EXTENDED REPORT

Global ultrasound assessment of structural lesions in osteoarthritis: a reliability study by the OMERACT ultrasonography group on scoring cartilage and osteophytes in finger joints

H B Hammer,1 A Iagnocco,2 A Mathiessen,1 E Filippucci,3 F Gandjbakhch4,5 M C Kortekaas,6,7 I Möller,8 E Naredo,9 R J Wakefield,10 P Aegerter,11,12 M-A D’Agostino13,14

ABSTRACT
Objective Ultrasonography is sensitive for the evaluation of cartilage pathology and degree of osteophytes in patients with hand osteoarthritis (OA). High consistency of assessments is essential, and the OMERACT (Outcome Measures in Rheumatology) ultrasonography group took the initiative to explore the reliability of a global ultrasonography score in patients with hand OA using semiquantitative ultrasonography score of cartilage and osteophytes in finger joints.

Methods Ten patients with hand OA were examined by 10 experienced sonographers over the course of two days. Semiquantitative scoring (0–3) was performed on osteophytes (carpo-metacarpal 1, metacarpo-phalangeal (MCP) 1–5, proximal interphalangeal 1–5 and distal interphalangeal 2–5 joints bilaterally with an ultrasonography atlas as reference) and cartilage pathology (MCP 2–5 bilaterally). A web-based exercise on static cartilage images was performed a month later. Reliability was assessed by use of weighted κ analyses.

Results Osteophyte scores were evenly distributed, and the intraobserver and interobserver reliabilities were substantial to excellent (κ range 0.68–0.89 and mean κ 0.65 (day 1) and 0.67 (day 2), respectively). Cartilage scores were unevenly distributed, and the intraobserver and interobserver reliability was fair to moderate (κ range 0.46–0.66 and mean κ 0.39 (day 1) and 0.33 (day 2), respectively). The web-based exercise showed acceptable agreement for cartilage being normal (κ 0.47) or with complete loss (κ 0.68), but poor for the intermediate scores (κ 0.22–0.30).

Conclusions Use of the present semiquantitative ultrasonography scoring system for cartilage pathology in hand OA is not recommended (while normal or total loss of cartilage may be assessed). However, the OMERACT ultrasonography group will endorse the use of semiquantitative scoring of osteophytes with the ultrasonography atlas as reference.

INTRODUCTION
Osteoarthritis (OA) is the common joint disease in the middle aged to elderly population. Pathological features of the disease involves a gradual breakdown of the hyaline cartilage, formation of excess bone at the joint margins (osteophytes) and inflammation of the synovial membrane.1 According to a recent prevalence study of persons aged 40–84 years, radiographic hand OA is widespread (women 44% and men 38%),2 and it may cause pain and stiffness. Indeed, patients with hand OA have been found to have similar levels of symptoms in the joints as in patients with rheumatoid arthritis.3 To date, there is no effective disease modifying medication for these patients, but it is hoped that an increased interest in understanding the pathology of this disease might result in new treatment paradigms and the development of new therapeutic agents. Imaging is one new avenue undergoing evaluation in OA which might provide added value in the assessment of patients with OA.4

The Osteoarthritis Research Society International (OARSI) has previously published guidelines for conducting clinical trials of hand OA, recommending conventional radiography as the standard for assessing structural outcomes. However, they acknowledge that other novel imaging techniques such as ultrasound and MRI may play an additional role but require further validation.5 There is a growing interest in the use of ultrasound for the assessment of OA, as modern ultrasonography allows high-resolution planar and dynamic imaging of joints and it is also well suited for assessing multiple joints.6,7 The validation of ultrasonography as an outcome measure for evaluating hand OA is an area of interest for the OMERACT (Outcome Measures in Rheumatology) ultrasonography group, and exploring the reliability of evaluating the elementary lesions of structural damage such as cartilage abnormalities and osteophytes, is an important step in the research agenda.

The OMERACT ultrasonography group in collaboration with OARSI has previously obtained consensus on the definitions of cartilage pathologies and demonstrated moderate to good reliability for detecting these changes in hyaline cartilage at the metacarpo-phalangeal (MCP) joints in patients with hand OA.6 Based on a previously developed scoring system for osteophytes in finger joints,8 a recent study achieved consensus on evaluating osteophytes in MCP, proximal and distal interphalangeal (PIP and DIP) joints as well as carpo-metacarpal (CMC)1 joints and found ultrasonography to be highly sensitive for detecting osteophytes compared with MRI and radiography.9 Additionally, the study included an ultrasonography atlas of semiquantitative scores
(0–3) of finger joint osteophytes, and the use of this atlas resulted in excellent reliability for scoring of static ultrasonography images. As a further stage in the ultrasonography evaluation of hand OA, the OMERACT ultrasonography group decided to perform a reliability study on the global assessment of structural lesions where grading of pathology was introduced for the first time. The objective of the present study was to explore the reliability of highly experienced sonographers in performing semiquantitative ultrasonography scoring of cartilage pathology and osteophytes in the finger joints of patients with hand OA.

