Review

Advances in Imaging for Clinical Trials in Rheumatic Diseases

Walter P. Maksymowych

University of Alberta, Edmonton, Canada; CARE Arthritis, Edmonton, Canada

The successful execution of clinical trials for novel anti-rheumatic compounds is increasingly approaching the limits of what can be achieved using radiographic outcomes for the assessment of disease modification. Moreover, there is a growing need for more objective tools to assess joint inflammation, especially for disorders such as axial spondyloarthritis where spinal symptoms are often non-specific and physical findings may be minimal until later stages of disease. The use of MRI to evaluate inflammation in the synovium and bone marrow as well as erosions in peripheral joints of patients with RA and PsA represents a major new advance that should now be routinely implemented in clinical trials of RA. MRI-based scoring systems have been well validated and demonstrate that, for RA, MRI changes after therapeutic intervention may be observed in a month and precede findings on radiography that only become evident after a year. The assessment of disease activity on MRI of the sacroiliac joints and spine using a standardized and well-validated method, such as the SPARCC instruments, is indispensable to the evaluation of efficacy for new agents aimed at the treatment of spondyloarthritis. Further advances include the use of whole-body MRI evaluation to assess inflammation in both the axial and peripheral skeleton as well as sequences that dispense with the requirement for the use of contrast agents, such as gadolinium, and data processing techniques that permit full automation and absolute quantification. This review will discuss how imaging is transforming clinical trials in rheumatic diseases.

Keywords: Magnetic Resonance Imaging, Rheumatoid Arthritis, Spondyloarthrits, Ankylosing Spondylitis, Psoriatic Arthritis, Synovitis, Effusion, Bone Marrow Lesion, Validation, Scoring System, Clinical Trials.
Introduction

The past two decades have witnessed unprecedented development of novel therapeutics for chronic inflammatory joint disorders such as rheumatoid arthritis (RA), psoriatic arthritis (PsA), and axial spondyloarthritis (axSpA). This has led to regulatory authorities granting indications for amelioration of symptoms as well as structural damage disease-modification labels for numerous new agents. These advances have been facilitated by clinical trial designs that include an array of validated outcome measures evaluating diverse clinical outcomes, but also objective outcomes provided by imaging. However, the availability of this increasing number of effective therapeutics has led to growing challenges in the conduct of clinical trials due to factors such as poor recruitment of patients, high placebo response rates, ethical constraints on the duration of the placebo-controlled phase, and the large patient sample size and lengthy duration of follow up required to provide evidence of disease modification when the degree of structural progression is small. There is evidence for the emergence of similar challenges for clinical trials of chronic disorders in other fields of medicine. Moreover, there has been an increasing focus on the evaluation of a variety of patient self-reported measures considered by patients with arthritis as being of particular importance to their well-being, such as pain and fatigue. But these are often non-specific and may be influenced by non-disease related factors. Objective measures reflecting the pathophysiology of disease are therefore necessary, and these become even more important in a condition such as axSpA because of the frequency of non-specific back pain in the general population and the lack of disease-specific findings on physical exam. Advances in the development of outcomes based on imaging increasingly demonstrate how we might overcome these challenges. What are the advances in imaging outcomes that have already been evaluated in chronic inflammatory joint disease, how have they been used to enhance the design and outcome of clinical trials, what are their limitations, and what can we expect for future enhancements?

IMAGING OUTCOMES IN RHEUMATOID ARTHRITIS

Radiography

A key requirement for a regulatory authority to grant a label for structural disease modification in RA is evidence that the new agent can prevent, halt, or reverse the structural damage to joints using standardized radiographic scoring methods, such as the Sharp/van der Heijde score (mTSS) [1,2,3]. This is a composite assessment of joint space narrowing, indirectly denoting loss of articular cartilage, and joint erosion in the joints of the hands and feet with a scoring range of 0-448.

