendly, low cost, not invasive, and uses no ionising radiation. In daily practice, ultrasound has the advantage of being more sensitive for detection of erosions compared to x-rays and more importantly, allows detection of joint inflammation in times when we are focusing on pre- or early RA. When compared to MRI, US has the capacity to assess multiple joints (for instance hands and feet) during the same examination, whereas MRI evaluation generally focuses on one location (whole body MRI is still undergoing validation). EULAR recommendations for RA imaging highlight that US is superior to clinical examination to detect inflammation, and could be useful for both RA diagnosis and monitoring (50). Recently a group of experts has proposed algorithms for the use of US in RA (51).

Validity of ultrasound measures of disease activity and damage
Most studies have focused on the validity of US to detect inflammation, especially synovitis. Different studies have demonstrated that US appears to have good criterion and construct validity.
US-synovitis is correlated significantly with histological scores for vascularity and inflammation, and with acute phase reactants (52-54), although the correlation between US-synovitis and C-reactive protein is not consistently observed, perhaps due to differences in the methods of quantification. Comparison with other imaging modalities and clinical examination have also demonstrated good construct validity of US for the detection of synovitis (55-57). While there was (expectedly) moderate agreement between clinical examination and US for detection of synovitis (58-61), the agreement between US-synovitis and MRI-synovitis appeared good, especially for small joints (56, 62, 63). Like MRI, US has higher sensitivity compared to the clinical examination to detect synovitis (57, 64). Furthermore, it allows detection of tenosynovitis which may be difficult to assess by clinical examination (63).

US presents also good validity for detection of erosion; different studies have shown a good agreement between US and conventional radiography, computerised tomography and MRI to detect erosions in small joints (55, 57, 65). In particular, US has higher sensitivity compared to conventional radiographs to detect erosions, especially in early RA (65, 66).

Ultrasound and RA diagnosis
Different authors have studied the added value (over clinical examination) of US for the diagnosis of RA. Filer et al. showed that the addition of US to clinical examination in patients with early RA improved the Leiden score (Leiden score: p<0.001; AUC 0.905 vs. Leiden score and PD10 (Naredo-knee): p<0.006; AUC 0.962) (67). In a cohort of 109 early arthritis patients followed during one year, Nagakomi et al. demonstrated that grade ≥1 had higher sensitivity compared to clinical examination for the diagnosis of RA (USGS1: sensitivity 78%, specificity 79.4% vs. clinical examination: sensitivity 58.5% specificity 79.4%) and that higher US grade of synovitis (GS ≥2, PD ≥1) provided almost the same sensitivity than clinical examination with higher specificity (sensitivity 56.1%, specificity 93.7%) (68). The authors reported that US was particularly useful in patients with suspected RA who did not meet the ACR-EULAR criteria for RA. For these patients, the use of US improved the AUC for the diagnosis of RA compared to clinical examination (Clinical: 0.457(0.292–0.622); US GS ≥1: 0.736 (0.595-0.878); US GS ≥2, PD ≥1: 0.800 (0.673–0.927) (68).

Questions on use of US in routine practice remain. There is no consensus concerning the number and sites of joints to assess for diagnosis and monitoring of RA. Most authors have studied wrists, MCP,PIP and MTP joints. It is possible that ultrasound examination for RA diagnosis should include a higher number of joints when compared to RA assessments in follow-up. Moreover, the threshold for the grade of synovitis which should be taken into account to distinguish pathological and physiological findings remains uncertain. Recent studies have shown that low grade synovitis (grade 1 for both B mode and Doppler) can be detected in healthy subjects (likely some of whom have osteoarthritis), especially at some anatomical sites (wrist, MTP 1 to 4) (69). Interpretation of US findings therefore should be taken with caution and importantly (as for MRI diagnosis) in clinical context.

In a study including early arthritis patients from the ESPOIR cohort and age/sex matched healthy subjects, Milloet al. showed that the presence one joint GS-synovitis (grade 1 for both B mode and Doppler) can be detected in healthy subjects (likely some of whom have osteoarthritis), especially at some anatomical sites (wrist, MTP 1 to 4) (69). Interpretation of US findings therefore should be taken with caution and importantly (as for MRI diagnosis) in clinical context.

Ultrasound and RA monitoring
Ultrasound has been validated for evaluation and monitoring of inflammatory activity in RA for both synovitis and tenosynovitis with good reliability and sensitivity to change. Although a common criticism of the use of US in RA is that it is an operator-dependent technique, the reliability of US for the detection of synovitis (kappa = 0.61-0.97) is as good as the reliability of clinical examination (kappa = 0.53-0.82) and metrological properties (construct validity and sensitivity to change) are at least equivalent (58).

Different scoring systems have been proposed, mostly semi-quantitative from 0 to 3 for individual joints (normal, mild, moderate or severe) for B mode and Doppler mode. Recently, the OMERACT PDUS score has been proposed, which has the advantage to be applicable in all joints, whereas previous scores like Szkudlarek’s score has been validated only in particular joints [metacarpophalangeal joints, proximal interphalangeal joints and metatarsophalangeal joints (71, 72)]. The OMERACT PDUS score for synovitis has shown good sensitivity to change, with early responsiveness post-therapy (73).