**METHODS**

**Design of the study**

The study was performed on three consecutive days with consensus and training the first day followed by a reliability exercise on patients over the next 2 days. One month after the meeting, a further reliability exercise was performed on still images of cartilage lesions via a website.

**Training session**

**Osteophyte evaluation**

The sonographers agreed to use the previously described semiquantitative scoring system for grading osteophytes (0=none, 1=minor, 2=moderate, 3=large size of osteophytes). The proximal and distal parts of the joints were evaluated as a whole with the largest osteophyte defining the score independently of the number and location of additional osteophytes in the same joint. The ultrasonography atlas of representative images of each score for CMC, MCP, PIP, and DIP joints was used as reference. Since only two of the present sonographers were familiar with this scoring system, a reliability test on static ultrasonography images of osteophytes was performed 3 months prior to the practical exercise. Ten images of each of the 15 joints to be examined (CMC1, MCP 1–5, PIP 1–5, and DIP 2–5), that is, a total of 150 images, were sent by e-mail and scored individually by all the sonographers two times (several weeks apart and in a new order the second time) using the ultrasonography atlas of osteophytes as reference. Excellent reliability was found (κ values >0.9) and, therefore, no practical training on scoring osteophytes was performed before the exercise on patients.

**Cartilage evaluation**

On the basis of the previously published study assessing hyaline cartilage abnormalities according to a dichotomous score, a combination of different abnormalities was set up and the sonographers agreed on a semiquantitative scoring system for evaluating the severity of cartilage involvement on a 4-point scale (see box 1). Since the reliability for detecting the presence/absence of different forms of cartilage lesions had previously been found to be good, a web-based exercise on static cartilage ultrasonography images before the practical exercise was not undertaken. However, on the first day of the study, the sonographers reached consensus on the scanning technique and on how to use the new semiquantitative scoring system by grading formerly collected static images of different cartilage lesions as well as performing practical ultrasonography training on the evaluation of cartilage in OA joints including scan positioning of the joint.

**Patients**

Ten patients with hand OA were recruited from the outpatient clinic at the Diakonhjemmet Hospital (Oslo, Norway). All were women (median (range) age 74.5 (53–77) years) fulfilling the American College of Rheumatology classification criteria for hand OA. The presence of any inflammatory joint disease was an exclusion criterion. The study was approved by the regional ethics committee, and the patients gave their written informed consent.

**Equipment**

Five identical General Electric logic E9 machines (GE, Medical Systems, Milwaukee, Wisconsin, USA), equipped with two

---

**Box 1 Ultrasound semiquantitative scoring system for cartilage abnormalities in hand osteoarthritis:**

- 0=normal cartilage (anechoic structure, normal margins of cartilage);
- 1=loss of anechoic structure and/or focal thinning of cartilage layer OR irregularities and/or loss of sharpness of at least one cartilage margin;
- 2=loss of anechoic structure and/or focal thinning of cartilage layer AND irregularities and/or loss of sharpness of at least one cartilage margin;
- 3=focal absence or complete loss of the cartilage layer.

---

**Figure 1** (A) Ultrasound scoring of osteophytes was performed on the dorsal side and sliding the probe from side to side on metacarpophalangeal (MCP), proximal and distal interphalangeal (PIP and DIP) joints, and on the radio-palmar side of the carpo-metacarpal (CMC) 1 joint. (B) Ultrasound scoring of cartilage in the MCP joints was performed with maximal flexion of the joint, abundant gel and scanning of the mid-portion of the metacarpal head.
multifrequency linear probes (hockey stick 8–18 MHz used for scoring of cartilage and regular probe 6–15 MHz used for scoring of osteophytes) both operating at a frequency of 15 MHz were used (the ultrasonography machines were supplied free of charge by GE). Previous to the reliability test, the same B-mode setting with 50% gain and positioning of the focus at the level of interest was fixed in all machines and not modified during the study.