A sample size of 1500 patients was estimated in a recent trial of upadacitinib, a Janus Kinase 1 (JAK1) selective inhibitor, versus placebo or adalimumab, a tumor necrosis factor inhibitor (TNFi) monoclonal antibody, in patients with RA and an inadequate response to methotrexate (MTX), the primary conventional synthetic disease modifying anti-rheumatic drug (csDMARD) used for RA, to provide at least 90% power to detect treatment differences compared to placebo in the primary and several key secondary endpoints, including a difference of 0.39 between upadacitinib and placebo in the change from baseline in the mTSS at week [4]. The duration of the placebo-controlled portion was 26 weeks but for ethical reasons, initiation of rescue therapy occurred after the primary endpoint at week 12 between weeks 14 and 26 in 47%, 24%, and 19% of patients in the placebo, adalimumab, and upadacitinib groups, respectively. Linear extrapolation was applied for the handling of missing radiographic data and for patients who initiated rescue therapy at week 14. The mean change from baseline in the mTSS was 0.97 for the placebo plus background methotrexate group and 0.24 for the upadacitinib plus background MTX group, a statistically significant difference. The total recruitment period for this trial across 286 sites in 41 countries, including Ukraine, was about 18 months.

This trial is typical of current pivotal phase 3 trials in RA. The placebo-controlled phase is ethically restricted to 24 weeks with rescue therapy from 14 weeks, the change in the structural damage endpoint over this time frame is very small, despite including patients at higher risk for developing joint damage, and the change is of no clinical significance. This necessitates a costly study with large numbers of participants where a substantial proportion of participants are unable to remain on their original randomized treatment assignment and substantial data extrapolation is required for patients receiving rescue therapy. FDA regulatory guidance has not changed since 2013 while EMA guidance was last updated in 2018 [1,2]. The EMA guidance requires that an “anti-inflammatory effect is demonstrated unequivocally by the above-mentioned composite endpoints, accompanied by observational data of X-rays providing reassurance that structural bone damage does not deteriorate during treatment, e.g. in comparison with an active control with established efficacy regarding the prevention of structural damage.” This guidance leaves the door open to misinterpretation and suggests a non-inferiority study design versus an active control would be acceptable.
Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) has been well established as providing superior sensitivity to conventional radiography for the measurement of structural joint damage in RA clinical trials [5-12]. Moreover, it can detect and quantify synovial and bone marrow inflammation. The Outcome Measures in Rheumatology Clinical Trials (OMERACT) RA MRI score (RAMRIS) is a semi-quantitative scoring method that evaluates inflammatory (i.e., osteitis/bone marrow edema, synovitis, tenosynovitis) and destructive changes (i.e., erosions, cartilage loss/joint space narrowing (JSN)) in the metacarpophalangeal (MCP) joints, hand, and wrist [13]. A study of a randomized placebo-controlled trial of MTX naïve patients who had MRI evaluation compared placebo plus MTX, golimumab (a TNFi monoclonal antibody) 100 mg plus placebo, golimumab 50 mg plus MTX, or golimumab 100 mg plus MTX every 4 weeks, and used the RAMRIS to assess synovial inflammation, bone edema/osteitis, and erosion at weeks 12 and 24 [14]. The mTSS was used to assess radiographic progression at week 28. Combined therapy with golimumab plus MTX versus placebo plus MTX significantly improved RAMRIS scores for synovitis, osteitis, and erosion as soon as 12 weeks. By comparison, there was no significant difference between these groups in the mTSS score at week 28. A significant difference in the mTSS score between these groups was observed in the total trial population but required double the number of patients (637 versus 318) and double the length of followup (28 versus 12 weeks) as needed for MRI to demonstrate this. MRI measures of inflammation and structural damage correlated independently with physical function, pain and patient global assessments [15].

Additional trials provide evidence of the ability of MRI to discriminate progression and treatment effect in a more effective manner than radiography [14,16-24]. In the JADA study of baricitinib, dose-dependent suppression of synovitis, osteitis, erosion, and cartilage loss was evident with MRI within 12 weeks, despite a sample size of fewer than 30 patients per active treatment arm [24]. One study evaluating an interleukin-6 (IL6) targeted therapy even demonstrated significant suppression of synovitis by 2 weeks [22]. The change in MRI findings also correlated with changes in lab parameters of disease activity [25] and predicted subsequent radiographic findings [14,16,21,23,26]. Erosions detected by MRI have been found to correlate well with the findings from direct visualization during arthroscopy and computed tomography (CT) [7,27,28]. Detection of synovitis by contrast-enhanced MRI has also been correlated with macroscopic evidence of synovitis by arthroscopy and histopathologic evidence of synovial inflammation [29,30]. Osteitis detected by MRI closely correlated with postsurgical histologic findings of osteitis [31,32].

