The Value of an Automated Ultrasound System in the Detection of Synovitis

Authors
Ruediger Mueller1,2, Mathias Grunke2, Jörg Wendler3, Florian Schuch4, Karina Hofmann-Preiss5, Ina Boettger5, Rüdiger Jakobs6, Hendrik Schulze-Koops2, Johannes von Kempis1

Affiliations
1 Division of Rheumatology, Kantonsspital St. Gallen, St. Gallen, Switzerland
2 Munich University Hospital, Division of Rheumatology, Munich, Germany
3 Schwerpunktpraxis Rheumatologie, Rheumatologie, Erlangen, Germany
4 Schwerpunktpraxis Rheumatologie, Schwerpunktpraxis Rheumatologie, Erlangen, Germany
5 Institut für bildgebende Diagnostik und Therapie, BDT · MVZ Träger GmbH, Erlangen, Germany
6 Siemens Healthcare GmbH, Ultrasound, Erlangen, Germany

Key words
ultrasound, arthritides, hand

ABSTRACT
Background The detection of joint swelling caused by synovitis is important for the diagnosis of inflammatory arthritis. Ultrasound (US) and MRI have proven to be more sensitive and reliable than physical examination, but they are time-consuming and expensive. The automated breast volume scanner was developed to acquire serial B-mode pictures of the female breast and these can be analyzed in all three dimensions.

Objectives To analyze the value of automated B-mode ultrasound employing the ABVS system in detecting synovitis of the finger joints compared to manual ultrasound (mUS) and physical examination, using MRI as the gold standard.

Methods 19 consecutive patients suffering from active rheumatoid (n = 15) or psoriatic (n = 4) arthritis were included. Automated and mUS were conducted with a linear array (ACUSON S2000TM, 11 MHz). Multiplanar reconstruction enabled examination of the images for the presence of synovitis.

Results 90% of the hand joints were assessable by automated ultrasound. Automated US detected 12.0, mUS 14.2, MRI 13.4, and clinical examination 4.1 positive joints - i.e. joints with synovitis - on average per patient. The inter-observer reliability of both assessors for automated and mUS, MRI, and physical examination, was 66.9 %, 72.7 %, 95.1 %, and 88.9 %, respectively. 84.3 % of the joints classified as positive on MRI were confirmed by automated ultrasound, 85.5 % on mUS, and 36.0 % on physical examination. This translated into a sensitivity of 83.5 %, 85.5 %, and 36.0 % for the three methods, respectively.

Conclusion: Automated ultrasound is a promising ultrasound method for assessing small joints in patients with inflammatory arthritis.
Introduction

Background

In rheumatic diseases, the detection and quantification of synovitis is one of the most important diagnostic challenges. Detection of synovitis is the backbone not only of classification criteria for RA and other kinds of inflammatory arthritis [1–3], but also of composite scores indicative of disease activity [1, 4, 5]. To target remission in rheumatoid arthritis [6, 7], the assessment of synovitis is essential for therapeutic decisions. In daily practice, however, the reliable detection of synovitis can pose difficulties for the treating physician [8]. Physical examination, which has been the mainstay of diagnosis over decades, may lead to substantially different results between independent examiners [9, 10]. Even with standardized training programs, these differences can variable [11, 12]. Consequently, physical examination has limitations especially for the quantification of joint swelling. Ultrasound may be useful in this respect.

In recent years, high-resolution ultrasound has proven to be a very sensitive method for the detection of even small amounts of synovitis [13]. In this respect it is clearly superior to physical examination [14, 15]. Moreover, with the addition of the Color and Power Doppler technique, it can visualize inflammatory hyperperfusion of the synovium. In addition, bony erosions can be traced much earlier than with conventional radiography [16, 17]. Disadvantages of this method are its dependency on a skilled examiner, the difficult standardization and reproducibility, and the amount of time needed for a comprehensive examination [18, 19].