The OMERACT group has also proposed and validated a semi-quantitative scoring method from 0 to 3 (normal, mild, moderate and severe) for the evaluation of tenosynovitis with good reliability and sensitivity to change (74-76). These OMERACT scoring systems for synovitis and tenosynovitis are currently being used in clinical trials and could be considered as reference for the evaluation of US inflammatory activity.

There remains no consensus concerning the number of joints to assess during RA monitoring. Different studies have shown that reduced joint US evaluation could be proposed with good correlation with a more extended joint US evaluation, with the advantage to increase feasibility, as it requires less time. This problem of feasibility remains an issue for implementation of US in daily practice. Mandl et al. showed that the feasible scoring systems developed by Naredo et al. (12 joints: bilateral wrist, MCP2, MCP3, knee, ankle and elbow) and by Backhaus et al. (7 joints: unilateral wrist, MCP2, MCP3, PIP2, PIP3, MTP2 and MTP 5) had good metrological properties including for sensitivity to change, and could be proposed for RA monitor-
ing (77). D’Agostino et al. compared different reduced US evaluation and concluded that they all provided good sensitivity to change (73). The sites of tendon assessments in RA monitoring is less controversial as the OMERACT group recommended these sites at the time of creation of the scoring system for tenosynovitis (76).

Detection of both US synovitis and tenosynovitis is important, as both appear to be independently predictive of structural damage (78-80). A group of experts proposed that targeting therapy to PD-activity provides superior outcomes compared with treating to clinical targets alone, and introduced the rationale for new randomised trials using targeted US in RA (81). However the role of US monitoring in RA daily practice has been controversial. In the TASER study, the use of US for monitoring a cohort of early and active RA did not result in a higher frequency of patients in remission as compared to clinical evaluation (82). However, this study applied US to everyone in the US arm of the study, and did not apply US only to cases in which there was clinical uncertainty about disease activity – which may be the setting in which US should be used in routine practice to guide therapy. So this study may not have answered the question about which patients would benefit from US, but has suggested US should not be used in all patients in whom a tight-control treatment strategy is being implemented.

There is again no consensus concerning US monitoring of structural damage. While US RA erosion has been well defined, there is no consensus on how erosion size should be measured, nor if a threshold for the size of the erosion should be used or not (83).

US may detect subclinical inflammation in RA patients in clinical remission with synovitis demonstrated in B mode and in Doppler mode (84). Different studies have shown that this subclinical inflammation detected by US is predictive of flare and of structural damage progression in patients with RA (38, 85-88). These data suggest that the use of US may be valuable in the management of RA patients in apparent clinical remission, especially when tapering is considered. Presence of Doppler-detected synovitis may predict biologic tapering failure in RA patients in sustained clinical remission (85).

The ARTIC study is a randomised trial examining the benefit of ultrasonography in a clinical tight control regimen aiming for remission in rheumatoid arthritis (DAS<1.6, absence of swollen joints and no radiographic progression) (89). No difference for the primary endpoint was seen between the groups after 24 months follow-up. Yet, a tendency for lower radiographic progression was observed in the US-group at 24 months (p=0.05). A sub-analysis of the ARCTIC study showed that an association between 6-month remission and no radiographic progression was observed for ACR/EULAR Boolean remission (44 joints, OR 3.2, 95% CI 1.2 to 8.4), absence of ultrasound power Doppler (OR 3.6,95% CI 1.3 to 10.0) and grey-scale remission (OR 3.2,95% CI 1.2 to 8.0) (90).

Conclusions

In summary, the recent and growing literature on MRI and US warrant consideration of their use in both clinical trials and in routine care. Currently MRI appears the tool of choice for clinical trials, offering central reading and smaller, shorter duration studies compared to those with radiographic structural endpoints, and objective assessment of pre-erosive inflammatory changes. The role of ultrasound in clinical trials may become more apparent with ongoing trials to be reported in the next year. In terms of routine care, both tools have a role in diagnosis in determining subclinical inflammation and the extent of such inflammation. However, US is often more feasible. Their role in monitoring RA should be considered carefully in patients in whom there is clinical uncertainty over disease activity, especially in low disease activity states, though much of these benefits remain to be established in robust studies. However, if used appropriately, both tools are likely to positively impact care of patients with RA and positively impact health care systems.

Take home messages

- While MRI measures of inflammation and damage have good construct validity and responsiveness, and provide prognostic value above clinical assessments in a research context, strategies to employ them in the clinic have not been widely studied.
- Ultrasound is highly feasible for well-trained users in the clinic, adds value for a proportion of diagnostic cases, and may be useful in monitoring disease if there is clinical uncertainty about disease activity and in clinical low disease activity states.

References

8. MACISAAC KD, BAUMGARTNER R, KANG J et al.: Pre-treatment whole blood gene ex-


42. MANGNUS L, VAN STEENBERGEN HW, REINIERSE M, VAN DER HELM-VAN MIL AH:


70. MANDL P, NAREDO E, WAKEFIELD RJ, CON- Aghnan PG, D’AGOSTINO MA, OMERACT ULTRASOUND TASK FORCE: A systematic lit- erature review analysis of ultrasound joint count and scoring systems to assess syno-