Is the use of MRI feasible for the widespread implementation required for clinical trials? Several multicenter trials cited in this review have demonstrated that the quality of images acquired is sufficient to ensure reliability of scoring evaluations between central readers. The greater sensitivity of MRI to detect structural joint damage means that fewer patients are required together with shorter duration of exposure to placebo. The higher costs of MRI evaluation are offset by lower costs for patient recruitment and evaluation during the trial.

There have been recent adaptations of RAMRIS that may further increase feasibility. RAMRIS scores MRI lesions in 23 joints of the dominant hand and wrist and therefore consumes time and resources. A streamlined MRI score, RAMRIS-5, focuses on only 5 joints of the hand and wrist, and has been shown to correlate substantially with the full RAMRIS [33,34].

Additional studies have explored omission of the use of gadolinium. Intravenous administration of this contrast agent results in its passage into the interstitial space and the degree of tissue penetration is dependent on blood flow and local microvascular permeability. Inflamed tissues, such as the synovium in RA, demonstrate enhancement of signal intensity on T1-weighted post-gadolinium injection images. Gadolinium is therefore administered intravenously to demonstrate the degree of synovitis that can be readily quantified on MRI. However, the use of this agent markedly prolongs the MRI examination time for the patient, thereby increasing patient discomfort, significantly increases costs, and has raised safety concerns [36]. Omission of gadolinium would also permit the evaluation of additional joints and further increase the information derived from the examination. Additional MRI sequences that do not require administration of contrast, such as T2-weighted fat-saturated (T2 FS) and short-tau inversion recovery (STIR) sequences, are fluid-sensitive and thereby allow the display of areas with high water content, such as edema and inflammatory cell infiltration, which appear bright on the MR image. Furthermore, bone marrow inflammation can also be visualized using these sequences due to suppression of the signal from fat that may be physiologically present in bone marrow. One study compared the two types of MRI sequences using 0.2 Tesla dedicated extremity MRI and demonstrated that omission of gadolinium did not change RAMRIS scores for bone erosions and bone edema but decreased the sensitivity and reliability of RAMRIS synovitis scores [35]. Pilot studies have evaluated gadolinium free
STIR and 3D gradient echo sequences and termed this version RAMRIS Sine-Gadolinium-For-Experts (RAMRISSAFE) [37]. High correlations between the full RAMRIS and RAMRIS-SAFE synovitis subscores was observed, which might be explained by the higher field strength of 3Tesla used in this study, but the correlation for tenosynovitis was only modest [38]. This is a promising direction requiring larger studies comparing MRI sequences with and without gadolinium.

A significant limitation of the semi-quantitative RAMRIS scoring method is that its reproducibility relies on trained and calibrated readers. A fully automated RAMRIS scoring method, the RA MRI Quantification (RAMRIQ) system, has been developed (Imorphics, Manchester, UK) which makes use of computer vision Active Appearance Models (AAMs) and Active Shape Models (ASMs) [39,40] to generate bone surfaces for all the bones in the hand and apply anatomical masks which correspond to the synovial capsule of each joint. These 3D regions of interest are then examined for the volume of synovitis which enhances after application of gadolinium, indicating active joint inflammation. Similar principles are used to assess the volume around tendons that enhances after gadolinium, quantifying tenosynovitis, and the volume of bone that is bright on a fat-suppressed sequence, quantifying bone marrow edema. A multivariate analysis of an early RA population recruited to a phase 2 clinical trial comparing tofacitinib 10 mg BID monotherapy, tofacitinib 10 mg BID with MTX, or MTX monotherapy, demonstrated that RAMRIS erosion change at 1 month, as well as RAMRIQ osteitis changes at months 1 and 3, were significant independent predictors of 1-year radiographic progression [41]. Conversely, clinical changes in disease activity at 1 and 3-months did not predict radiographic progression. RAMRIQ may be helpful in phase 2 development to determine whether a novel agent is likely to be efficacious and also to provide information on appropriate dosing.

It is important to emphasize that RAMRIS and its variants have been primarily assessed in clinical trial settings where image acquisition parameters are stipulated, and quality assurance has been implemented to ensure correct implementation of the imaging. These methods are not intended for use in clinical practice, where implementation of the RAMRIS/RAMRIQ protocol may be subject to variability due to different magnetic field strengths and potential inter-operator differences.