Magnetic resonance imaging (MRI) has a comparable sensitivity to manual ultrasound in the detection of synovitis and bony erosions [20, 21]. Inflammation can be illustrated by using the appropriate sequences and contrast media. The reproducibility is better than with manual ultrasound [22]. For the interpretation of both MRI and ultrasound generated images, a skilled assessor is needed. Moreover, this technique is expensive and constantly changing. This brings the RAMRIS score in clinical practice has to be shown and validated in the near future [23].

The “automated breast volume scanner” (ABVS, ACUSON S2000TM ABVS; Siemens Medical Solutions, Inc., Mountain View, CA) has recently been developed for the acquisition of serial consecutive B-mode pictures of the female breast [24, 25]. These pictures can be analyzed in all three dimensions and be reconstructed to 3D data sets. The ABVS can also be operated by an assistant medical technician. The acquired data are sent to a separate workstation to be independently analyzed by a Ultrasonologist with an interest in rheumatological diseases. Similar to conventional ultrasound, ABVS has shown a high sensitivity in detecting breast abnormalities and an excellent prediction of lesion size [26, 27]. Even the differentiation between benign and malignant lesions was comparable between ABVS and manual ultrasound [28]. In patients with suspected breast tumors, the detection rate was even higher compared to manual ultrasound [27]. The existence of a device that has been validated in the automatic generation of ultrasound pictures of the female breast led to the idea of modifying this machine to be useful in another type of soft tissue, i.e., the synovium.

The aim of our study was to investigate the value of automated three-dimensional ultrasound (referred to as automated US in the following). For this purpose we used the ABVS system in patients with inflammatory joint diseases. The focus was set on sensitivity and inter-observer reliability in comparison to physical examination and manual ultrasound (mUS), with MRI as the gold standard.

Methods

Study population and design

19 consecutive patients suffering from rheumatoid arthritis (n = 15) or psoriatic arthritis (n = 4) with a minimum of one swollen MCP (metacarpophalangeal) or PIP (proximal interphalangeal) joint were recruited. Patients fulfilled either the 2010 ACR/EULAR criteria for rheumatoid arthritis [29] or the CASPAR criteria for psoriatic arthritis [30]. Patients had to be older than 18 years of age. All of them gave written informed consent for participation in the trial, which was approved by the Ethics Committee of the University of Munich. The overall time needed for MRI (1 h), manual joint assessment (5 min), manual (30 min) and automated (20 min) US was around 2 h including the time needed for preparation and documentation (estimate). Manual ultrasound and manual joint assessment were conducted independently by two assessors. The interpretation of MRI and automated US data was also performed separately by two investigators each (s. "Manual Ultrasound" and "MRI", respectively).

Outcome parameter

The primary end point was to compare the detection of joint swelling caused due to synovitis by clinical examination, manual and automated US in relation to MRI as the gold standard.

Manual joint assessment

All individuals were clinically assessed according to the EULAR examination technique performed by two rheumatologists [11, 12]. This joint count includes the knees and all joints of the upper extremities including the distal interphalangeal joints. Every single joint is evaluated for the presence or absence of tenderness and swelling without graduation.

Manual ultrasound

Manual ultrasound was performed using the ACUSON S2000™ ultrasound system (Siemens Medical Solutions, Mountain View, USA) equipped with a large-format 50 mm linear array transducer with a bandwidth of 6–18 MHz.

Grayscale ultrasound was performed on all MCP, PIP and DIP joints of both hands from the dorsal and palmar view [31]. Each joint was analyzed independently by each of the assessors. Examinations were conducted without knowledge of the results obtained by the other assessors or those obtained using ABVS. All scanning data were recorded and stored. Both assessors were trained according to the guidelines of the DEGUM (Deutsche Gesellschaft für Ultraschall in der Medizin, the German ultrasound society) and had performed ultrasound as rheumatologists on a daily basis for many years.

Automated ultrasound (ABVS, aUS)

Automated ultrasound was also conducted with the ACUSON S2000™ ABVS (Siemens Medical Solutions, Mountain View, USA). The ABVS transducer was equipped with a linear array (5–14 MHz
angles for spin-echo images were used (90°/180°). MRI was interpreted by two board-certified radiologists with extensive experience in the interpretation of MRI results.