Sonographers
Of the 10 sonographers from six countries participating in the study, nine were rheumatologists, experts in musculoskeletal ultrasonography and members of the OMERACT ultrasonography group, while one was a trainee fellow in rheumatology, highly experienced in scoring of osteophytes and had participated in the development of the ultrasonography atlas. Additionally, six of the sonographers had previously performed ultrasonography reliability studies on cartilage in MCP joints.

Reliability exercise on patients
On each of the 2 days of reliability examination on patients with hand OA, all the sonographers assessed five patients in two rounds, always with an interval of at least 3 h between the two evaluations of the same patient. The patients were positioned in separate examination rooms with their hands resting on a small table close to the ultrasonography machine. The sonographers rotated between the rooms, and they were given a maximum of 20 min to complete scoring of both cartilage and osteophytes. Each ultrasonography examination included bilateral scoring of cartilage in MCP 2–5 (ie, 8 joints) and osteophytes in CMC1, MCP 1–5, PIP1–5 and DIP 2–5 (ie, 30 joints). The results of the examinations were given orally to a student seated in the same room, and the scores were immediately punched into an Excel file.

Osteophytes were assessed by longitudinal scanning on extended joints with swiping of the probe from side to side of the dorsal aspect of MCP, PIP and DIP joints and at the radiopalmar side of CMC1 joints (figure 1A).

For the evaluation of cartilage, the MCP joints were kept in maximal flexion (ie, close to 90°) and a longitudinal dorsal scan was performed at the level of the mid-portion of the metacarpal head (figure 1B). There was particular attention to keep the probe perpendicular to the surface of the hyaline cartilage, which was obtained by performing slight sweeping movements with the transducer over the region of interest. All examinations were performed by applying abundant amounts of gel to the skin to provide an appropriate acoustic interface.

Reliability web-based exercise on cartilage pathology
Due to unsatisfactory results on the scoring of cartilage in patients with hand OA, a web-based exercise on static images of cartilage was subsequently performed 1 month after the reliability exercise on patients. Since scoring of static images was supposed to be more reliable than scanning and scoring on patients, the aim of this additional exercise was to examine whether the presently used semiquantitative grading system of cartilage was appropriate for scoring of cartilage pathology. In order to ensure a better adherence to the proposed scoring system, a new ultrasonography atlas of MCP cartilage was developed (figure 2) and sent to the sonographers prior to the reliability exercise and used as reference. This atlas included representative images of each of the four levels (0–3) of the semiquantitative score that had been agreed upon. The web-based exercise contained 125 representative images of MCP cartilage, including a high number of images that was scored two times for intraobserver calculation.
Statistical analysis

The mean prevalence of lesions observed in each scoring session was calculated for both cartilage and osteophytes. The observed agreement between ultrasonographers during the two rounds of semiquantitative scoring of cartilage and osteophytes, as well as the intraobserver and interobserver reliabilities in patients, was calculated. Reliability was assessed by using standard Cohen’s $\kappa$ for binary evaluation (0 vs 1, 2, 3) and weighted $\kappa$ with absolute weights for the semiquantitative evaluation (0–3). While intraobserver coefficients were evaluated on pairs of measures performed by the same sonographer, calculation of interobserver coefficient was exclusively based on the first measure of these pairs. Global intraobserver reliability was obtained by calculating the mean $\kappa$ for all $n$ (n−1)/2 pairs, $n$ being the number of sonographers (ie, Light’s $\kappa$). $\kappa$ Coefficients were interpreted according to Landis and Koch (poor=0; slight=0.01–0.20; fair=0.21–0.40; moderate=0.41–0.60; substantial=0.61–0.80; excellent=0.81–1.00).

RESULTS

Tables 1 and 2 show the results of the intraobserver and interobserver reliabilities on scoring osteophytes and cartilage in patients.

Reliability of scoring osteophytes

The prevalence of osteophyte scores in patients with hand OA was quite evenly distributed (table 2). Both the intraobserver and interobserver reliabilities were highly satisfactory. The intraobserver reliability for scoring 300 joints varied from substantial to excellent (table 1). The interobserver reliability was also very good for both binary and semiquantitative scoring during the two study days (table 2).

Reliability of scoring cartilage abnormalities

The prevalence of cartilage scores were unevenly distributed (table 2). Quite variable intraobserver reliability was found for the scoring of 80 joints regarding both binary and semiquantitative scoring, and it ranged from poor to moderate (table 1). Additionally, the interobserver reliability for scoring of cartilage pathology was only fair on the two study days independently of the scoring (binary or semiquantitative) (table 2).