**IMAGING OUTCOMES IN PSORIATIC ARTHRITIS**

**Radiography**

Regulatory guidance for granting a label for structural disease modification in PsA has been more limited. The FDA has not issued a guidance document specifically for PsA. The last guidance issued by the EMA was in 2006 and this was non-committal with respect to scoring methods but still recommended radiography to assess structural modification [42].

Clinical trials of disease-modifying therapies for PsA encounter the same limitations as those using radiographic evaluation in RA trials. A typical contemporary example was the phase 3, double-blind, placebo-controlled study of guselkumab, an IL-23 inhibitor that specifically binds the IL-23 p19 subunit. It was conducted in biologic-naive patients from 118 sites in 13 countries (including Ukraine) with active psoriatic arthritis (at least five swollen joints, at least five tender joints, and C-reactive protein ≥0.6 mg/dL) despite standard therapies [43]. Patients were randomized in a 1:1:1 ratio to receive guselkumab every 4 weeks; guselkumab at week 0, week 4, and then every 8 weeks; or placebo over 24 weeks. This trial evaluated radiographs at baseline and 24 weeks using the van der Heijde modified-Sharp score modified for psoriatic arthritis (distal interphalangeal joints of hands added, (range 0–528)) (mTSS-PsA) [44]. Required patient sample size was estimated to be 684 patients for 90% power to detect a treatment difference in change from baseline in the mTSS-PsA, assuming mean changes from baseline at week 24 of 0.9 in the placebo group and 0.3 in the guselkumab groups. At week 16, 63 patients met the criteria for early escape to additional therapy having failed to meet pre-specified criteria for minimal improvement, while an additional 23 withdrew from therapy before 24 weeks. Patients treated with guselkumab every 4 weeks, but not those receiving guselkumab every 8 weeks, demonstrated significantly less progression of structural damage as reflected by the least squares mean change of 0.29 versus 0.95 in the placebo group.

The minimal progression of only 0.95 in the placebo group (out of a maximum score of 528) and the continuing improvement in availability of effective therapies for PsA indicates that radiographic progression of joint damage in the placebo group will be even less in future trials and will require increasingly larger sample sizes rendering trials using radiographic outcomes to assess disease-modification non-feasible.
Magnetic Resonance Imaging

A similar MRI-based scoring method to RAMRIS has been developed for the assessment of inflammation and structural damage in PsA by the Outcome Measures in Rheumatology (OMERACT) MRI Inflammation Group and named the Psoriatic Arthritis Magnetic Resonance Image Scoring System (PsAMRIS) [45]. This method scores synovitis, edema, tenosynovitis, periarticular inflammation, erosion, and bone proliferation on MRI in the hands and feet and differs from the RAMRIS in scoring lesions in the distal interphalangeal joints as well as periarticular inflammation and new bone growth at enthesal locations.

Experience with this method in clinical trials is far more limited than for the RAMRIS method. In a recent small placebo-controlled trial of 40 patients with PsA evaluating abatacept, a biologic disease-modifying therapy impairing co-stimulation of T-cells by antigen-presenting cells, the PsAMRIS method showed good reliability and demonstrated change in inflammatory scores [46].

Patients with PsA have diverse musculoskeletal manifestations that include enthesal inflammation, tenosynovitis, and arthritis of both axial and peripheral joints. More recently, a whole-body MRI scoring method has been developed to assess arthritis and enthesitis and the methodology is detailed at www.carearthritis.com/mriportal/wbmri/index/[47,48]. This method may detect inflammation in PsA at both joints and entheses before it is evident clinically. A total of 83 joints are scored separately for synovitis and osteitis, and 33 entheses are scored separately for soft tissue inflammation and osteitis [49,50]. The first data using this method in an open label trial of apremilast was recently presented [51]. Significant amelioration of peripheral and enthesal inflammation was demonstrable after 24 weeks of treatment.

The OMERACT MRI in Arthritis Group has also recently developed a Heel Enthesitis MRI Scoring System (HEMRIS) which scores enthesal lesions at the Achilles and plantar fascia insertions by evaluating bone marrow edema, retrocalcaneal bursitis (Achilles insertion only), intratendinous and peritendinous inflammation [52]. Details of the scoring methodology are provided at www.carearthritis.com/mriportal/heelenth/index/.

