

Osteoarthritis and Cartilage



Review

OARSI Clinical Trials Recommendations: Hand imaging in clinical trials in osteoarthritis



D.J. Hunter †*, N. Arden ‡§, F. Cicuttini ||, M.D. Crema ¶#, B. Dardzinski ††, J. Duryea ††, A. Guermazi ¶, I.K. Haugen §§, M. Kloppenburg ||||, E. Maheu ¶¶, C.G. Miller ##, J. Martel-Pelletier †††, R.E. Ochoa-Albíztegui †††, J.-P. Pelletier †††, C. Peterfy §§§, F. Roemer ¶||||, G.E. Gold ¶¶¶

† Institute of Bone and Joint Research, Kolling Institute, University of Sydney, and Rheumatology Department, Royal North Shore Hospital, Sydney, NSW, Australia

‡ NIHR Musculoskeletal Biomedical Research Unit, University of Oxford, UK

§ Arthritis Research UK Centre for Sport, Exercise and Osteoarthritis, University of Oxford, UK

|| School of Public Health and Preventive Medicine, Monash University, Alfred Hospital, Melbourne 3004, Australia

¶ Quantitative Imaging Center, Department of Radiology, Boston University School of Medicine, Boston, MA, United States

Department of Radiology, Hospital do Coração (HCor) and Teleimagem, São Paulo, SP, Brazil

†† Department of Radiology & Radiological Sciences, Uniformed Services University of Health Sciences, Bethesda, MD, USA

††† Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, USA

§§ Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway

|||| Departments of Rheumatology and Clinical Epidemiology, Leiden University Medical Centre, 2300 RC, Leiden, Netherlands

¶¶ Department of Rheumatology, Hôpital Saint-Antoine, AP-HP, Paris, France

Bioclinica, Newtown, PA, USA

††† Osteoarthritis Research Unit, University of Montreal Hospital Research Centre (CRCHUM), Montreal, Quebec, Canada

‡‡‡ Department of Radiology, The American British Cowdray Medical Center IAP, Mexico

§§§ Spire Sciences, Inc., Boca Raton, FL, USA

||||| Department of Radiology, University of Erlangen-Nuremberg, Erlangen, Germany

¶¶¶ Stanford University, Departments of Radiology, Bioengineering, and Orthopedic Surgery, USA

ARTICLE INFO

Article history:

Received 29 November 2014

Accepted 9 March 2015

Keywords:

Imaging
Hand osteoarthritis
Guidelines

SUMMARY

Tremendous advances have occurred in our understanding of the pathogenesis of hand osteoarthritis (OA) and these are beginning to be applied to trials targeted at modification of the disease course. The purpose of this expert opinion, consensus driven exercise is to provide detail on how one might use and apply hand imaging assessments in disease modifying clinical trials. It includes information on acquisition methods/techniques (including guidance on positioning for radiography, sequence/protocol recommendations/hardware for MRI); commonly encountered problems (including positioning, hardware and coil failures, sequences artifacts); quality assurance/control procedures; measurement methods; measurement performance (reliability, responsiveness, validity); recommendations for trials; and research recommendations.

© 2015 Osteoarthritis Research Society International. Published by Elsevier Ltd. All rights reserved.

Introduction

Substantial advances have occurred in our understanding of the pathogenesis of hand osteoarthritis (OA) and these are beginning to be applied to trials targeting modification of the disease course¹.

* Address correspondence and reprint requests to: D.J. Hunter, Rheumatology Department, Royal North Shore Hospital and Institute of Bone and Joint Research, Kolling Institute, University of Sydney, Sydney, NSW Australia. Tel.: 61-2-9463-1887; Fax: 61-2-9463-1077.

E-mail address: David.Hunter@sydney.edu.au (D.J. Hunter).

Current regulatory requirements for disease modification require that alongside the assessment of structural effects that symptom improvement be demonstrated². The previous guidelines for OA clinical trials published in 1996³ included recommendations for imaging with a predominant focus on radiography (consistent with the era); some details are provided in the appendices on methods of acquiring radiographs and use of MRI. A more recent iteration focused on hand OA recommended conventional radiographs to image hand OA both for selection of patients and severity assessment and also for monitoring progression in structure modifying

trials⁴. In neither of these recommendation documents was there detail on the different imaging methods available, the pitfalls inherent nor performance metrics of different imaging markers. Of note, an updated guideline on the conduct of clinical trials in persons with hand OA is included in this issue for OARSI clinical trial guidance however it does not contain comprehensive detail on imaging (Maheu *et al.*).

The purpose of this expert opinion, consensus driven exercise is to provide detail to anyone performing or planning clinical trials, imaging scientists and their respective teams on how one might use and apply this knowledge in disease modifying clinical trials utilizing hand imaging assessments. It includes information on acquisition methods/techniques (including guidance on positioning for radiography, sequence/protocol recommendations/hardware for MRI); commonly encountered problems (including positioning, hardware and coil failures, sequences artifacts); quality assurance/control procedures; measurement methods; measurement performance (reliability, responsiveness, validity); recommendations for trials; and research recommendations.

Method

This review began with a search of Pubmed using terms of OA, hand, and imaging. We also conducted general searches for manuscripts covering general randomized control trial (RCT) methods, covering each of the sub-topics below; from this literature we identified designs and methods, as well as other manuscripts of high relevance. It should be noted that the vast majority of this manuscript is based upon expert opinion of the diverse multidisciplinary group involved in this exercise. Authors of this review included multiple experts familiar with different approaches for imaging of the hand joint, including radiologists, rheumatologists, and engineers. This expert opinion was generated via a series of teleconference and email exchanges followed by generation of a series of recommendations. This correspondence allowed the working group to identify additional topics and manuscripts for inclusion and to develop and reach concurrence on a set of recommended principles for inclusion of hand imaging methods in OA implementation trials. Given the potential for divergent perspectives for the trial recommendations and research recommendations a survey was conducted of these members to determine the strength of recommendation for each point raised. Final recommendations (Tables I and II) were obtained by averaging the responses to a survey among the 14 authors for the strength of recommendation from 0 to 100. At the commencement of this exercise all members of the working group were asked about conflicts of interest and the results from those who were conflicted were not included in the survey results. The focus of the content is on radiography and MRI as the preferred imaging techniques with some content on ultrasound when appropriate.

Acquisition methods and techniques

Radiography

Typical hand radiographs can be acquired with multiple views^{5,6}. The hand is usually evaluated with postero-anterior (PA) and oblique views, occasionally with more detailed magnified views of the joint of interest⁷. For symptom-modifying trials, where a radiograph is obtained for diagnostic purposes, a single PA radiograph of both hands imaged side-by-side on the same cassette is acceptable. For structure-modifying trials, however, a PA radiograph of each hand should be obtained on a separate cassette, with the palm of the hand on the film, the fingers extended and adducted, and the entire forearm resting on the X-ray table. Beam-

centering and angulation (i.e., centered on the third metacarpophalangeal joint (MCP-3), perpendicular to the film-screen or receptor), focus-to-film distance and exposure should be standardized across all participating imaging facilities.

Measures should be taken to ensure that the right and left hands are accurately labeled. This is important because if a right hand at baseline is compared to a left hand at follow-up that was erroneously labeled right, differences in the degree and distribution of abnormalities can be misinterpreted as disease progression or regression. Image labeling can be unreliable if radiopaque markers are placed on the film-screen manually. Anatomical asymmetries, when noticed, can help identify labeling errors, but these are often subtle, and many cases go unrecognized. Most clinical trials of rheumatoid arthritis (RA) have overcome this error by utilizing positioning frames with permanently embedded left-right markers designed in such a way as to make it impossible to image a hand with the wrong marker. As with knee radiography, changes in the position of the hand can change JSW values^{8,9} and are likely to affect measurements of osteophytes. The use of a hand positioner is therefore recommended so that consistency is maintained both cross-sectionally and longitudinally.

There are no published reports that quantify the effect of radiographic technique (mA and kVp) on measurement error; however these factors are unlikely to alter measurements of JSW or osteophytes unless the quality is significantly compromised. Huetink *et al.*⁹ reported no systematic change in JSW as a function of the film-to-focus distance (FFD) for FFD covering 110–120 cm. Indeed as long as the hand is placed directly on top of the detector and the FFD is consistent at all time points for all patients, the magnification factor will change negligibly as a function of FFD. Current positioners usually come with a calibration marker built in.

While no studies have specifically examined the effect on detector pixel spacing on measurement accuracy or precision, studies have shown comparable reproducibility of Genant-modified Sharp scoring using screen-film radiographs digitized at a pixel spacing of 0.1 mm with that using the original film-screen radiographs¹⁰, and several studies have quantified reproducibility and average JSW change from which it may be possible to draw conclusions about requirements for spatial resolution. In a study of RA subjects Neumann *et al.*¹¹ reported a long term repositioning reducibility of 0.10 mm for screen-film radiographs digitized at a pixel spacing of 0.1 mm. Finckh *et al.*¹² reported an average JSW change of 0.16 mm in RA subjects with a 4 year median follow-up time. In a phantom study Huetink *et al.*⁹ reported a smallest detectable distance of 0.028 mm for JSW measured using a “standard digital X-ray imaging system”. Angwin *et al.*⁸ reported a value of 0.11 mm to represent a threshold for “an actual physical change in joint space width”. Together these studies suggest that clinical trials which use radiographic JSW as an endpoint should use a detector with the smallest pixel spacing possible and that the same imaging protocol should be used for all time points.

MRI

Key factors in determining which MRI protocol to use to evaluate OA of the hand include the measurement method to be used to analyze the images, patient tolerance with respect to examination time, specific needs or preferences of the readers, the degree of heterogeneity of technology and experience among imaging centers participating in the study, whether the study design is cross-sectional or longitudinal, and of course, any regulatory requirements and budgetary constraints.

To support scoring with OMERACT Hand Osteoarthritis Magnetic Resonance Scoring System (HOAMRIS), MRI images must cover at least the distal interphalangeal (DIP) and proximal

interphalangeal (PIP) joints with as high spatial resolution as possible, given the small size of these joints^{13,14}. To optimize scoring the sequences must be obtained in at least two planes or with a three-dimensional technique with small isometric voxels in one plane, 1 mm slice thickness and subsequent reconstruction¹⁵. Accordingly, 1.5T or 3T scanners are recommended. Coverage of the PIP and DIP joints can be accomplished with most commercial surface coils, which are widely available and therefore applicable to multicenter clinical trials. If additional coverage is needed, for example to include the first carpometacarpal (CMC-1) (see also Research Recommendations), two separate scans with a surface coil or a single large field of view scan with a knee coil will be necessary. Multicenter imaging of the hand and wrist with larger coils has been successfully performed in multicenter clinical trials of RA^{16–18}. Using a knee coil necessitates imaging the patient prone with the arm extended over the head, which may be problematic for patients with shoulder problems. If multicenter availability of imaging technology is not a constraint, as in single-center studies, specialized MRI systems, coils and set ups can be used. The focus in this review, however, is the more constrained setting of multicenter studies in which images acquired from diverse facilities must ultimately be pooled for analysis. A key objective in this context is minimizing technical differences among images from the different centers, particularly spatial resolution and image contrast.