Reliability of the web-based exercise on cartilage abnormalities

The results of the reliability on the web-based exercise are shown in table 3. In order to examine the disagreement observed in detecting cartilage lesions during the reliability exercise on patients, the cartilage abnormalities in each image during the web-based exercise were evaluated as previously described during the exercise on patients (see box). Despite inclusion of a reference atlas, the results were variable (table 3). A good agreement was found for defining grade 0, normal cartilage (even if there was a low $\kappa$ due to the presence of few normal images, ie, $\kappa$ paradox) and grade 3, complete loss of cartilage. However, agreement was only fair, and $\kappa$ was poor for the two intermediate scores (table 3).

DISCUSSION

To our knowledge, this is the first reliability study on a global ultrasonography scoring of structural lesions (ie, osteophytes and cartilage abnormalities) in hand OA finger joints. Encouraging results were found for the scoring of osteophytes, where substantial to excellent intraobserver and interobserver reliabilities were seen using an ultrasonography atlas of osteophyte score as reference. High agreement has been found between ultrasonography and MRI for assessments of osteophytes in finger joints in hand OA, and ultrasonography has been shown to detect more joints with osteophytes than conventional radiography even if the detection of osteophytes by ultrasonography is limited to joint margins accessible for the beam. We found high reliability for the scoring of osteophytes by sonographers who were not trained in advance, and the present use of a reference atlas may have contributed to this result. An atlas can be helpful by permitting the ultrasonographer to have a direct comparison between the scanned ultrasonography findings and examples of defined scoring level images in the atlas. It is thus plausible that the use of a reference atlas can facilitate a multicentre reliability during therapeutic clinical trials.

Even if the DIP and PIP joints are primarily involved in finger OA, the cartilage in these joints was not examined in the present study because of technical ultrasonography limitations due to the severity of OA in this population. To evaluate cartilage by ultrasonography, the probe has to be perpendicular to the cartilage surface, which is not obtainable in patients with hand OA who have limited flexion of these joints. Additionally, frequently, there are dorsal osteophytes in DIP and PIP joints in these patients, which is a hindrance for the ultrasonography beam. However, since the MCP joints are often involved in advanced hand OA, these joints are representative for OA cartilage as well as being usually accessible for ultrasonography examinations.

For cartilage scoring, we found moderate intraobserver, but only fair interobserver reliability. The interobserver reliability was not significantly improved when static images were scored, even if a new reference ultrasonography atlas of cartilage pathology was used as background. However, the present study showed good agreement for ultrasonography scoring on static images of cartilage when scoring normal cartilage (ie, no abnormalities, grade 0) as well as for complete loss of the cartilage layer (grade 3). The poor reliability of the two intermediate scores can be explained by the fact that the proposed definitions could not help to sufficiently

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Intraobserver reliability of cartilage pathology and osteophytes in patients with hand osteoarthritis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cartilage</td>
<td>Osteophytes</td>
</tr>
<tr>
<td>Range of observed agreement</td>
<td>Range of $\kappa$ (95% CI)</td>
</tr>
<tr>
<td>0–3</td>
<td>0.46–0.66</td>
</tr>
<tr>
<td>0.57–0.81</td>
<td>(0.35–0.68) to (0.67–0.86)</td>
</tr>
</tbody>
</table>
distinguish between the two grades (ie, minimal and moderate cartilage abnormalities).

Since ultrasonography may be used to evaluate the thickness and quality of cartilage, it is a promising method for detecting early cartilage pathology in hand OA joints.20 However, high reliability is a requirement for inclusion of ultrasonography evaluation in the clinical armamentarium on these patients. Since the previous OMERACT study experienced good reliability for evaluating the presence of cartilage pathology in MCP joints,6 the aim of the present study was to extend the evaluation into a semiquantitative grading system of cartilage pathology. A previous study had shown good reliability for evaluation of different degrees of pathology in MCP joint cartilage in patients with rheumatoid arthritis.1 In the present study, however, even if highly experienced sonographers agreed on the scoring system, the interobserver reliability was not satisfactory. Practical difficulties in scanning of cartilage may explain the result.21 For example, to obtain exactly the same ultrasonography image, the scanning must include the same ultrasonography beam angle. However, this may not have occurred for all scanning positions, which would influence the scoring.22 Additionally, since the cartilage may not be uniformly damaged, obtaining high reliability would require the sonographers to evaluate exactly the same part of the cartilage which is difficult to control for. Hence, even if this study did not support the use of the present semiquantitative scoring system of cartilage pathology, there was good agreement for scoring cartilage with normal cartilage as well as complete loss of the layer. This supports the previous results of the OMERACT ultrasonography group.6