Definition of joint swelling and synovitis on ultrasound and MRI
Joint swelling in connection with ultrasound, either manual or automated, was defined for manual ultrasound according to Szkuła et al. [31]. For the purpose of this study, synovitis was only described as present or absent, but not graded. Each joint was analyzed independently by each of the ultrasonographers (RM and MG). Examinations were conducted without knowledge of the results obtained by the other ultrasonographer for conventional ultrasound or those obtained using automated US or MRI. All data obtained by ultrasound scanning were recorded and stored.

Missing data
If one joint could not be assessed by one of the four methods, this joint was completely removed from the analysis. If only one assessor defined a joint as not analyzable, this joint was also completely removed from the analysis.

Statistical analysis
Statistical comparisons were performed using the paired Student’s t-test or Fisher’s exact test. The analysis included all MCP,PIP, DIP and the carpal joint (total number: 40 joints). A Bonferroni correction was applied if necessary. A p-value of <0.05 was considered to be statistically significant. The inter-rater reliability was calculated employing Cohen’s kappa (k) test. The k was interpreted as follows: < 0 = poor agreement, 0-0.20 = slight agreement, 0.21-0.40 = fair agreement, 0.41-0.60 = moderate agreement, 0.61-0.80 = substantial agreement, 0.81-1.00 = almost perfect agreement. The analyses were performed using Excel (version 14.2.2) and/or the Graph pad Prism 5 software.

Results
Patient data
A total of 532 joints of 19 consecutive arthritis patients were analyzed in this cross-sectional study employing automated and manual US, MRI, and physical examination. Patients with at least one swollen MCP or PIP joint detected by physical examination of the recruiting physicians were included in the study. Patients had a mean age of 51 years (range: 19–71 years). None of the patients was in clinical remission, as defined by DAS 28 < 2.6. The mean DAS 28 was 4.5 (median: 4.5, range: 2.7–6.3). Patient data are summarized in Table 1.

Detection of joint swelling
All patients were analyzed for joint swelling by automated and manual US, MRI, and physical examination. Every joint was rated for joint swelling by two independent assessors for each of the techniques. A typical image of a swollen joint as detected by automated US is depicted in Fig. 2 and a movie with a typical film can be found as supplementary material. This figure also demonstrates that screening of multiple joints by automated US may also enable the detection of erosions. On average, 4.1 swollen and 6.8 tender joints were detected per patient by physical examination. By all three methods, automated and manual US and MRI, significantly more joints...
were classified as swollen as compared to physical examination (Table S1 supplementary material, Student’s T test, for all three p<0.00005). On average, 12.0, 14.2, and 13.4 swollen joints per patient were classified as swollen by automated and manual US, and MRI respectively (paired Student’s T test: manual US – automated US p = 0.0024, MRI – automated US p = 0.32, manual US – MRI p = 0.44, Fig. 3a).

Missing data

If a joint could not be assessed by one of the four methods (manual and automated US, MRI, physical examination) by at least one of the assessors, this joint was completely removed from the analysis. All joints were assessable by MRI. Only one joint could not be examined by manual US. 118 of a total of 1064 joint aspects (dorsal and palmar examination of 532 joints) could not be analyzed by automated US. The reason for this was mainly insufficient contact between the gel matrix and the scanner. This problem occurred predominantly with thumbs and heavily deformed joints.

Inter-observer reliability

As mentioned before, every joint was analyzed by automated and manual US, MRI, and physical examination by two assessors. This was done at the time of examination for manual US and physical examination, ABVS, and MRI, while ABVS and MRI were interpreted at another time. To determine the inter-observer reliability, the percentage of joints rated equally as swollen or not swollen by both assessors was determined. Joints were equally defined as swollen or not swollen by both assessors (automated and manual US, MRI, physical examination) in 66.9 % and 72.7 %, 95.1 %, and 88.9 % of cases, respectively (Fig. 3b). The inter-observer reliability was significantly greater for MRI and physical examination than for both ultrasound assessment methods (p = 0.0005, Fischer’s exact test). Cohen’s kappa for inter-observer reliability was 0.90 for MRI, 0.45 for manual US, 0.34 for automated US, and 0.77 for physical examination.