If longitudinal analyses are to be done, reproducible positioning of the finger joints is critical, as even slight variations in their orientations relative to each other will result in variable sectioning of each bone and joint from scan to scan, regardless of how meticulously the slices are prescribed. Reproducible positioning can be facilitated by the use of specially designed acrylic frames that fix the orientation of finger joints, as has been the standard in clinical trials of RA for years¹⁹.

Because of the anatomy of the finger joints, sagittal images are most useful for evaluating articular cartilage. However, as in the knee, coronal images may be adequate as well, and require less than half the imaging time to achieve full coverage of the hand. Axial images are not very useful for evaluating cartilage in the fingers, but are excellent for monitoring synovitis and tendons, as well as associated ligamentous changes. Axial images, however, like sagittal images, require longer scan times to cover the necessary anatomy in the hand. Sagittal images are also important in the DIP

and PIP joints for evaluating dorsal–palmar osteophytes, which are often prominent in hand OA. Bone erosions, subarticular cysts and bone marrow lesions are all well visualized on coronal and sagittal images (Figs. 1 and 2). Thus, MRI protocols including all three orthogonal planes are ideal. However, because of imaging time constraints, one may be practically limited to only one or two planes (we would recommend sagittal and coronal). For example, almost all randomized controlled clinical trials of RA using MRI that have been reported to date have been able to include only a single plane of section (coronal). Despite this, most of these studies were highly discriminative of progression of joint damage over short follow-up intervals and with relatively small numbers of subjects per arm. While it is difficult to broadly generalize about resolution given tradeoffs with imaging time, joint/s studied and SNR we would make the following suggestions for sequence criteria for high resolution and clinical studies. For clinical studies would suggest minimal slice thickness: 1 mm for the 3D (SPGR, FLASH); 1.5 mm for proton density- and T1-weighted, 3 mm for the STIR; Gap: no gap is preferable; and Field of view (FOV): 140 mm for the hand, and 180 mm if wrist is included in coronal plane (120 mm in axial plane). For high resolution studies would suggest: Minimal slice thickness: 1 mm; Gap: no gap is preferable; and FOV: 40 or 45 mm.

With respect to pulse sequences, T1-weighted 3D gradient echo images with as much in-plane and through-plane resolution as time will allow, should be used to evaluate cartilage, cysts, erosions and osteophytes. T1-weighted 3D gradient echo images are suited for quantitative assessment of cartilage, but are probably not ideal for semi-quantitative assessment since it shows suboptimal conspicuity in the detection of focal cartilage defects, as demonstrated in other joints²⁰. Fat suppression²¹ or selective water excitation are helpful for visualizing the articular cartilage and subchondral bone marrow alterations^{22,23}. This can be challenging in the fingers, as irregular shapes perturb magnetic fields, which can in turn result in heterogeneous or completely failed spectral fat saturation or water excitation. Careful calibration of the MRI magnet for each examination and strategic placement of water bags can improve this problem in many cases. Short Tau Inversion Recovery (STIR) images offer more robust fat suppression, but lack the spatial resolution required to evaluate small changes in these features in finger joints. STIR images are highly sensitive for evaluating

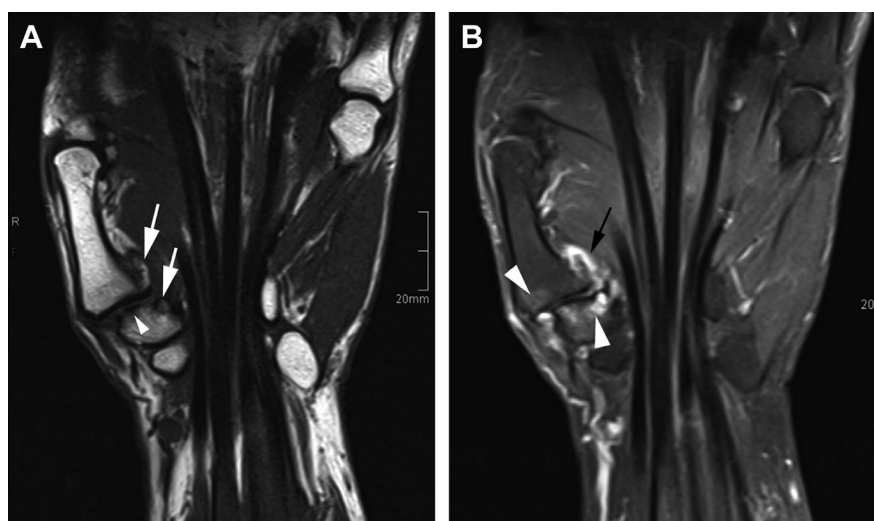


Fig. 1. Visualization of thumb base OA. A. T1-weighted spin echo MRI shows ulnar-sided osteophytes at the base of the first metacarpal and the articulating os trapezium (arrows). In addition there is subchondral sclerosis ideally visualized on T1-weighted images (arrowhead). B. Corresponding T1-weighted fat-suppressed contrast-enhanced MRI shows concomitant synovitis surrounding the metacarpal osteophyte (black arrow) and subchondral bone marrow lesions (white arrowheads).



Fig. 2. Severe OA of the DIP joints. Coronal T1-weighted MRI shows large osteophytes at the ulnar-sided heads of the second and third middle phalanges. In addition there is a subchondral erosion at the head of fourth middle phalanx visualized as a rounded hypointense lesion on T1-weighted imaging (arrowhead).

bone marrow lesions. STIR, especially in the axial plane, can be used to indirectly evaluate synovitis and tenosynovitis. Intermediate-weighted fast (or turbo) spin-echo images with fat saturation acquired with relatively thin slices (2–3 mm) may also offer an excellent contrast assessing synovitis, tenosynovitis and bone marrow. However, fat-suppressed, T1-weighted sequences following intravenous gadolinium-based contrast agent should be ideally performed for the detection and quantification of synovitis and tenosynovitis, and also helps distinguish bone marrow lesions from erosions and cysts because of superior resolution (Fig. 1)¹³.

If contrast-enhanced MRI is used, care should be taken to exclude patients with renal insufficiency to minimize the risk of nephrogenic systemic fibrosis. The current recommendation for clinical practice is to exclude patients with glomerular filtration rate (GFR) < 30 mL/min/1.73 m². For clinical research, higher thresholds, closer to 60 mL/min/1.73 m², are often used²⁴.

Recently, compositional MRI techniques have also been introduced to assess biochemical changes within the cartilage of fingers these are difficult and time consuming. First reports have focused on the delayed gadolinium enhanced MRI (dGEMRIC) technique but other techniques may be of use as well^{25,26}. These techniques may be useful for evaluating early changes in articular cartilage before significant morphological destruction has occurred. At present, however, compositional MRI techniques have not been used in clinical trials due to challenges applying them consistently in multicenter settings and because of limited data on validation.

In the end, any MRI protocol must be optimized around the scientific objectives and unique practical constraints of the specific study in question, particularly with respect to the study centers, which usually have different equipment, internal processes and local cultures to consider (Fig. 2). It is difficult, therefore, to define in great detail a basic imaging protocol that can be generalized to all studies. In the opinions of the authors, however, a longitudinal multicenter study of OA in the fingers would ideally use ≥ 1.5 T MRI with a positioning frame and a surface coil or a knee coil if additional joints will be included. Pulse sequences should include sagittal and coronal, T1-weighted, 3D gradient-echo scans with fat-suppression or water excitation and at least $350 \mu \times 700 \mu$ in-plane and 1500μ through-plane resolution, coronal STIR ($470 \mu \times 470 \mu$ in-plane and 3000μ through-plane resolution) and if possible axial T1-weighted, 3D gradient-echo with fat-suppression or water excitation following intravenous contrast administration ($1170 \mu \times 470 \mu$ in-plane and 2000μ through-plane resolution). If contrast is not included, axial STIR ($1170 \mu \times 470 \mu$ in-plane and 3000μ through-plane resolution) should be used instead (Fig. 3).

Ultrasound

Ultrasound can be used in OA to visualize osteophytes, synovitis, effusions and erosions. Positioning and interpretation of ultrasound imaging of the hand is an interactive process with the ultrasonographer using visual feedback to scan the areas of interest. Many joints can be individually imaged providing increased coverage. In addition to grey-scale ultrasound imaging, power Doppler (Fig. 4) can be used to assess neovascularity in inflammatory arthropathies and erosive/inflammatory hand OA. A multi-frequency linear transducer, typically 8 MHz up to 17 MHz is used to evaluate hand joints.

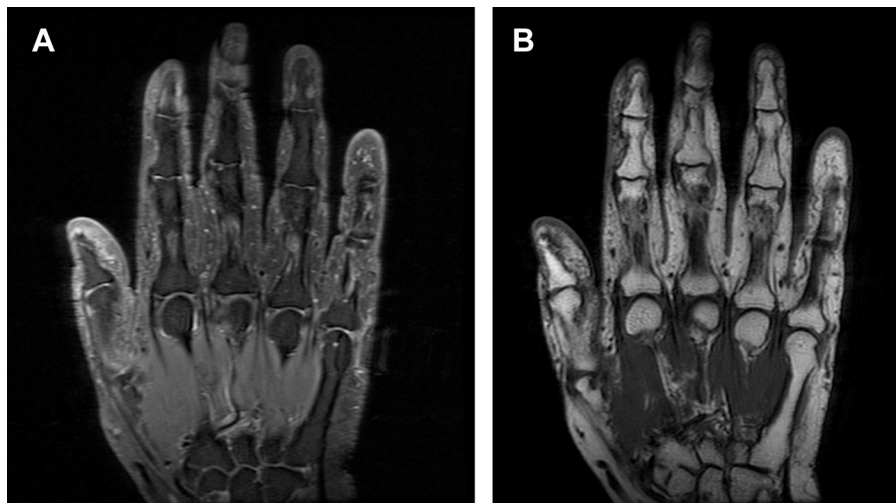


Fig. 3. Representative clinical MRI in a person with hand OA. A. Coronal fat-suppressed proton density-weighted MRI reveals a small subchondral cyst in the third PIP joint. B. Coronal T1-weighted MRI depicts small osteophytes, subchondral sclerosis and cartilage loss in multiple joints including thumb IP, second DIP and third PIP joints.

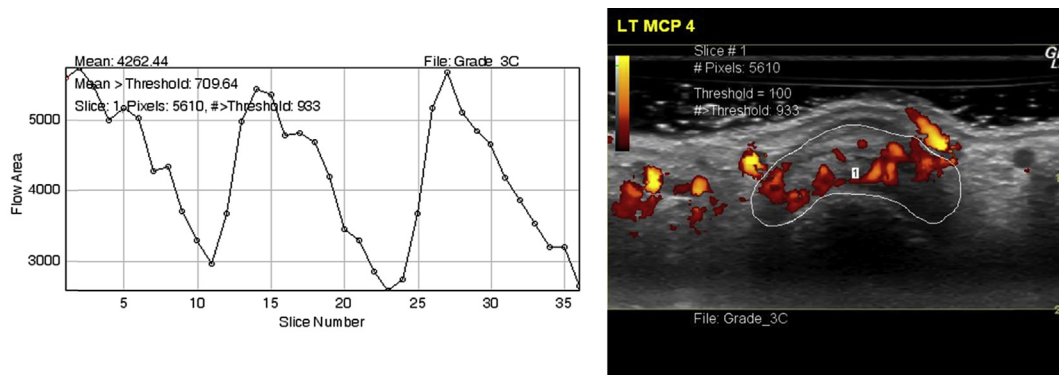


Fig. 4. Quantifying power-Doppler ultrasound over 36 frames of acquisition of the metacarpophalangeal joint.