A detailed MRI-based scoring system for assessment of inflammation in the hip and knee in osteoarthritis has been developed and can also be used in PsA and other inflammatory disorders affecting these joints [53-55]. This constitutes an important unmet need because clinical assessment of hip joints for inflammation is limited and insensitive while involvement of weight-bearing joints in PsA contributes greatly to functional impairment. Moreover, radiographic evaluation is only capable of detecting structural lesions once damage is already extensive. The Hip Inflammation MRI Scoring System (HIMRISS) and the Knee Inflammation MRI Scoring System (KIMRISS) scoring methods have now been formally validated in osteoarthritis [56,57] and methodological details are outlined in detail at www.carearthritis.com/service/osteoarthritis-imaging/.

Bone marrow lesions (BML) in the hip are assessed on fluid-sensitive MRI scans using an electronic overlay on a web-based interface positioned over the femoral head and extending into the acetabulum on consecutive coronal slices through the hip joint (scoring range 0–100) (Figure 1) [57]. It has been shown to have very good to excellent reliability for detecting change in BML for patients with OA receiving glucocorticoid injections [58]. In the KIMRISS method, BML is scored using specific web-based electronic overlays for the femur, tibia, and patella on consecutive sagittal slices through the knee joint (scoring range 0–500) (Figure 2) [56]. It has also been shown to have very good reliability for detecting change in BML in patients with knee OA followed for 1 year in the Osteoarthritis Initiative study and in a pilot evaluation of adalimumab for inflammatory knee OA [54,56]. Synovitis effusion is scored on consecutive sagittal slices of fluid-sensitive non-contrast-enhanced sequences in all compartments of the knee and has a scoring range of 0-100 (Figure 3). Both methods have recently been validated by the OMERACT MRI in Arthritis Group for scoring inflammation in hip and knee joints on WB-MRI scans of patients with SpA [59,60].
Figure 1. The web-based overlay used in the Hip MRI Inflammation Scoring System (HIMRISS) is positioned over the cortex of the femoral head and the femoral portion is divided into 9 sectors while the acetabular portion is divided into 3 sectors [57]. The presence/absence of bone marrow lesions is scored for each sector by mouse clicking over each sector (for further details go to www.carearthritis.com/service/osteoarthritis-imaging/)

Figure 2. The web-based overlays used in the Knee MRI Inflammation Scoring System (KIMRISS) are positioned over femoral, tibial, and patellar articular bone [56]. The presence/absence of bone marrow lesions is scored for each sector by mouse clicking over each sector (for further details go to www.carearthritis.com/service/osteoarthritis-imaging/)
Figure 3. The Spondyloarthritis Research Consortium of Canada (SPARCC) MRI Sacroiliac Joint Inflammation Scoring System is based on the assessment of bone marrow edema in sacroiliac joint quadrants [70] (for further details go to www.carearthritis.com/service/mri-scoring-modules/)

**IMAGING OUTCOMES IN AXIAL SPONDYLOARTHRITIS**

Radiography

Radiography of the sacroiliac joints is still the cornerstone of diagnostic evaluation for evaluating features of spondylitis indicative of axSpA and the extent of change in each joint is assessed using the well-established modified New York criteria (mNY) grading system with a range of 0 (normal) to 4 (complete ankylosis). Spondylitis is defined as being definitively present (mNY+) where there is at least bilateral grade 2 or unilateral grade 3 severity on radiography. However, it is well-established that reliability for detection of sacroiliitis on radiography is poor, even among expert readers, and especially in early axSpA. The lack of reliable detection of early radiographic spondylitis has proven to be a major challenge in the use of radiography for quantification of structural progression.

An assessment of radiographs over 2 years where readers scored the images unaware of time point from 449 patients in the DESIR cohort, which recruited patients with inflammatory back pain of less than 3 years duration, reported progression from mNY criteria negative to positive after 2 years in only 4.9% of patients and in the reverse direction in 5.7% of patients [61]. In a 4.4-year follow up of patients in the Assessments in Spondyloarthritis international Society (ASAS) classification cohort, progression from mNY criteria negative to positive was reported in 18.3% and in the reverse direction in 58.1% [62]. It seems implausible that most patients would have reversion of structural changes without major anti-inflammatory therapy. Challenges in reader interpretation of the radiographs most likely accounts for this finding. Some studies have quantified radiographic spondylitis according to the mNY grade of 0-4 for each sacroiliac, which then provides a theoretical scoring range of -8 to +8 for both joints. However, this scoring methodology did not improve sensitivity to change as highlighted by the limited progression of only 5.1% over 5 years follow up in the DESIR cohort [63]. Moreover, since the grading is based on a composite of both increased new bone formation (sclerosis, ankylosis) as well as bone degradation (erosion), it is unclear what accounts for the change in radiographic spondylitis.