Table 1

<table>
<thead>
<tr>
<th>Number</th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex (female: male)</td>
<td>15:4</td>
</tr>
<tr>
<td>Age (years)</td>
<td>51 (19–71)</td>
</tr>
<tr>
<td>Diagnosis (RA: PsA)</td>
<td>15:4</td>
</tr>
<tr>
<td>Disease duration (years, range)</td>
<td>12.6 (0.3–39)</td>
</tr>
<tr>
<td>Rheumatoid factor pos.</td>
<td>80%</td>
</tr>
<tr>
<td>ACPA pos. *</td>
<td>60%</td>
</tr>
<tr>
<td>SJC (mean, range)</td>
<td>4.8 (1–20)</td>
</tr>
<tr>
<td>TJC (mean, range)</td>
<td>8.2 (1–20)</td>
</tr>
<tr>
<td>DAS 28 (mean, range)</td>
<td>4.5 (2.7–6.3)</td>
</tr>
<tr>
<td>HAQ (mean, range)</td>
<td>1.1 (0–2.37)</td>
</tr>
<tr>
<td>Patients’ global of disease activity (VAS)</td>
<td>41.1 (4–77)</td>
</tr>
<tr>
<td>Patients’ global of pain (VAS)</td>
<td>39.1 (5–75)</td>
</tr>
<tr>
<td>Physicians’ global of disease activity (VAS)</td>
<td>35.9 (19–57)</td>
</tr>
<tr>
<td>ESR (mm/h, range)</td>
<td>20.4 (5–39)</td>
</tr>
<tr>
<td>CRP (mg/l)</td>
<td>5.8 (0–14.6)</td>
</tr>
</tbody>
</table>

RA: rheumatoid arthritis
TJC: tender joint count (28 joint count)
SJC: swollen joint count (28 joint count)
PsA: psoriatic arthritis
ACPA: anti-citrullinated peptides antibodies
ESR: erythrocyte sedimentation rate
CRP: C-reactive protein
VAS: visual analogue scale (range 0–100)
DAS: disease activity score
HAQ: health assessment questionnaire
* calculated for RA patients only

Fig. 2

Three-dimensional visualization of synovitis (black arrows) and bony erosions (white arrow) of the second metacarpophalangeal joint in a patient with rheumatoid arthritis on aUS and MR images. The upper left picture shows the sagittal view a, the axial b and coronal c views are calculated by the integrated software. The axial view represents the original grayscale scan across the digit. The images D to F are corresponding T2-weighted MR images: d sagittal, e coronal, and f axial planes.
Positive confirmation of the gold standard (MRI) by the other methods

A double positive detection by the two MRI assessors was used as the gold standard. Since the analysis of joint swelling by automated US is not yet standardized, double and single positive detection in automated US by the observers was taken as comparison to the gold standard. For adequate comparability between the three methods of detection (physical examination, automated and manual US), double and single positivity was also used for physical examination and manual ultrasound.

In detail, 84.3% of the joints classified as positive by the gold standard (MRI) were confirmed by automated US, 85.5% by manual US, and 36.0% by physical examination (Fig. 3c). This translates into a sensitivity of 83.5%, 85.5%, and 36.0% for the three methods, respectively (Table 2).

Discussion

This pilot trial demonstrates for the first time the value of an automated ultrasound system in the detection of synovitis. It evaluates the practicability, sensitivity and inter-observer reliability of this system in comparison to manual ultrasound, physical examination, and MRI, with the latter taken as the gold standard. The patients chosen for the trial suffered from either rheumatoid or psoriatic arthritis and showed moderate disease activity.

Advantages

As previously demonstrated by several authors [14, 15], US was more sensitive for the detection of swollen joints than physical examination. Our data confirmed the high sensitivity for the detection of swollen joints, which was significantly higher for both ultrasound techniques as compared to physical examination. This finding is important, since automated US is not a fully developed tool for the detection of synovitis. Despite this fact, the number of joints detected by automated and manual US and MRI did not differ. The detection of synovitis may improve with further development of automated US.