Commonly encountered problems

Radiography

Challenges in radiography have to do primarily with the projectional viewing perspective of the technique. Since even slight variations in hand positioning or of beam centering or angulation among serially acquired images can obscure or reveal osteophytes, erosions and other structural abnormalities or simulate joint-space narrowing, meticulous care must be taken to ensure reproducible projection. Additionally, the shift from film-based to digitally acquired radiography has complicated multicenter, longitudinal analyses. These challenges, along with problems related to right-left mislabeling, are discussed in detail in “Recommendations for Clinical Trials.”

MRI

Positioning within the MR system is crucial to achieve adequate chemical fat saturation. Therefore the preferred position is that of the “Superman” position where the patient lays prone head first into the MR system bore with their hand palm side down above their head. This allows the hand/wrist to be positioned at the isocenter of the magnet providing superior fat saturation when needed. However this can be an uncomfortable position. If the study protocol is not too extensive and the total examination time can be kept short, e.g., 30–45 min, this usually is not a problem. However, if the examination is prolonged or the patient is in too much pain, the patient can be imaged in the more comfortable “off-center” position, in which the patient enters the MR system supine and head first with the arm at the side. The problem with this position is that since the hand is in the periphery of the magnet bore, where the magnetic field is less homogeneous, signal-to-noise ratio (SNR) is typically poorer and spectral fat saturation or selective water excitation needed to identify edema might not be optimal. Other techniques for fat suppression, such as inversion recovery (STIR), tend to be less affected by field heterogeneities, and therefore may be a better choice under such conditions. Hence, thought must be given to the trade-offs among comfort, robustness and efficiency when designing an imaging protocol for clinical research.

Typically circumferential MR coils for the wrist are quite small, providing coverage of only the wrist, but achieving high SNR. Surface coils are more versatile, but still commonly restricted to field of views allowing anatomical coverage of only the wrist or the fingers of the hand but not both. Coil technology has advanced over the years with the introduction of phased array configurations that provide extremely high SNR, and allow knee coils or other large coils to image both the hand and wrist simultaneously. Most MR

coils are of receive-only type, which increases the Specific Absorption Rate (SAR) since the body coil must be used to transmit. There are specific transmit/receive coils for the knee and wrist which not only reduce SAR but allow faster imaging times and increased SNR employing smaller echo times, which can be important for acquisitions like T2 mapping.

The most commonly encountered problem besides positioning and anatomical coverage is poor fat saturation. This is because the irregular shape of the hand perturbs the magnetic field irregularly, and since resonant frequency is dependent on field strength, spectral fat suppression techniques based on pre-irradiating the field of view with the average frequency of fat (chemical fat saturation), will fail in certain locations, usually at the fingers or other protuberances, where the frequency of fat has shifted, leading to heterogeneous fat saturation. The same is true of selective water excitation. Incomplete fat saturation artifacts can be minimized by carefully shimming the magnet with the hand in the bore, reducing the anatomical irregularities by strategically placing water bags or other volume supplements over or between the fingers, or by using more robust fat-suppressing pulse sequences, such as inversion recovery.

Ultrasound

In ultrasound each joint has to be evaluated separately and with different ultrasound modalities such as gray scale, power Doppler and real time change in position; this increases volume of data which comes at the expense of time. Therefore operator experience is crucial as well as trying to establish reproducibility when imaging is done at multiple sites using equipments of various manufacturers. Images can be quantified based on intensity and number of voxels that are present. This can be performed on the instrument and with offline specialized analysis tools. One of the major issues is the centralization of reading and analysis of images. Since acquisition is operator dependent the image capture protocol (joints and position analyzed) need to be standardized, each imaging facility should indicate in its policy and standard procedure measure. Standard procedure images should be recorded, and standard plus additional images should be stated in the report when recorded such as findings with additional imaging characterization (power Doppler, harmonics, elastography and etc). for any central reading for large clinical trials.

Quality assurance/Control procedures

Radiography

A hand map or specially designed positioning frames that ensure the correct right-left labeling of the hand should be provided to each

trial site to aid reproducibility of positioning and to ensure consistency of hand positioning among centers. In order to assure a standardized technique, data on the initial radiograph should include type of radiographic film, type of radiographic cassette, kilovolts, milliseconds, and milli-ampere. These should be identical in subsequent studies. For all trials, the radiographs should be masked for the patient name and the radiograph should be identified by a code. Unless there is a special reason, the first carpometacarpal (CMC-1) joint does not need to be imaged separately⁴.

For both symptom and structure-modifying trials, the pre-study films should be evaluated, prior to entry into the study, to confirm the diagnosis and define eligibility⁴. For structure-modifying trials, all films from a single patient should be evaluated at the same time after the final films have been obtained. Although blinding to sequence is the convention right now, its necessity has been questioned. For most studies, the radiographs of each patient are presented to the readers in random order, with the date of acquisition masked, so as to blind the readers to visit sequence. This is particularly important in longitudinal trials without placebo- or inactive-comparator control. However, some argue that blinding to sequence is not necessary in studies that include inactive control as long as the readers are blinded to treatment allocation, and that reading with known chronology under such circumstances may even improve sensitivity to change²⁷. However, in studies that show progression in all treatment arms but no discrimination among them, unless the readers were blinded to sequence, one cannot be sure that the progression was not based entirely on bias and therefore that there was no potential for discriminating treatment effect in the first place, i.e., type-2 error. Thus, the decision whether or not to blind to sequence must be made with care. Further, if the protocol allows the readers to be aware of the time sequence, this potential source of bias should be identified in the protocol, and an appropriate explanation and justification provided. For intra-reader reproducibility, a pre-specified number of pairs of baseline and endpoint radiographs should be re-read at each session and between sessions. Re-reading anything less than all of cases usually does not offer sufficient power to discriminate a meaningful discrepancy, so the value is usually not informative. Since the re-read sample size usually needs to be all or close to all of the cases, to be meaningful, one might as well use two readers. Re-reading of radiographs allows calculation of the intra-reader cross-sectional and longitudinal variability. If there is more than one reader, inter-reader variability should also be calculated using the same sample pairs of radiographs. In Phase III pivotal studies in other therapeutic areas using nominal or semi-quantitative scoring methods, the regulatory agencies (FDA and EMA) usually have 100% of all radiographs read by two readers and a third reader adjudicate any differences between the first two. This is therefore recommended for clinical trials of OA as well.

Before starting the readings, the readers should train together with the same radiographs to identify and hopefully resolve major discrepancies and inconsistencies in the way each reads. Studies have shown that there are variations in the threshold for defining pathological findings even among expert readers^{28–30} and that variability is minimized when each subject's radiographs are assessed at the same time³¹. In multicenter studies, initial recruitment radiographs are sometimes read by site readers to determine subject eligibility. It is critical in such cases that the site readers also be trained and validated prior to beginning. Such training can be performed by expert readers from the core laboratory where the efficacy readings will be conducted. A preferable alternative is to have the films centrally read to avoid false positive patients from diluting the cohort. Once the patients are recruited, all radiographs should be read by centralized reader(s) in order to contain measurement variability.

MRI

The image quality goals for MRI of the hand in OA are 1) complete anatomical coverage (both in-plane and through-plane), 2) absence of artifacts obscuring regions of interest, 3) homogeneous signal with adequate contrast-to-noise ratio and spatial resolution, 4) identical contrast and resolution among serially acquired images, and most importantly, 5) identical tomography among serially acquired images. The latter is particularly important for accurately identifying small changes between visits. Since the hand contains 24 independently mobile bones, no matter how meticulously the tomographic sections are aligned at each visit, each bone will be sectioned slightly differently each time unless its position relative to the other bones is fixed in a reproducible way. The hand must therefore be immobilized with a positioning frame¹⁹ and the space around the frame and hand filled with rubber sponge to ensure extension of the fingers and to reduce motion¹⁴. Further, MRI sections should be carefully aligned in both orthogonal planes using reproducibly identifiable anatomical landmarks. Hardware and pulse sequences should not be changed between visits.

Careful central monitoring of image quality is highly recommended. As noted above, reproducibility of between-visit alignment of tomographic anatomy and therefore slice orientation is particularly important for accurate, sensitive assessment of change.

As was the case for radiography, MRI readers must be trained on the scoring method to be used and work together to calibrate their readings prior to initiating study readings, so as to minimize variability. All images should be centrally read by two independent readers blinded to treatment allocation and usually to visit order. Some percentage of the largest discrepancies between the two readers' scores for all or some features can be adjudicated by a third independent reader who was previously calibrated to the original two readers. If such a third reader is not available, as is often the case with novel reading methods, adjudication can be performed by consensus review of the discrepant cases by the original two readers. The latter approach has been used in the majority of randomized controlled MRI trials of RA reported thus far.

Ultrasound

Ultrasound is operator dependent, and it is likely the same operator is not able to acquire all the ultrasounds in the trial. Thus standardizing the sequence and order of images acquired will help if different operators are involved. Using the same equipment and transducer should also help to enhance the homogeneity of images.

Measurement methods

Radiography

Semi-quantitative scoring

Several scoring systems for assessment of radiographic hand OA have been developed. The scales differ in assessed joints and whether they provide a global estimation of OA or assess individual features. Currently, there is no consensus on the preferred scale.

The Kellgren and Lawrence (KL) scale published in 1957³² was the first proposed scoring system for hand OA, and is still the most widely used³³. The scale provides a global score of OA severity (0–4 scale, of which grade ≥ 2 represents definite radiographic OA) based on the presence/severity of osteophytes/ossicles, joint space narrowing (JSN), sclerosis of the subchondral bone, pseudocystic areas, and altered shape of bone ends³² for each joint (MCPs, PIPs, DIPs, thumb base, i.e., trapeziometacarpal and scaphotrapezoid). In 1963 the authors published an atlas with radiographic images and legends describing the features in each particular film³⁴. The

interpretation of the scoring system has been complicated by different written definitions of the grades across joint groups as well as different descriptions in various publications. Further, the KL scale has been criticized for too much emphasis on osteophytes³⁵, as narrowed/sclerotic joints cannot be classified as having OA unless osteophytes are present. Thus, several studies have used modified KL scales to overcome these deficiencies.

Other scoring systems with detailed assessment of individual hand OA features have been developed in order to address the deficiencies of the KL scale and optimize agreement^{36–38}. Among these, the Osteoarthritis Research Society International (OARSI) atlas published in 1995 (revised in 2007) is most commonly used^{37,38}. With this atlas as a reference, the presence/severity of individual features such as osteophytes, JSN, subchondral erosions (pseudowidening), cysts, subchondral sclerosis, and malalignment are assessed on semi-quantitative scales. A grading system has been proposed by Kallman³⁶ assessing six features: osteophytes, JSN according to a 0–3 scale and sclerosis, cysts, lateral deviation and erosion in presence/absence.