Despite these challenges with radiography, it has been possible to demonstrate significantly less radiographic progression in the sacroiliac joints after 2 years in patients receiving TNFi therapy compared to patients in an observational cohort not on TNFi where both groups were matched for prognostic variables [64]. However, 2 years is clearly much longer than the 12-16-week placebo-controlled portion of clinical trials and this method provides no real insights into the understanding of the basis for changes in appearance of the sacroiliac joints on x-ray. MRI now offers
a more precise and sensitive method that is capable of defining change in specific structural features in a 12-16-week time frame (vide infra).

Spinal radiographic assessment of abnormalities at anterior vertebral corners of the cervical and lumbar spine represents the current gold standard for quantitative assessment of radiographic progression in the spine. In the modified Stoke Ankylosing Spondylitis Spine Scoring method (mSASSS), lateral radiographs of the cervical and lumbar spine are evaluated [65]. The thoracic spine cannot be visualized on conventional radiography because of overlapping structures. A score of 1 is assigned to any one of vertebral corner erosion, squaring of the vertebral body, or vertebral corner sclerosis, a score of 2 is assigned to a vertebral corner syndesmophyte, and a score of 3 to each vertebral corner where there is intervertebral ankylosis. The total scoring range for the 12 vertebral corners in the cervical and lumbar spinal segments is therefore 0-72. These features appear slowly over time and at least 2 years follow up is required before change can be reliably detected. It is therefore not possible to detect disease modification in the 12-16-week time frame of placebo-controlled trials.

A recent active-controlled RCT compared an IL17-targeted monoclonal antibody, secukinumab, at monthly doses of 150mg or 300mg versus a TNF inhibitor biosimilar monoclonal antibody, adalimumab, to determine superiority of secukinumab as a structural disease-modifying therapy for ankylosing spondylitis using the mSASSS as the primary outcome [66]. A sample size of 846 patients was required to demonstrate with adequate power that the proportion of patients with no radiographic progression (change in mSASSS≤0.5) was greater on secukinumab versus adalimumab at week 104. Patients were selected for a higher risk of disease progression either because they had an elevated CRP and/or at least one syndesmophyte on spinal radiography. At 2 years, there was no significant difference between the proportion of patients with no radiographic progression between any of the treatment arms and the mean change from baseline in the mSASSS was 0.54, 0.55, and 0.72 in the secukinumab 150 mg, 300 mg, and adalimumab arms. This study illustrates the challenges of using a scoring instrument with such a limited sensitivity to change even in the setting of a head-to-head active comparator study. It is unlikely that such a study will ever be contemplated again in an industry-sponsored clinical trial.

Despite these limitations, a recent international based consensus of the ASAS group has endorsed the mSASSS as the preferred instrument for assessment of disease modification in trials of axSpA [67]. A key principle for the use of a method where the degree of change is so small is the importance of readers being adequately trained and calibrated to reduce measurement error and enhance reliability between readers in scoring assignment. A training module for the mSASSS has been developed by international consensus of experts in the field and this has been validated by demonstrating improvement in reliability of mSASSS scores after its use [68].

In view of the limitations posed by conventional radiography, increasing use is being made of low radiation CT (ldCT) of the whole spine at centers specializing in axSpA. Unlike conventional radiography, this technique permits visualization of bony structures in the thoracic spine, and syndesmophytes can be evaluated in both the sagittal and coronal planes. The amount of radiation is about 20-30% that of conventional CT. The Sensitive Imaging of Axial Spondyloarthritis (SIAS) Scoring Method is based on ldCT and scores new bone growth according to quadrants of each vertebral endplate [69]. Scoring of new bone growth is conducted as follows: 1 for a syndesmophyte that does not extend beyond 50% of the intervertebral disc height (IVDH), 2 for a syndesmophyte that extends beyond 50% of the IVDH, 3 for bridging ankylosis. This method was compared with the mSASSS for the assessment of structural progression in the spine over 2 years in 37 patients with typical axSpA. Significantly more patients were noted to have progression using the ld-CT versus mSASSS methods (84% versus 46%). The ld-CT SIAS method will require further evaluation in clinical trials to determine how much change can be reliably detected over shorter time frames because trials of 2 years duration are often associated with substantial patient drop-out and/or modifications in treatment which may influence the primary outcome of structural progression.