Another advantage is that automated ultrasound allows the treating physician to go back to the original data and examine any plane more closely. With standard pictures from manual ultrasound, the physician is always restricted to the previously generated image without the opportunity of a later revision of the scan.

Sensitivity/specificity

In detail, the two ultrasound-based methods were two and half times more sensitive than physical examination. The specificity,
however, was more than two times higher for physical examination. Proliferation of synovial tissue and fluid accumulation as the major contributors to joint swelling are only detectable by physical examination if they reach a certain extent, leading to the detection of less swollen joints by physical examination and increased specificity as a consequence. This effect may have been accentuated because the patients in this study had only moderately active disease as confirmed by a higher negative predictive value for the two ultrasound-based methods (Table 2). Further development of automated US for the hand and other joints may increase its sensitivity and specificity.

Limitations and future work
Some restrictions of this method should be considered. About one tenth of the joints could not be adequately examined by automated US, mainly due to insufficient contact between the gel matrix and the scanner. It has to be acknowledged that the automated US system has been developed for the female breast and not for joints. The technical setting in this pilot trial was therefore preliminary and the number of not assessable joints may increase with further development. Grossly deformed joints, however, will presumably remain as problematic as they are for manual ultrasound. Combining automated ultrasound with manual ultrasound and the use of a hockey stick may help to lessen the effects of deformity.

Large joints and feet were not analyzed for this pilot study. Hands are commonly used to analyze disease activity in RA patients. It was shown that erosions correlated well when comparing small and large joints [32, 33]. Therefore, we focused for this first analysis on the small joints of the hands. Feet were not analyzed for hygienic and other practical reasons.

The inter-observer reliability was significantly better for physical examination than for the US-based methods. The extensive experience and training of both assessors may be one reason. Secondly, the lower number of swollen joints by physical examination, as discussed above for specificity, may cause this higher inter-observer reliability.

Our study with automated US has only shown its capacity for the detection of joint swelling. A great advantage of MRI and manual ultrasound is the ability to demonstrate inflammation by gadolinium uptake or power Doppler. Power Doppler, however, has not yet been integrated into this automated US system. The reason may be the great pressure the system builds up. A possibility to include Power Doppler would be to use a water bath or, maybe, combining the method with the fluorescence optical imaging technology, e.g., Xiralite scan. However, it should be taken into account that a Power Doppler-driven treatment strategy does not necessarily lead to an improved outcome of RA patients as shown by e.g. Dale et al. in the TaSER study [34] or by Haavardsholm in the ARCTIC study [35]. We think that, despite these two studies, a solution to integrate Power Doppler or Xiralite scan needs to be found, as Power Doppler adds important information on establishing the diagnosis [36, 37], staging and predicting disease course [38–40], monitoring and managing therapeutic response [41], [40, 41].

A cost effectiveness analysis is not possible at this point in the development of automated US for the detection of synovitis in inflammatory joint diseases.

This investigation was planned as a pilot study only. Currently, we think that automated US is not sufficiently developed to be used in the standard evaluation of arthritis patients. The automated ultrasound system now needs to be evaluated systematically in further studies. In the first place, a clear definition of joint swelling detected by ABVS is needed. Secondly, the contact problem of automated ultrasound with the matrix needs to be solved. Only if these problems are resolved can a definite analysis of metric parameters, such as sensitivity and specificity, be conducted to show the real value of this automated ultrasound system. In addition Power Doppler or Xiralite scan needs to be included in the system. Automated ultrasound may offer some opportunities as a screening and follow-up method. Focused manual US could be employed in select joints detected by ABVS and this technology could also prove useful for clinical trials where the study can be read independently by external assessors.

Conclusion
Automated US is a promising device for screening for synovitis and its documentation in patients with inflammatory joint diseases. It is simple, fast, can be conducted by non-physician staff, and yields three-dimensional joint images.

Conflict of Interest
The authors declare no conflict of interest.

References