Other scoring systems have been proposed, but have hardly been used in any studies after publication. Lane *et al.* proposed a new global scale for presence and severity of hand OA³⁹. A summary grade (0–2 scale) was derived directly from the assessment of osteophytes and JSN according to a modified Kallman scale³⁶. Similarly, Kessler *et al.* proposed a global scale for presence (not severity) of hand OA based on osteophytes, JSN and sclerosis according to the OARSI atlas^{37,40}. Both the Lane and Kessler scales gave more emphasis on JSN in comparison to the KL scale, and did not include other radiographic features such as cysts/erosions and deformity in the definition of hand OA.

Verbruggen *et al.* developed two numerical scoring systems for the progression of erosive and non-erosive hand OA^{41,42}. The first scoring system, the anatomical lesion progression system, was based on changes in osteophytes, JSN and cysts (± 0.5 point for increase/decrease in size and ± 1 point for appearance/disappearance). The system was limited by no evaluation of the magnitude of change. Further, they noticed that joints that developed erosions often had increasing joint space width (JSW), and therefore got negative scoring values despite disease progression. Hence, they developed the anatomical phase progression system to comprise the erosive evolution. This scoring system was based on an assumption of hand OA as a disease that undergoes predictable phases; in the non-erosive stationary phase (S phase), the joint has classical hand OA features such as osteophytes, JSN, and subchondral bone changes. When the joint progresses into the destructive phases, the joint space completely disappears (J phase), and thereafter the subchondral plate becomes eroded (E phase). These destructive phases are then followed by repair or remodeling (R phase). Recently, Verbruggen *et al.* also proposed a more complex scoring system; the Ghent University Scoring System (GUSS), which has shown higher sensitivity in detection of progression during the destructive phases⁴³. In GUSS, the proportions of normal subchondral bone, subchondral plate and joint space are assessed on an 11 point rating scale (range 0–100 with 10 unit increases).

Quantitative measurement

Image processing software methods can be used to provide fully quantitative assessment of structures on radiological images. These measurements rely less on reader subjectivity than semi-quantitative scoring systems. Much of the earlier work in the field addressed measuring the loss of JSW to assess RA progression^{12,44–47}. Most of the methods include software that puts a grid (mesh) over the hand x-rays and calculates the distance of joint space. More recently these methods have been evaluated on OA subjects to measure reliability⁹, correlation with semi-

quantitative scoring⁴⁸, and in a case–control study, a comparison to clinical features⁴⁹. With these techniques the software automatically delineates the opposing margins of the joint and calculates the JSW as the distance between the bone margins. Since both OA and RA cause loss of joint space, software methods developed for RA can be applied to the OA hand, although the commonly involved joints are different for the two types of arthritis, and the presence of osteophytes in the OA joint may require special consideration. The only problem with these computer-based methods of measuring JSW is that whilst they offer a quantitative, reliable, and more objective means by which to assess JSN in patients with RA, they tend to be very time consuming, which has limited their clinical use¹².

MRI

In a study with IV contrast MRI has higher sensitivity to detect osteophytes and erosions than conventional radiography, and also synovitis⁵⁰. On MRI, osteophytes are detected as abnormal bone protuberance at joint margins or surfaces on T1-weighted fat-suppressed sequence^{13,14}. Erosion can be detected as a sharply marginated bone lesion with typical signal characteristics (i.e., increased signal intensity in the area of cortical bone and bone marrow fat on T1-weighted fat-suppressed images)^{13,14}. Synovitis can be detected as an area in the synovial compartment that shows post-contrast enhancement of a thickness greater than the width of normal synovium^{13,14}.

Haugen *et al.* proposed a comprehensive MRI scoring system and atlas for hand OA (Oslo hand OA scoring system), including how to assess osteophytes, JSN, erosions/attrition, cysts, malalignment, synovitis, flexor tenosynovitis, bone marrow lesions and collateral ligament abnormalities¹⁴. Most features were scored on 0–3 scales in the proximal and distal part of the joint separately, if applicable. Cysts, malalignment and collateral ligament abnormalities were scored as absent/present¹⁴. The scoring system covers the distal and PIP joints, and future studies should examine whether the scoring system can be similarly applied to the metacarpophalangeal and thumb base joints. Haugen *et al.* found good intra- and inter-reader reliability for the majority of hand OA features in the scoring system¹⁴. Using the same scoring system, good reliability was also confirmed in another cohort⁵¹, suggesting that MRI can reliably assess OA pathology in the small finger joints.

The scoring system was comprehensive in order to include all potentially relevant features in hand OA. However, the authors have discovered that certain features, such as collateral ligament pathology and flexor tenosynovitis, were infrequently present, did not correlate with OA severity and/or were not associated with pain in the Oslo hand OA cohort^{15,50}. Hence, based on this experience, the scoring system was further optimized by the MRI interest group in the Outcome Measures in Rheumatology (OMERACT)¹³. OMERACT HOAMRIS¹³ is a semi-quantitative scoring system for hand OA features that was iteratively developed based on Oslo hand OA MRI scoring system published in 2011¹⁴. The changes made by OMERACT included exclusion of collateral ligament pathology and flexor tenosynovitis, scoring of the joint as a whole (i.e., not the distal and proximal part separately) and adding half scores (0.5 increments) for bone marrow lesions, synovitis and erosions in order to better capture small changes in longitudinal studies.

Ultrasound

In order to facilitate a reliable scoring of ultrasound features in the finger joints, a preliminary scoring system was developed.

Three features were included in the scoring system; grey-scale synovial hypertrophy/effusion, power Doppler activity and osteophytes. All features were scored as absent/present as well as on semi-quantitative scales (grade 0–3)⁵². A large “real-life” reliability exercise was arranged in order to test the reliability of the proposed scoring system. Despite divergent results between ultrasonographers and for the various features (in general lowest reliability for power Doppler activity and highest for osteophytes), the authors concluded that the results were satisfactory and that the scoring system could be a good basis for further development of ultrasound as an outcome tool.

To facilitate the scoring of osteophytes, Mathiessen *et al.* developed an atlas showing examples images of different grades of osteophyte severity in the interphalangeal joints (grade 0–3)⁵³. Using the atlas for scoring of stored ultrasound images, the authors found excellent inter-reader reliability. There is currently no published atlas for assessment of grey-scale synovitis and power Doppler activity in hand OA. However, the synovitis found in OA can be scored according to the published comprehensive ultrasound atlas for grey-scale synovitis and power Doppler activity in RA⁵⁴.

The first scoring system by Keen *et al.* did not include assessment of erosions and JSN due to concerns about reliable definitions, contemporary ultrasound technology and feasibility related to the duration of the scanning⁵². Recently, definitions of cartilage pathology in the metacarpophalangeal joints were proposed: 1. Global cartilage abnormalities (absent/present); 2. Loss of anechoic structure and/or cartilage thinning; 3. Irregularities and/or loss of sharpness of >1 cartilage margin⁵⁵. Results from a “real-life” reliability exercise showed varying inter-reader reliability from fair (irregularities and/or loss of sharpness of) to very good (global assessment). A further limitation of the system is the focus on metacarpophalangeal joints only, which are less frequently affected by OA.

Measurement performance: reliability, responsiveness, validity

Radiography

As mentioned earlier, the most commonly used semi-quantitative radiographic scoring methods for hand OA are KL grading³², Verbruggen–Veys (anatomical phases)⁴² and Kallman³⁶ grading systems, and the OARSI atlas^{37,38}. They vary regarding radiographic features and joints assessed. Two recent comparison studies^{29,30} of patients followed over 1 and 6 years, respectively found no real performance difference among these methods. In fact, KL, which is a global rather than feature-based score, showed the greatest change, although this scoring system was not designed for assessing change. A further systematic review evaluated discrimination (reliability, sensitivity to change), feasibility and validity of the available radiographic scoring methods and found comparability across studies and called for a consensus on preferred scoring method, the examined joints and the used presentation of data (Visser *et al.* O&C in press). In addition, none of these scoring methods, except the Verbruggen anatomical phase scoring system has been developed and validated to be analyzed as a global scoring summing scores of all hand joints into one global score (patient level). Whether this is appropriate rather than considering and analyzing each score at the joint level remains an open issue.

Quantitative radiographic methods have demonstrated good reliability⁹, substantive correlation with semi-quantitative scoring⁴⁸, and in a case–control study, a good correlation to clinical features⁴⁹.

MRI

OMERACT HOAMRIS¹³ is a semi-quantitative scoring system for hand OA features that was iteratively developed based on the Oslo hand OA MRI scoring system published in 2011¹⁴. Haugen *et al.* found good intra- and inter-reader reliability for the majority of hand OA features in the scoring system¹⁴. Using the same scoring system, good reliability was also confirmed in another cohort⁵¹, suggesting that MRI can reliably assess OA pathology in the small finger joints. These studies have also shown that MRI can reliably assess most OA features, including structural changes as well as inflammation, with the current technology^{14,51}. Although good reliability has been shown in cross-sectional studies^{14,51}, longitudinal studies using MRI in hand OA are needed to examine the reliability of serial assessments as well as the sensitivity to change of the MRI features.

Ultrasound

Recent studies have provided insight on the validity with an association with clinical features, such as pain⁵⁶ or radiographic and MRI feature^{57–60} and with radiographic progression over time⁶¹. Ultrasound has been shown to be more accurate to detect osteophytes and erosions in hand OA than conventional radiographs^{53,58,59}.

Recommendations for clinical trials (Table I)

General comments

The goals of imaging in clinical trials can include subject selection, monitoring disease progression and treatment effect, and/or identifying complications of the disease or the treatment. An imaging biomarker may be excellent for diagnosis but not useful for monitoring disease progression. To be applicable to the latter, the biomarker must change over time. For example, the presence of a particular receptor in certain joint tissues may be critical to a patient's responsiveness to a treatment specifically targeting that receptor. Perhaps one could use a radiolabeled probe for that receptor using scintigraphy or PET to identify patients applicable to the therapy. However, if the receptor itself doesn't change in response to the treatment, the imaging technique cannot be used to monitor response. In the case of hand OA, the presence of osteophytes in the DIP/PIP joint or CMC-1 joint at the base of the thumb, are important for diagnosis, but it is not clear how changes in osteophyte size relate to clinical outcomes, besides cosmetically. It is possible that osteophyte formation actually represents an adaptive, potentially beneficial response to OA, and that removal or suppression of osteophytes under certain circumstances could be deleterious. Alternatively, a biomarker that normally can be used to monitor disease progression or treatment response, such as cartilage loss or JSN, can become useless if it has reached its ceiling, and cannot progress further. Thus, severity of such an otherwise responsive marker could serve as a criterion for excluding patients from certain studies. The same could be said for subchondral bone attrition or erosion. Once morphological distortion is too great and becomes a significant driver of further damage itself, some treatments may lose efficacy.