Magnetic Resonance Imaging

A major advance for the evaluation of therapeutics in clinical trials of axSpA has been the use of MRI to objectively quantify the degree of inflammation in the sacroiliac joints and spine. This requires the use of fluid sensitive MRI sequences, such as STIR or T2 fat-sat, that also suppress the signal from bone marrow fat which is found in healthy adults. The fluid that is detected may reflect edema and/or cellular infiltration. The most widely used methods to quantify inflammation in the sacroiliac joints and spine are the Spondyloarthritis Research Consortium of Canada (SPARCC) scores [70,71].
MRI assessment of the sacroiliac joints is conducted in the oblique coronal plane and a standardized image acquisition protocol has been developed by international consensus of musculoskeletal radiologists and rheumatologists with special expertise in imaging of axSpA (available at www.carearthritis.com/service/mri-spa-imaging-acquisition-protocols/). In the SPARCC sacroiliac joint method, bone marrow edema is quantified by dividing each joint into quadrants and then scoring the presence/absence of BME in subchondral bone marrow on 6 consecutive semicoronal slices (Figure 3). An added score is applied for BME that is intense and/or extends over at least one centimeter for each joint, leading to a scoring range of 0-72. Recent studies confirm its high degree of reliability and consistency in discriminating between active therapies and placebo at the primary endpoint of 12-16 weeks in recent clinical trials across the spectrum of axSpA and in the assessment of MRI inflammation of the sacroiliac joints in children [72-74]. SPARCC MRI sacroiliac joint scores demonstrate moderate correlation with clinical indices of disease activity, especially C-Reactive Protein (CRP), this being more evident in patients with early axSpA [75]. The degree of MRI inflammation also predicts clinical response to treatment as well as the development of structural damage in the sacroiliac joints. There is even evidence that MRI inflammation in the sacroiliac joints is a predictor of radiographic progression in the spine. It is also relatively easy for untrained readers to achieve high reliability for SPARCC scores with expert readers using an online calibration module (available at www.carearthritis.com/service/mri-scorings-modules/).

Inflammation in the spine can also be assessed objectively using fluid sensitive fat-suppressed MRI scans in the sagittal orientation. The SPARCC method quantifies BME in the vertebral bodies according to discovertebral units (DVU), each unit being delineated by two horizontal lines across the middle of adjacent vertebrae (Figure 4). Each of the 23 spinal DVU is divided into quadrants and the presence/absence of BME is recorded on 3 consecutive sagittal slices with an added score for intensity and extension of BME over at least one centimeter, leading to a scoring range of 0-18 for each DVU and 0-414 for the entire spine. It has also been used to score only the 6 DVU with the most severe inflammation. The SPARCC spine method has been shown to discriminate between active therapy and placebo at the primary endpoint of 12-16 weeks in several clinical trials across the spectrum of axSpA [74-77]. MRI inflammation in the spine has been associated with development of syndesmophytes and achieving remission of MRI inflammation in the spine has been associated with a reduced risk of developing structural progression in the spine. Consequently, there is the potential for MRI inflammation in the SIJ and spine to be used as a surrogate for structural progression in much the same way as bone density assessment has been used as a surrogate to evaluate risk of fracture. The SPARCC MRI inflammation methods have recently been endorsed by international consensus of the ASAS group as the preferred instruments to be used in the evaluation of disease-modifying therapies in clinical trials of axSpA [67].

Figure 4. The Spondyloarthritis Research Consortium of Canada (SPARCC) MRI Spine Inflammation Scoring System is based on the assessment of bone marrow edema in discovertebral unit quadrants [71] (for further details go to www.carearthritis.com/service/mri-scoring-modules/)
It is also important to quantify MRI inflammation in the posterolateral regions of the spine as these can be affected in axSpA and lead to impairment of spinal mobility. A consortium of investigators from Canada and Denmark have developed a scoring system that evaluates MRI inflammation on an anatomical basis in all regions of the spine. This CAN-DEN MRI inflammation score has recently been shown to discriminate between active therapy and placebo in clinical trials of axSpA [78].