Our study has several strengths. First of all, only highly experienced sonographers participated in the study and only high-end ultrasonography machines with similar settings were used. Additionally, it was the first reliability study on hand OA with inclusion of an ultrasonography atlas as reference. However, there were a number of weaknesses including the fact that as only five ultrasonography machines were available, the study had to be performed during two consecutive days. On the other hand, the reliability results from the two days were similar. A major limitation could be the absence of a scoring exercise on static ultrasonography images of cartilage pathology prior to the scoring on patients. It could be speculated that an early finding of low reliability on scoring of static images could have changed the protocol for the cartilage part of the reliability study.

Based on the present study, for clinical trials, the OMERACT ultrasonography group will currently not recommend the use of the presently described semiquantitative ultrasonography score for assessing cartilage pathology in patients with hand OA. However, ultrasonography may be used to detect absence of abnormalities or very severe damage (ie, complete loss of cartilage). On the other hand, the group will endorse the use of a semiquantitative scoring of osteophytes with the ultrasonography atlas as reference. The present finding of high reliability for scoring of osteophytes appears to suggest that ultrasonography is a promising tool for diagnosing and follow-up of finger joints in patients with hand OA.

**Author affiliations**

1Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway
2Rheumatology Unit, Sapienza Università di Roma, Rome, Italy
3Clinica Reumatologica, Università Politecnica delle Marche, Jesi (Ancona), Italy
4Department of Rheumatology, AF-HP, Pitié-Salpêtrière Hospital, Paris, France
5UPMC Université Paris 06, Paris, France
6Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
Provenance and peer review

Abbvie and Bristol-Myers Squib. consultancy work for AbbVie and BMS. M-AD has received speaker fees from
funding from MSD. RJW has received lecture honoraria from AbbVie and has done
writing the manuscript.

Contributors

HBH has organised the practical part of the study and has made substantial contributions to conception and design, acquisition of data, analysis and interpretation of data and writing the manuscript; AI organised the ultrasonography


Wittevek R, Lans J, Lambech V, et al. Reliability and construct validity of ultrasonography of soft tissue and destructive changes in erosive osteoarthritis of the interphalangeal

Kortekaas MC, Kock W-Y, Reijers M, et al. Osteophytes and joint space narrowing are independently associated with pain in finger joints in hand


Torp-Pedersen S, Bartels EM, Wilhelm J, et al. Articular cartilage thickness measured with US is not as easy as it appears: a systematic review of measurement


Clinical and epidemiological research

REFERENCES


3 Stålkovsksy-Christensen B, Mowinkel P, Loge JH, et al. Health-related quality of life in women with symptomatic hand osteoarthritis: a comparison with rheumatoid


6 Iagnocco A, Conaghan PG, Aegerter P, et al. The reliability of musculoskeletal ultrasound in the detection of cartilage abnormalities at the metacarpophalangeal


16 Wittveicter R, Lans J, Lambech V, et al. Reliability and construct validity of ultrasonography of soft tissue and destructive changes in erosive osteoarthritides of the interphalangeal

17 Kortekaas MC, Kock W-Y, Reijers M, et al. Osteophytes and joint space narrowing are independently associated with pain in finger joints in hand


21 Torp-Pedersen S, Bartels EM, Wilhelm J, et al. Articular cartilage thickness measured with US is not as easy as it appears: a systematic review of measurement


Global ultrasound assessment of structural lesions in osteoarthritis: a reliability study by the OMERACT ultrasonography group on scoring cartilage and osteophytes in finger joints

H B Hammer, A Iagnocco, A Mathiessen, E Filippucci, F Gandjbakhch, M C Kortekaas, I Möller, E Naredo, R J Wakefield, P Aegerter and M-A D'Agostino

Ann Rheum Dis 2016 75: 402-407 originally published online December 17, 2014
doi: 10.1136/annrheumdis-2014-206289

Updated information and services can be found at: http://ard.bmj.com/content/75/2/402

These include:

References
This article cites 22 articles, 8 of which you can access for free at: http://ard.bmj.com/content/75/2/402#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections

- Clinical diagnostic tests (1282)
- Radiology (1113)
- Radiology (diagnostics) (750)
- Degenerative joint disease (4641)
- Musculoskeletal syndromes (4951)
- Osteoarthritis (931)

Notes

To request permissions go to: http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to: http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to: http://group.bmj.com/subscribe/