With respect to monitoring disease progression or treatment response, not only does the linkage between changes in the biomarker and how patients feel and function matter, but also the power with which the biomarker discriminates change. This relates to the actual rate of change of the structure or process that the biomarker targets, but also to the effect size and the measurement error. The latter includes variability introduced during image

Table 1
Summary of recommendations for clinical trials

Process	Strength of recommendation (range 0–100)*
Site technologists should receive training on study-specific imaging technique, the quality of their images should be centrally verified and regular feedback provided	95
Patient-grouped images from multiple visits should be viewed simultaneously, blinded to treatment allocations and usually to chronological order as well	72
Readings should be performed centrally by two independent readers, and adjudicated	79
Readings should be performed by a musculoskeletal radiologist	49
Radiography	
Each hand should be imaged separately	56
Standardized, reproducible positioning, beam centering and exposure are critical	83
Acquisition should include the use of a positioning frame	69
Regarding scoring methods, consistent with the consensus from OMERACT 12 we encourage use of the most widely used and currently best validated measures in a core set for structural damage. We would advocate use of either the Kellgren–Lawrence method, or the OARSI atlas or the Verbruggen–Veys method or the Kallman method, as preliminary instruments for the structural damage domain.	83
MRI	
Coverage should include the fingers, especially DIP and PIP joints, and possibly CMC-1	91
Reproducible tomography is critical including the use of an acrylic positioning frame and biplanar slice alignment	72
Triplanar sections are ideal, but if not feasible, coronal and possibly sagittal sections should be acquired	89
Pulse sequences should include STIR and high-resolution, fat-suppressed 3D gradient echo \pm post-contrast T1-weighted sequence	83
Most developed reading method today is HOAMRIS and we would recommend its use in MRI trials	84

* Mean response from the 14 persons on the working group who responded to the survey. Results of participants who were conflicted are not included. Strength of recommendation scale ranged from 0 (don't recommend) to 100 (strongly recommend).

analysis but also that introduced during image acquisition. Variability stemming from the images themselves is minimized by judicious protocol design and meticulous quality control, including technologist training, ongoing image quality verification, and repeat imaging whenever needed. Variability associated with image analysis is contained by using central readers (ideally two) who are specially trained, calibrated to each other and highly experienced in reading for clinical trials, and by using reading workstations that optimally display the serial scans, minimize data entry errors and limit reader fatigue.

Safety monitoring with imaging depends on the mechanism of action of the intervention, but may include looking for signs of osteonecrosis, infection or stress fracture.

Explicit decisions about any of these factors depends on the scientific, regulatory and business objectives of the study as well as practical considerations, such as patient tolerance and safety, patient availability, time to market goals and budgetary constraints. Central to all of these is the role that the project will play in the treatment development program. Design considerations are very different if the study is a pivotal phase-III trial intended to establish definitive proof of efficacy and safety for regulatory approval, than if the intention is to inform internal decision making and portfolio management, such as in phase-II proof of concept or studies aimed

at determining how subsequent phase-III trials should be designed with respect to patient type, biomarker type, sample size, follow-up interval, intensity of quality control, etc. Phase-IV studies that explore special circumstances, such as patient subtypes or different practice settings similarly entail different protocol design tradeoffs.

Image acquisition and quality control

How the images for a study will be analyzed, including the particular biomarkers that will be used and the strengths and preferences of the readers who will perform the readings, ultimately dictates how the images should be acquired and the specific quality criteria that must be prioritized. No matter how experienced, skilled and dedicated the readers for a clinical trial may be, their readings can only be as accurate and sensitive to change as the quality of the images is able to support. Thus, image quality can make or break a trial. Unlike in clinical practice, images for clinical trials are acquired from multiple different sites, often from different countries, each with different hardware, software, local processes, languages and cultures, and the images from these diverse facilities must be pooled with no discernable differences among them except those related to the patients and the treatments. There is no other context in radiology in which this is necessary, and it is thus counterintuitive to clinical radiologists and technologists, who have been taught to optimize image acquisition around individual patient idiosyncrasies, local hardware/software strengths and weaknesses, and their local radiologists' preferences. Each cohort of imaging sites included in a study poses unique challenges to designing an imaging protocol that accomplishes uniform image quality, and thus it is difficult to generalize. Designing such protocols require in depth multivendor technical knowledge as well as an understanding of how radiology departments function in different parts of the world.

Once the imaging protocol is decided on, site technologists must be instructed on the procedure, particularly why the imaging protocol was designed the way that it was, what elements of the protocol are most important to accurate image analysis and therefore require the most attention, and how to avoid and solve problems that may arise. How this information is conveyed depends on the experience and motivation of the technologists involved, but typically involves a combination of printed manuals and didactic or interactive training sessions conducted in groups or one-on-one by webinar or face-to-face meetings. Test scans with volunteers or even phantoms can be helpful, but continuous feedback on quality and acquiring repeat scans when necessary, if feasible, are critical elements of a successful training/quality control program.

As outlined in "Acquisition Methods/Techniques-Guidance on positions for radiography, sequence/protocol recommendations/hardware for MRI," radiography of each hand should be performed independently with the hand positioned palm down and flat and the fingers fully extended. Measures should be taken to ensure that the right and left hands are accurately labeled. Image labeling can be unreliable if radiopaque markers are placed on the film-screen manually. Key radiographic image quality considerations are complete anatomical coverage, reproducible projection of the joint spaces and bones among serially acquired images, absence of artifacts obscuring target locations, and sharp delineation of cortical and trabecular bone. Once screen-film based radiographs are felt to be of adequate quality by the radiology technologist, they should be digitized to 100 μ pixel spacing, and graphic quality verified centrally using uniform criteria.

Recently, there has been a trend at most clinical facilities away from screen-film radiography to digital radiographic acquisition. While this has been a positive change for image archival, patient

throughput and environmental impact, it has added complexity and variability to multicenter clinical trials. Not only is the spatial resolution of digitally acquired radiographs generally lower than that of images generated by digitizing originally film-based radiographs, verifying the quality of digitally acquired radiographic images is harder for technologists. In the past, technologists could visually inspect film radiographs directly on backlit view boxes. However, digital radiography generates 12-bit electronic images comprising 4096 grayscale units. Since the human eye can discriminate only a small fraction of these, technologists must infer image quality indirectly from the exposure index, which is a parameter that is computed differently by different manufacturers. Not only has this added technical variability to multicenter clinical trials, but, because some imperfections in digital image quality can be corrected through post-processing by the reading radiologist after the patient has left, the focus of radiology technologists at busy medical centers has shifted from meticulous image acquisition to rapid patient throughput. Unfortunately, reproducible patient positioning on serial radiographs, which is critical for reliably assessing change in clinical trials, cannot be corrected by digital post-processing. Accordingly, studies employing digital radiography must invest additional effort to ensure uniformity and high quality of images pooled from multiple sites.

As outlined in “Acquisition Methods/Techniques-Guidance on positions for radiography, sequence/protocol recommendations/hardware for MRI,” MRI of the hand should employ an acrylic frame that fixes the position of the fingers and hand properly and reproducibly. It should also include meticulous biplanar alignment of slices using reliable anatomical landmarks to ensure reproducible tomography of the bones and joints on serially acquired scans in longitudinal studies. Anatomical coverage should include at least the fingers, particularly the DIP and PIP joints. This can be accomplished with most commercial surface coils. If the CMC-1 joint is of interest, which currently is not formally included in HOAMRIS, two separate scans with surface coils or a single scan with a knee coil will be needed. Ideally, high-resolution, fat-suppressed, T1-weighted, 3D gradient-echo scans in all three orthogonal planes should be acquired. However, if this is not feasible because of time constraints, as is usually the case, coronal and possibly sagittal scans should be acquired to assess osteophytes, cartilage loss and subarticular cysts/bone erosions. To evaluate bone marrow lesions and synovitis, coronal STIR or fat-suppressed intermediate-weighted spin echo scans should be included. These generally lower-resolution pulse sequences could be substituted with contrast-enhanced fat-suppressed, T1-weighted, 3D gradient-echo scans, which are more reliable for synovitis assessment. However, gradient-echo techniques are less sensitive for bone marrow lesions even with enhancement. Thus, the decision between STIR and contrast-enhanced 3D gradient-echo depends on the relative importance of synovitis and bone marrow lesions to the study objectives. Tendons and synovitis are best assessed in the axial plane. As noted above, contrast-enhanced 3D scans are optimal for synovitis but due to time constraints, STIR is often the most feasible option for axial scanning.

As detailed in the section on Quality Assurance/Control Procedures, key image quality criteria for MRI include adequate anatomical coverage, absence of artifacts or signal drop-off in target locations, adequate tissue contrast and spatial resolution, and importantly, comparability of cross-sectional anatomy among serial scans.

Image analysis

Since the experience with clinical trials of hand OA is so limited, there is no consensus yet on which scoring method should be used.

As discussed in “Research Recommendations for Hand osteoarthritis,” the most commonly used radiographic scoring methods for hand OA are KL grading³², Verbruggen anatomical phases scoring⁴², Kallman scale³⁶ and the OARSI atlas^{37,38}. The most recently developed MRI scoring method is the Outcome Measures in Rheumatology (OMERACT) Hand OA MRI Score (HOAMRIS)¹³.

Regardless of the scoring method used, however, readings for clinical trials should be performed centrally, ideally by two trained readers experienced in the scoring methods to be used and in reading for clinical trials of arthritis. As is the standard in RA trials, serially acquired images for a patient should be displayed simultaneously and possibly in random visit order, so that the readers remain blinded to chronology. This is particularly important in studies without an inactive comparator arm. Note that re-reading of some timepoints may be required when interim analyses are performed in extended, multi-visit longitudinal studies.

Each reader should independently score all images for all patients on regulatory compliant electronic case report forms (eCRF) without knowledge of the other readers' values. Additionally, in most studies, cases in which the readers disagree more than a certain degree are rescored either by a third independent reader or by consensus review by the original two readers. The purpose of this adjudication is not to statistically smooth variability further, as that would require re-reading of all of the cases in order to be unbiased, but rather to identify data entry errors or other mistakes that would be expected to show up as discrepancies but not related simply to ambiguity of scoring criteria, as the latter is already dealt with by score averaging. Regulators are provided both the original and adjudicated results.

Research recommendations

General comments

Clinical research in imaging is often divided into technical innovation or application development. The two are interdependent in that development of new capabilities, enhanced performance of existing capabilities, or improved convenience, accessibility or cost-effectiveness can reveal previously unanticipated application possibilities, and new or unmet existing needs drive demand for technological innovation.

In the case of hand OA, one can anticipate a similar course as was seen in RA, in which imaging has been used successfully in clinical trials for more than two decades. As was the case for RA 20 years ago, the current lack of effective structure modifying therapy for hand OA has limited demand from clinicians for detailed imaging information about joint structure in these patients. Why determine whether a patient is an appropriate candidate for a structure-modifying therapy or whether that therapy is working properly if the therapy does not yet exist? However, precise and fully validated methods for doing both of these are necessary to gain regulatory approval of any putative structure-modifying therapy for hand OA. Thus, the demand for imaging in this disease, as in most others, emerges first during the clinical development of potential therapies by pharmaceutical, biotechnology and medical devices companies, and it is therefore the scientific, logistical, regulatory and business needs of the drug development process that shape the early evolution of imaging for hand OA.