MRI has also been used to evaluate structural lesions in the SIJ and spine. The SPARCC group has developed a method to quantify the primary structural lesions observed on MRI, as defined by the ASAS group, and these include erosion, fat metaplasia, backfill, and ankylosis [79]. Fat lesions and erosions are assessed in sacroiliac joint quadrants while backfill and ankylosis are assessed in sacroiliac joint halves of 5 consecutive MRI slices leading to a scoring range of 0-40 for erosion and fat metaplasia and 0-20 for backfill and ankylosis (Figure 5). Fat metaplasia and backfill are the MRI lesions observed when bone marrow edema and inflammation in an erosion cavity resolve after the institution of an effective anti-inflammatory therapy, such as agents targeting TNF or IL17. There is evidence that these lesions are an intermediary form of tissue repair in the pathway from inflammation to erosion to ankylosis. Moreover, this phenotype of tissue repair may predispose to the development of ankylosis both in the sacroiliac joints and spine. Consequently, it is important to quantify the impact of novel treatments on these lesions. Several recent placebo-controlled trials have demonstrated a significant reduction in MRI erosion in the sacroiliac joints of patients receiving therapeutics that inhibit TNF or IL17 and also the JAK-1 selective inhibitors, filgotinib and upadacitinib, as soon as 12-16 weeks at the primary endpoint of the trial [80]. A version of this method has also been developed for children (available at www.carearthritis.com/service/mri-scoring-modules). Training of readers for reliable detection of structural lesions is essential because of the diverse morphology of these lesions, especially erosions, and formal reading exercises should be preceded by the use of this online self-learning tool. (available at www.carearthritis.com/service/mri-scoring-modules). This SPARCC MRI Sacroiliac Joint Structural Score (SPARCC SSS) instrument has been endorsed by the international consensus of the ASAS group as one of the instruments that can be used to assess structural changes in axSpA [67]. Further studies should examine the predictive capacity of these MRI structural lesions in the SIJ and the development of new bone evident on radiography in patients with axSpA.

![Figure 5](https://example.com/figure5.png)

Figure 5. The Spondyloarthritis Research Consortium of Canada (SPARCC) MRI Sacroiliac Joint Structural Scoring System is based on the assessment of erosion and fat lesions in sacroiliac joint quadrants and backfill plus ankylosis in sacroiliac joint halves [79] (for further details go to www.carearthritis.com/service/mri-scoring-modules/)

MRI structural lesions in the spine can be quantified using the CAN-DEN method (available at www.carearthritis.com/service/mri-scoring-modules). This assesses fat lesions, erosion, and ankylosis according to
anatomical location and may allow the further scrutiny of the impact of potential disease-modifying therapies at different stages in the evolution of lesions from inflammation to fat metaplasia to new bone.

In conclusions: Recent publications demonstrate that we have reached a point where the continuing use of conventional radiography to objectively assess the efficacy of new therapeutic agents for inflammatory forms of arthritis is no longer helpful nor feasible. This is especially relevant to the assessment of disease-modifying therapies. The standard of care has advanced to the point where radiographic progression of disease in the placebo or active comparator arms is minimal in the time frame of ethically acceptable clinical trials. Fortunately, there is now ample evidence that MRI offers a more sensitive and responsive imaging modality that should now supplant radiography for the assessment of disease modification. Further advances aimed at automation and absolute quantification of key disease parameters will further enhance its utility in future clinical trials.

References


45. Østergaard M, McQueen F, Wiell C, et al. The OMERACT psoriatic arthritis magnetic resonance imaging scoring system (PsAMRIS): definitions of key pathologies, suggested MRI sequences, and preliminary scoring system for PsA hands. J Rheumatol 2009; 36: 1816-1824.


64. Dougados M, Maksymowycz WP, Landewe R, et al. Evaluation of the change in structural radiographic sacroiliac joint damage after 2 years of etanercept therapy (EMBARK trial) in comparison to a contemporary control cohort (DESIR cohort) in recent onset axial spondyloarthritis. Ann Rheum Dis 2018; 77: 221-227
78. Maksymowycz WP, Østergaard M, Landewé R, et al. Filgotinib decreases both vertebral body and posterolateral spine inflammation in ankylosing spondylitis: results from the TORTUGA trial. Rheumatology 2022; 61: 2388-2397