Also as for RA, once effective therapies become accessible, withholding treatment for prolonged periods in placebo controlled trials becomes unethical. This increases demand for methods that can discriminate treatment effects more rapidly and sensitively, both for clinical trials and clinical practice. Further, as the therapies become available in the clinic, the measurement methods must be adapted for that context, which differs in a number of ways from that of clinical trials. In clinical research, there is a greater demand

Table II
Summary of research recommendations (the strength of recommendation refers to the priority this group gives to this research question/topic)

General comments	Strength of recommendation (range 0–100)*
Further work on the specific joint (DIP vs PIP vs CMC-1 vs MCP) to be assessed	79
Conventional radiography	
Which structural features (e.g., osteophytes, JSN erosion etc.) should be assessed?	61
Reliability of longitudinal JSW assessment (including the reproducibility of radiographic projection)	73
Sensitivity to change (JSW)	78
Timing of imaging assessment(s) (depends on biological rate of change, which needs to be assessed in epidemiological studies, sensitivity to change of the imaging modality and anticipated effect size of the intervention)	80
Ultrasound	
Which structural features and joints (e.g., the CMC-1 may be more challenging to examine than the DIP joints) could reliably be assessed with the current technology	63
Methods to optimize variability introduced by operator dependency	87
Reliability of longitudinal assessment	82
Timing of imaging assessment(s) (depends on biological rate of change, which needs to be assessed in epidemiological studies, sensitivity to change of the imaging modality and anticipated effect size of the intervention)	70
Sensitivity to change of the ultrasound features	78
MRI	
Which structural features could reliably be assessed with the current technology	85
Protocols for assessment of the interphalangeal and CMC-1 joint (of which the latter is a complex joint and could be target for non-systemic treatment)	76
Reliability of longitudinal assessment	90
Sensitivity to change of the MRI features	95
Timing of imaging assessment(s) (depends on biological rate of change, which needs to be assessed in epidemiological studies, sensitivity to change of the imaging modality and anticipated effect size of the intervention)	84
Use of contrast for assessment of synovitis	69
Additional research questions suggested by survey respondents	
Whether paired radiographs of a same patient should be read in known chronological order or blinded to time sequence.	
Is there sufficient measurement performance data validating MRI and ultrasound as imaging methods in hand OA that they can replace or supplant the need for radiography.	
How best to analyze the scorings and radiographic progression; whether at the patient level or joint level.	
The utility of elastography (both on ultrasound and MRI) as a tool to evaluate musculoskeletal tissues.	

* Mean response from the 14 persons on the working group who responded to the survey. Scale ranged from 0 (don't recommend) to 100 (strongly recommend).

for quantitative data rather than subjective data. Images in clinical practice must answer questions about the management of individual patients, where as in clinical trials images acquired from multiple different sites must be pooled to answer research questions applicable to large groups of patients. Imaging for hand OA is currently in the first stage of this development cycle. Accordingly, it is important to accumulate as much long-term data as possible on placebo-treated patients while structure modifying therapies are still not available.

As for RA, radiographic and MRI scoring methods have been developed for diagnosing hand OA, quantifying the severity of

structural damage associated with it and for monitoring change and potential improvement over time, including radiographic KL^{32,34} and OARSI^{37,38} scales, Verbruggen⁴² and Kallman³⁶ grading and the HOAMRIS system¹³. Image-quality requirements for supporting these analyses and thus the protocols for acquiring the images are in turn driven by the needs and preferences of the radiologists who ultimately will perform the readings.

As in RA, radiographic methods for assessing hand OA developed first. As newer technologies, such as MRI, ultrasound, etc., became available, new methods for quantifying disease severity and change capitalizing on the unique strengths of these innovations followed.

Technical optimization is thus needed around each of the following factors:

- Which disease features should be assessed with each candidate imaging modality
- Which semi-quantitative scales or quantitative measures should be applied to each feature, anticipating that the performance requirements for these biomarkers will likely change once they have verified the efficacy of new structure modifying treatments, and these therapies become clinically available
- Strengths and weaknesses of different image acquisition protocols in multi-site clinical trial settings

Each of these will be addressed below.

Radiography

Radiography's strengths in imaging hand OA are high contrast and two-point resolution for cortical and trabecular bone, widespread availability and technical familiarity, and relatively low unit cost. Its weaknesses are low contrast for discriminating non-calcified tissues, projectional viewing perspective, and ionizing radiation, which for hand imaging is relatively minor. Thus, radiography can visualize osteophytes, bone erosions and cysts, joint-space width, bone sclerosis and joint alignment in the hand, but it cannot see articular cartilage directly, bone marrow lesions or osteitis, synovitis, joint effusion, tenosynovitis, tendons or ligaments, all of which are relevant to pain and functional disability in hand OA, and which can be assessed with MRI (see below). Future studies should explore to which extent these additional MRI-visible, disease-based features link to pain and dysfunction in hand OA, and whether they add substantially more discriminative power to assessments of disease progression and treatment response than do the conventional radiographs.

Because it is a projectional rather than a tomographic technique, and multiple planes are superimposed in one image, radiography magnifies and distorts anatomy. This makes quantitative dimensional measurement more challenging, and necessitates meticulously reproducible hand positioning and projection on serially acquired radiographs in order to assess change accurately and sensitively. Projection can also obscure findings by superimposing overlying structures. Thus, osteophytes and erosions projected enface are less sensitively detected than those tangential to the X-ray beam. Even minimal rotation of the hands or variation in beam centering and alignment between serial examinations can therefore mimic or obscure changes in these features and joint-space narrowing. These limitations decrease sensitivity to change, necessitating larger numbers of patients and longer follow-up intervals to achieve statistical power. Moreover, as for RA, the recent shift from film-based radiography to digitally acquired radiography at most imaging facilities around the world has made high-quality multi-center radiography of the hand more challenging.

Finally, for each of the radiographic methods, the optimal timing of assessments needs to be determined. This depends on the pathological rate of change, which in turn needs to be determined in epidemiological studies. It also depends on sensitivity to change of the imaging modality, which in turn depends on the effect size of the intervention and the measurement error of the scoring method. Research into each of these would be helpful.

MRI

Very few studies have so far been performed on MRI in hand OA, and future studies are needed. A few studies have shown that MRI can reliably assess most OA features, including structural changes as well as inflammation, with the current technology^{14,51}. However, these findings should be confirmed in more studies. There are currently no protocols for how best to assess the finger joints, including the interphalangeal, metacarpophalangeal and thumb base. Furthermore, previous studies have focused on interphalangeal joints^{13,14}, and no scoring system is available for the metacarpophalangeal joints and thumb base.

Although good reliability has been shown in cross-sectional studies^{14,51}, longitudinal studies using MRI in hand OA are needed to examine the reliability of serial assessments as well as the sensitivity to change of the MRI features.

When it comes to the timing of MRI assessment, it depends on the biological rate of change, which needs to be assessed in epidemiological studies, sensitivity to change of the imaging modality and anticipated effect size of the intervention.

Assessment of synovitis according to the proposed MRI scoring systems for hand OA requires the use of intravenous gadolinium contrast^{13,14}. Future studies are needed in order to compare the sensitivity in detection of synovitis using non-contrast enhanced images and contrast-enhanced images.

Ultrasound

As with MRI, ultrasound has the capacity to detect inflammation, deformity (osteophytes) and damage (erosions, ligament abnormalities). It cannot assess subcortical bone abnormalities because of acoustic shadowing by cortical bone. For the same reason, some joint surfaces and synovial cavities in the hand and wrist are not accessible to US, particularly the ulnar surface of metacarpophalangeal joint (MCP) 2, the radial surface of MCP 5, both the radial and ulnar surfaces of MCPs 3 and 4, and the ulnar surface of the important CMC-1 at the base of the thumb.

Most OA studies using ultrasound have deployed semi-quantitative assessments and further work is still required on their basic psychometric properties. Criterion and construct validity may be assumed for synovitis (based on RA studies but also on osteoarthritic knee biopsy studies), and there has been comparison between ultrasound and radiographic osteophytes⁶². Further work is required on understanding the validity of JSW or cartilage loss when compared to other modalities. There is a need for more information on how ultrasound abnormalities relate to clinical symptoms⁶³ and also for predictive validity⁶⁴ on most ultrasound-detected OA pathologies.

Reliability has been examined for a number of scoring systems^{52,55}, and further work on definitions and standardized reader training will be required to improve this area, which underpins responsiveness. Few studies have examined responsiveness of various scoring systems⁶⁵; generally these focus on pathological features expected to show change.

Use of ultrasound in clinical trials is feasible, though lessons can be learnt from the more numerous RA trials. As is clear from the previous discussion, the optimal semi-quantitative score or

quantitative tool for (even selected) pathologies is not yet established. More work is required to standardize operating procedures for such studies. As well as understanding and recording inter-machine differences in pathology detection, studies with multiple readers or sites require consensus training pre-trial and well defined definitions of pathologies, which may reduce sensitivity while increasing inter-reader agreement⁶⁶. Issues on timing of studies will depend on the responsiveness of individual imaging biomarkers.

Summary and conclusion

The goals of imaging the hand in clinical trials can include subject selection, monitoring disease progression and treatment effect, and/or identifying complications of the disease or the treatment. For acquisition a PA radiograph of each hand on separate cassettes should be obtained, with the palmar aspect of the hand placed on the film with the fingers extended and adducted on a positioning frame. MRI protocols must be optimized around the scientific objectives and unique practical constraints of the specific study in question, particularly with respect to the study centers. For a longitudinal multicenter study of OA in the fingers ideally one would use ≥ 1.5 T MRI with a positioning frame and a surface coil. For radiographic studies patient-grouped images from multiple visits should be viewed simultaneously, blinded to treatment allocations. Readings should be performed centrally by two independent readers, and adjudicated. This manuscript includes a number of recommendations for clinical trials that we would advise anyone planning on using imaging in hand OA follow.

Author contributions

All authors were involved in collecting data, reviewing the literature and drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published.

Disclosure

The comments and editorial expressed herein represent those of the author/s and do not reflect those of any official scientific role or institution that the author/s may be hold or be affiliated with. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Conflict of interest

- 1 David Hunter has royalties from Donjoy for a patent granted for a patellofemoral buttress. Has a grant from the NIH, multiple grants from Australian Research Council and NHMRC. Associate Editor for Osteoarthritis and Cartilage.
- 2 Nigel Arden – Has a grant from the NIHR for outcomes of arthroplasty and Biomedical Research Unit, one from NIH for hip morphology and another one from ARUK for video, project and equipment. Received honoraria from Bioiberica, Scherring-Plough, Merck, Servier.
- 3 Flavia Cicuttini-Has grants from the NHMRC for Statin RCT, Zoledronic Acid RCT and ACL study and TasOAK cohort. Holds a volunteered paid position in Repatriation Medical Authority as a Medical Officer, and one as an Associate Editor in Arthritis and Rheum, BMC Musculoskeletal Journal, and ART Journal.
- 4 Michel D. Crema – has shares in Boston Imaging Core Lab, LLC a company providing image assessment services to academia and the pharmaceutical and medical device industry.

- 5 Bernard Dardzinski – no conflict of interest to disclose.
- 6 Jeffrey Duryea – has a grant from NIH/NIAMS for a Quantitative MRI analysis method for longitudinal assessment of knee OA.
- 7 Ali Guermazi receives honorary from Genzyme, TissueGene, OrthoTrophix and Merck Serono. He has shares in Boston Imaging Core Lab, LLC. Deputy Editor of Radiology.
- 8 Ida Kristin Haugen – Has a grant from the Norwegian Rheumatism Association (Extrastiftelsen) for imaging in hand OA.
- 9 Margreet Kloppenburg-received lecture/consultancy fees from Pfizer, Servier, Abbvie, UCB, BMS and has grants from Dutch Arthritis Foundation, Pfizer, TI Pharma, OMERACT, and is Associate Editor at Arthritis & Rheumatology.
- 10 Emmanuel Maheu-Has received in the past 5 years honoraria and/or been invited at congresses from/by Expanscience, Ibsa-Genevriev, Pierre Fabre Labs, Rottapharm, TRB-Chemedia. Editorial Board of Osteoarthritis and Cartilage.
- 11 Colin G. Miller- Senior vice president at Bioclinica.
- 12 Johanne Martel-Pelletier – receives honoraria from ArthroLab/ArthroVision, Bioiberica, Elanco, Ferring, Merck, Pfizer, Servier, TRB Chemedia. Has grants from The Arthritis Society (as co-investigators), Canadian Institutes of Health Research (CIHR) (as co-investigator). Has investment in non-medical industry at ArthroLab/ArthroVision and is a Shareholder. Holds a volunteered lecturer position at the ESCEO13-IOF European Congress on Osteoporosis and Osteoarthritis (travel grant).
- 13 Jean-Pierre Pelletier-receives honoraria from ArthroLab/ArthroVision, Bioiberica, Elanco, Ferring, Merck, Pfizer, Servier, TRB Chemedia. Has grants from The Arthritis Society (as co-investigators), Canadian Institutes of Health Research (CIHR) (as co-investigator). Has investment in non-medical industry at ArthroLab/ArthroVision and is a Shareholder. Holds a volunteered lecturer position at the ESCEO13-IOF European Congress on Osteoporosis and Osteoarthritis (travel grant).
- 14 R. Elena Ochoa Albíztegui – no conflict of interest to disclose.
- 15 Charles Peterfy – has shares in Spire Sciences, inc., which provides Image analysis for clinical trials to multiple pharmaceutical and medical device companies. He holds a volunteered paid position in The International Society for Musculoskeletal Imaging in Rheumatology as treasurer.
- 16 Frank Roemer – has shares in Boston Imaging Core Lab, LLC, a company providing image assessment services to academia and the pharmaceutical and medical device industry. Associate Editor for Osteoarthritis and Cartilage.
- 17 Garry E. Gold receives honoraria as a consultant from Zimmer Inc., Isto, Inc. and Boston Scientific. Has a grant from GE Healthcare for improved MR Systems and one from the NIH for advanced MRI of Osteoarthritis.

Acknowledgments

Dr Hunter is supported by an NHMRC (Grant no. 1079777) Practitioner Fellowship. Dr Gold is supported by an NIH K24 Grant (AR062068) and RO1 (EB002524).

We acknowledge the support and contribution of the entire OARSI Imaging Clinical Trials Recommendations working group which includes: Ali Guermazi, Bernard Dardzinski, Charles Peterfy, John Carino, Colin Miller, Diann Stern, Emmanuel Maheu, Felix Eckstein, Flavia Cicuttini, Frank Roemer, Garry Gold, Hollis Potter, Ida Haugen, Jean-Pierre Pelletier, Jean-Pierre Raynaud, Jeffrey Duryea, Joanne Jordan, Johanne Martel-Pelletier, Kim Bennell, Lee Simon, Marc Hochberg, Margreet Kloppenburg, Michael Nevitt,

Michel Crema, Mike Bowes, Nancy Lane, Philip Conaghan, Richard Kijowski, Roy Altman, Saara Totterman, Siegfried Trattng, Souhil Zaim, Thomas Link, Thomas Schnitzer, Tim McAlindon, Tim Mosher, Yves Henrotin, Frank Roemer, Goetz Welsch, Graeme Jones, Jordan Renner, Marie-Pierre Hellio Le Graverand and Nigel Arden. We also acknowledge the administrative support of Diann Stern.

OARSI gratefully acknowledges support to defer in part the cost of printing of the Clinical Trial Recommendations from Abbvie, BioClinica, Boston Imaging Core Lab, and Flexion. The funding sources for printing had no role in the outcome of this manuscript.

References

1. Matthews GL, Hunter DJ. Emerging drugs for osteoarthritis. *Expert Opin Emerg drugs* 2011;16(3):479–91.
2. Abramson SB, Berenbaum F, Hochberg MC, Moskowitz RW. Introduction to OARSI FDA initiative OAC special edition. *Osteoarthritis Cartilage* 2011 May;19(5):475–7.
3. Altman R, Brandt K, Hochberg M, Moskowitz R, Bellamy N, Bloch DA, *et al.* Design and conduct of clinical trials in patients with osteoarthritis: recommendations from a task force of the Osteoarthritis Research Society. Results from a workshop. *Osteoarthr Cartil* 1996 Dec;4(4):217–43.
4. Maheu E, Altman RD, Bloch DA, Doherty M, Hochberg M, Mannoni A, *et al.* Design and conduct of clinical trials in patients with osteoarthritis of the hand: recommendations from a task force of the Osteoarthritis Research Society International. *Osteoarthr Cartil* 2006 Apr;14(4):303–22.
5. Schreiber KL, Freeland A, Gilula LA, Yin Y. Imaging of the hand and wrist. *Orthop Clin N Am [Review]* 1997 Oct;28(4):537–82.
6. Bhat AK, Kumar B, Acharya A. Radiographic imaging of the wrist. *Indian J Plast Surg* 2011 May;44(2):186–96.
7. Swagerty Jr DL, Hellinger D. Radiographic assessment of osteoarthritis. *Am Fam Phys [Review]* 2001 Jul 15;64(2):279–86.
8. Angwin J, Heald G, Lloyd A, Howland K, Davy M, James MF. Reliability and sensitivity of joint space measurements in hand radiographs using computerized image analysis. *J Rheumatol* 2001 Aug;28(8):1825–36.
9. Huetink K, van 't Klooster R, Kaptein BL, Watt I, Kloppenburg M, Nelissen RG, *et al.* Automatic radiographic quantification of hand osteoarthritis; accuracy and sensitivity to change in joint space width in a phantom and cadaver study. *Skelet Radiol* 2012 Jan;41(1):41–9 [Research Support, Non-U.S. Gov't].
10. Genant HK, Jiang Y, Peterfy C, Lu Y, Redei J, Countryman PJ. Assessment of rheumatoid arthritis using a modified scoring method on digitized and original radiographs. *Arthritis Rheum* 1998 Sep;41(9):1583–90.
11. Neumann G, dePablo P, Finckh A, Chibnik LB, Wolfe F, Duryea J. Patient repositioning reproducibility of joint space width measurements on hand radiographs. *Arthritis Care Res* 2011 Feb;63(2):203–7 [Comparative Study Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't].
12. Finckh A, de Pablo P, Katz JN, Neumann G, Lu Y, Wolfe F, *et al.* Performance of an automated computer-based scoring method to assess joint space narrowing in rheumatoid arthritis: a longitudinal study. *Arthritis Rheum* 2006 May;54(5):1444–50 [Randomized Controlled Trial Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't].
13. Haugen IK, Ostergaard M, Eshed I, McQueen FM, Bird P, Gandjbakhch F, *et al.* Iterative development and reliability of the OMERACT hand osteoarthritis MRI scoring system.

- J Rheumatol 2014 Feb;41(2):386–91 [Research Support, Non-U.S. Gov't].
14. Haugen IK, Lillegraven S, Slatkowsky-Christensen B, Haavardsholm EA, Sesseng S, Kvien TK, *et al.* Hand osteoarthritis and MRI: development and first validation step of the proposed Oslo Hand Osteoarthritis MRI score. *Ann Rheum Dis* 2011 Jun;70(6):1033–8 [Research Support, Non-U.S. Gov't Validation Studies].
 15. Haugen IK, Bøyese P, Slatkowsky-Christensen B, Sesseng S, Van Der Heijde D, Kvien TK. Associations between MRI-defined synovitis, bone marrow lesions and structural features and measures of pain and physical function in hand osteoarthritis. *Ann Rheum Dis* 2011 Nov 25.
 16. Conaghan PG, Durez P, Alten RE, Burmester GR, Tak PP, Klareskog L, *et al.* Impact of intravenous abatacept on synovitis, osteitis and structural damage in patients with rheumatoid arthritis and an inadequate response to methotrexate: the ASSET randomised controlled trial. *Ann Rheum Dis* 2013 Aug;72(8):1287–94.
 17. Kamishima T, Tanimura K, Shimizu M, Matsuhashi M, Fukae J, Kon Y, *et al.* Monitoring anti-interleukin 6 receptor antibody treatment for rheumatoid arthritis by quantitative magnetic resonance imaging of the hand and power Doppler ultrasonography of the finger. *Skelet Radiol* 2011 Jun;40(6):745–55.
 18. Ostergaard M, Emery P, Conaghan PG, Fleischmann R, Hsia EC, Xu W, *et al.* Significant improvement in synovitis, osteitis, and bone erosion following golimumab and methotrexate combination therapy as compared with methotrexate alone: a magnetic resonance imaging study of 318 methotrexate-naïve rheumatoid arthritis patients. *Arthritis Rheum* 2011 Dec;63(12):3712–22.
 19. Peterfy CG, Olech E, Dicarolo JC, Merrill JT, Countryman PJ, Gaylis NB. Monitoring cartilage loss in the hands and wrists in rheumatoid arthritis with magnetic resonance imaging in a multi-center clinical trial: IMPRESS (NCT00425932). *Arthritis Res Ther* 2013 Mar 20;15(2):R44.
 20. Roemer FW, Kwok CK, Hannon MJ, Crema MD, Moore CE, Jakicic JM, *et al.* Semiquantitative assessment of focal cartilage damage at 3T MRI: a comparative study of dual echo at steady state (DESS) and intermediate-weighted (IW) fat suppressed fast spin echo sequences. *Eur J Radiol* 2011 Nov;80(2):e126–31.
 21. Del Grande F, Santini F, Herzka DA, Aro MR, Dean CW, Gold GE, *et al.* Fat-suppression techniques for 3-T MR imaging of the musculoskeletal system. *Radiographics* 2014 Jan–Feb;34(1):217–33.
 22. Strickland CD, Kijowski R. Morphologic imaging of articular cartilage. *Magnetic Reson Imaging Clin N Am* 2011 May;19(2):229–48.
 23. Crema MD, Roemer FW, Marra MD, Burstein D, Gold GE, Eckstein F, *et al.* Articular cartilage in the knee: current MR imaging techniques and applications in clinical practice and research. *Radiographics* 2011 Jan–Feb;31(1):37–61.
 24. Thomsen HS. Nephrogenic systemic fibrosis: history and epidemiology. *Radiol Clin N Am* 2009 Sep;47(5):827–31. vi.
 25. Miese F, Buchbender C, Scherer A, Wittsack HJ, Specker C, Schneider M, *et al.* Molecular imaging of cartilage damage of finger joints in early rheumatoid arthritis with delayed gadolinium-enhanced magnetic resonance imaging. *Arthritis Rheum* 2012 Feb;64(2):394–9.
 26. Williams A, Shetty SK, Burstein D, Day CS, McKenzie C. Delayed gadolinium enhanced MRI of cartilage (dGEMRIC) of the first carpometacarpal (1CMC) joint: a feasibility study. *Osteoarthritis Cartilage* 2008 Apr;16(4):530–2.
 27. Felson DT, Nevitt MC. Blinding images to sequence in osteoarthritis: evidence from other diseases. *Osteoarthritis Cartilage* 2009 Mar;17(3):281–3.
 28. Guermazi A, Hunter DJ, Li L, Benichou O, Eckstein F, Kwok CK, *et al.* Different thresholds for detecting osteophytes and joint space narrowing exist between the site investigators and the centralized reader in a multicenter knee osteoarthritis study—data from the Osteoarthritis Initiative. *Skelet Radiol* 2012 Feb;41(2):179–86 [Comparative Study Multicenter Study Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't].
 29. Maheu E, Cadet C, Gueneugues S, Ravaud P, Dougados M. Reproducibility and sensitivity to change of four scoring methods for the radiological assessment of osteoarthritis of the hand. *Ann Rheum Dis* 2007 Apr;66(4):464–9.
 30. Bijsterbosch J, Haugen IK, Malines C, Maheu E, Rosendaal FR, Watt I, *et al.* Reliability, sensitivity to change and feasibility of three radiographic scoring methods for hand osteoarthritis. *Ann Rheum Dis* 2011 Aug;70(8):1465–7 [Comparative Study Multicenter Study Research Support, Non-U.S. Gov't].
 31. Botha-Scheepers S, Watt I, Breedveld FC, Kloppenburg M. Reading radiographs in pairs or in chronological order influences radiological progression in osteoarthritis. *Rheumatology (Oxford)* 2005 Nov;44(11):1452–5 [Comparative Study].
 32. Kellgren JH, Lawrence JS. Radiological assessment of osteoarthritis. *Ann Rheum Dis* 1957 Dec;16(4):494–502.
 33. Marshall M, Dziedzic KS, van der Windt DA, Hay EM. A systematic search and narrative review of radiographic definitions of hand osteoarthritis in population-based studies. *Osteoarthritis Cartilage* 2008 Feb 1;16(2):219–26.
 34. Kellgren J, Jeffrey MR, Ball J. *The Epidemiology of Chronic Rheumatism Volume II Atlas of Standard Radiographs of Arthritis*. Oxford: Blackwell Scientific Publications; 1963 Feb 14.
 35. Spector TD, Cooper C. Radiographic assessment of osteoarthritis in population studies: whither Kellgren and Lawrence? *Osteoarthritis Cartilage* 1993 Oct 1;1(4):203–6.
 36. Kallman DA, Wigley FM, Scott WW, Hochberg MC, Tobin JD. New radiographic grading scales for osteoarthritis of the hand. Reliability for determining prevalence and progression. *Arthritis Rheum* 1989 Dec 1;32(12):1584–91.
 37. Altman RD, Hochberg M, Murphy Jr WA, Wolfe F, Lequesne M. *Atlas of individual radiographic features in osteoarthritis*. *Osteoarthritis Cartilage* 1995 Sep;(3 Suppl A):3–70 [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S. Research Support, U.S. Gov't, P.H.S.].
 38. Altman RD, Gold GE. *Atlas of individual radiographic features in osteoarthritis, revised*. *Osteoarthritis Cartilage* 2007;(15 Suppl A):A1–A56 [Research Support, Non-U.S. Gov't].
 39. Lane NE, Nevitt MC, Genant HK, Hochberg MC. Reliability of new indices of radiographic osteoarthritis of the hand and hip and lumbar disc degeneration. *J Rheumatol* 1993 Nov 1;20(11):1911–8.
 40. Kessler S, Dieppe P, Fuchs J, Stürmer T, Günther KP. Assessing the prevalence of hand osteoarthritis in epidemiological studies. The reliability of a radiological hand scale. *Ann Rheum Dis* 2000 Apr 1;59(4):289–92.
 41. Verbruggen G, Veys EM. Numerical scoring systems for the progression of osteoarthritis of the finger joints. *Rev Rhum Engl Ed* 1995 Jun 1;62(6 Suppl 1):275–32S.
 42. Verbruggen G, Veys EM. Numerical scoring systems for the anatomic evolution of osteoarthritis of the finger joints. *Arthritis Rheum* 1996 Feb;39(2):308–20.

43. Verbruggen G, Wittoek R, Cruyssen BV, Elewaut D. Morbid anatomy of 'erosive osteoarthritis' of the interphalangeal finger joints: an optimised scoring system to monitor disease progression in affected joints. *Ann Rheum Dis* 2010 May;69(5):862–7.
44. Dacre JE, Coppock JS, Herbert KE, Perrett D, Huskisson EC. Development of a new radiographic scoring system using digital image analysis. *Ann Rheum Dis* 1989 Mar 1;48(3):194–200.
45. Angwin J, Lloyd A, Heald G, Nepom G, Binks M, James MF. Radiographic hand joint space width assessed by computer is a sensitive measure of change in early rheumatoid arthritis. *J Rheumatol* 2004 Jun 1;31(6):1050–61.
46. Sharp JT, Gardner JC, Bennett EM. Computer-based methods for measuring joint space and estimating erosion volume in the finger and wrist joints of patients with rheumatoid arthritis. *Arthritis Rheum* 2000 Jun 1;43(6):1378–86.
47. Sharp JT, van der Heijde D, Angwin J, Duryea J, Moens HJ, Jacobs JW, et al. Measurement of joint space width and erosion size. *J Rheumatol* [Review] 2005 Dec;32(12):2456–61.
48. van 't Klooster R, Hendriks EA, Watt I, Kloppenburg M, Reiber JHC, Stoel BC. Automatic quantification of osteoarthritis in hand radiographs: validation of a new method to measure joint space width. *Osteoarthritis Cartilage* 2008 Jan 1;16(1):18–25.
49. Kwok WY, Bijsterbosch J, Malm SH, Biermasz NR, Huetink K, Nelissen RG, et al. Validity of joint space width measurements in hand osteoarthritis. *Osteoarthritis Cartilage* 2011 Sep 5.
50. Haugen IK, Boyesen P, Slatkowsky-Christensen B, Sesseng S, Bijsterbosch J, van der Heijde D, et al. Comparison of features by MRI and radiographs of the interphalangeal finger joints in patients with hand osteoarthritis. *Ann Rheum Dis* 2012 Mar;71(3):345–50 [Comparative Study Evaluation Studies Research Support, Non-U.S. Gov't].
51. Kwok WY, Kortekaas MC, Reijnierse M, van der Heijde D, Bloem JL, Kloppenburg M. MRI in hand osteoarthritis: validation of the Oslo hand osteoarthritis MRI-scoring method and association with pain. *Osteoarthritis Cartilage* 2011 Sep 29: 1–31 (abstract).
52. Keen HI, Lavie F, Wakefield RJ, D'Agostino MA, Hammer HB, Hensor E, et al. The development of a preliminary ultrasonographic scoring system for features of hand osteoarthritis. *Ann Rheum Dis* 2008 May;67(5):651–5 [Evaluation Studies Research Support, Non-U.S. Gov't Review].
53. Mathiessen A, Haugen IK, Slatkowsky-Christensen B, Bøyesen P, Kvien TK, Hammer HB. Ultrasonographic assessment of osteophytes in 127 patients with hand osteoarthritis: exploring reliability and associations with MRI, radiographs and clinical joint findings. *Ann Rheum Dis* 2013 Jan;72(1):51–6.
54. Hammer HB, Bolton-King P, Bakkeheim V, Berg TH, Sundt E, Kongtorp AK, et al. Examination of intra and interrater reliability with a new ultrasonographic reference atlas for scoring of synovitis in patients with rheumatoid arthritis. *Ann Rheum Dis* 2011 Nov 1;70(11):1995–8.
55. Iagnocco A, Conaghan PG, Aegerter P, Moller I, Bruyn GA, Chary-Valckenaere I, et al. The reliability of musculoskeletal ultrasound in the detection of cartilage abnormalities at the metacarpo-phalangeal joints. *Osteoarthritis Cartilage* 2012 Oct;20(10):1142–6.
56. Kortekaas MC, Kwok WY, Reijnierse M, Watt I, Huizinga TW, Kloppenburg M. Pain in hand osteoarthritis is associated with inflammation: the value of ultrasound. *Ann Rheum Dis* 2010 Jul;69(7):1367–9.
57. Kortekaas MC, Kwok WY, Reijnierse M, Huizinga TW, Kloppenburg M. Osteophytes and joint space narrowing are independently associated with pain in finger joints in hand osteoarthritis. *Ann Rheum Dis* 2011 Oct;70(10):1835–7.
58. Vlychou M, Koutroumpas A, Malizos K, Sakkas LI. Ultrasonographic evidence of inflammation is frequent in hands of patients with erosive osteoarthritis. *Osteoarthritis Cartilage* 2009 Oct;17(10):1283–7.
59. Wittoek R, Carron P, Verbruggen G. Structural and inflammatory sonographic findings in erosive and non-erosive osteoarthritis of the interphalangeal finger joints. *Ann Rheum Dis* 2010 Dec;69(12):2173–6.
60. Wittoek R, Jans L, Lambrecht V, Carron P, Verstraete K, Verbruggen G. Reliability and construct validity of ultrasonography of soft tissue and destructive changes in erosive osteoarthritis of the interphalangeal finger joints: a comparison with MRI. *Ann Rheum Dis* 2011 Feb;70(2):278–83.
61. Kortekaas MC, Kwok WY, Reijnierse M, Huizinga TW, Kloppenburg M. In erosive hand osteoarthritis more inflammatory signs on ultrasound are found than in the rest of hand osteoarthritis. *Ann Rheum Dis* 2013 Jun;72(6):930–4.
62. Keen HI, Wakefield RJ, Grainger AJ, Hensor EM, Emery P, Conaghan PG. Can ultrasonography improve on radiographic assessment in osteoarthritis of the hands? A comparison between radiographic and ultrasonographic detected pathology. *Ann Rheum Dis* 2008 Aug;67(8):1116–20 [Comparative Study].
63. Keen HI, Wakefield RJ, Grainger AJ, Hensor EM, Emery P, Conaghan PG. An ultrasonographic study of osteoarthritis of the hand: synovitis and its relationship to structural pathology and symptoms. *Arthritis Rheum* 2008 Dec 15;59(12):1756–63.
64. Conaghan PG, D'Agostino MA, Le Bars M, Baron G, Schmidely N, Wakefield R, et al. Clinical and ultrasonographic predictors of joint replacement for knee osteoarthritis: results from a large, 3-year, prospective EULAR study. *Ann Rheum Dis* 2010 Apr;69(4):644–7 [Multicenter Study Research Support, Non-U.S. Gov't].
65. Keen HI, Mease PJ, Bingham 3rd CO, Giles JT, Kaeley G, Conaghan PG. Systematic review of MRI, ultrasound, and scintigraphy as outcome measures for structural pathology in interventional therapeutic studies of knee arthritis: focus on responsiveness. *J Rheumatol* [Review] 2011 Jan;38(1):142–54.
66. D'Agostino MA, Conaghan P, Le Bars M, Baron G, Grassi W, Martin-Mola E, et al. EULAR report on the use of ultrasonography in painful knee osteoarthritis. Part 1: prevalence of inflammation in osteoarthritis. *Ann Rheum Dis* 2005 Dec;64(12):1703–9 [Multicenter Study Research Support, Non-U.S. Gov't].