The OMERACT Ultrasound Group: Status of Current Activities and Research Directions

RICHARD J. WAKEFIELD, MARIA-ANTONIETTA D'AGOSTINO, ANNAMARIA IAGNOCCO, EMILIO FILIPPUCCI, MARINA BACKHAUS, ALEXANDER K. SCHEEL, FREDRICK JOSHUA, ESPERANZA NAREDO, WOLFGANG A. SCHMIDT, WALTER GRASSI, INGRID MOLLER, CARLOS PINEDA, ANDREA KLAUSER, MARCIN SZKUDLAREK, LENE TERSLEV, PETER BALINT, GEORGE A.W. BRUYN, WIJNAND A.A. SWEN, SANDRINE JOUSSE-JOULIN, DAVID KANE, JUHANI M. KOSKI, PHILIP O'CONNOR, SANJA MILUTINOVIC, and PHILIP G. CONAGHAN

ABSTRACT. Ultrasound (US) is a relatively new imaging modality in rheumatology that offers great potential as a diagnostic and management tool. In 2004, an OMERACT Ultrasound Special Interest Group was formed to address the metric qualities of US as a potential outcome measure. A preliminary systematic review highlighted the deficiencies in the literature, particularly with regard to the reliability of interpreting and acquiring images; as a consequence, a number of exercises were proposed to address these issues. This report describes a series of iterative studies that have resulted in improved intra- and inter-reader reliability for detecting and scoring synovitis from both static and real-time images of the hand joints of patients with rheumatoid arthritis. The reliability of acquiring images was also enhanced using standardized positions. Future studies will assess the value of US in clinical trials. (J Rheumatol 2007;34:848–51)

Key Indexing Terms:
ULTRASONOGRAPHY VALIDITY RELIABILITY INFLAMMATORY ARTHRITIS

Musculoskeletal ultrasound (US) is a rapidly growing imaging modality used for the investigation and management of musculoskeletal disorders[1]. It has a number of practical advantages over other imaging techniques such as magnetic resonance imaging (MRI), including low cost, good accessibility, and ability to scan multiple joints in relatively short periods of time. As well, it can simultaneously image bone and soft tissue. In rheumatology, recent interest has focused on the ability of US to detect and monitor joint-related soft tissue inflammation and its structural sequelae. Pathologies that can potentially be visualised using US include bone erosions, synovitis, tenosynovitis, and enthesopathy.

From the Academic Unit of Musculoskeletal Disease, University of Leeds, Leeds, UK; Ambroise Pare Hospital, Boulogne-Billancourt, France; Rheumatology Unit, University La Sapienza, Rome, Italy; University of Ancona, Ancona, Italy; Charite University Hospital, Humboldt University, Berlin, Germany; Department of Medicine, Nephrology and Rheumatology, Georg-August University Gottingen, Gottingen, Germany; Department of Rheumatology, St. George Hospital, University of NSW, Sydney, Australia; Department of Rheumatology, Severo Ochoa Hospital, Madrid, Spain; Medical Center for Rheumatology Berlin-Buch, Berlin, Germany; Department of Rheumatology, Hospital de Bellvitge, Barcelona, Spain; Instituto Nacional de Cardiologia Ignacio Chavez, Mexico City, Mexico; Department of Radiology II, University Hospital Innsbruck, Innsbruck, Austria; University of Copenhagen Hvidovre Hospital, Copenhagen, Denmark; The Parker Institute, Frederiksberg Hospital, Frederiksberg, Denmark; Third Rheumatology Department, National Institute of Rheumatology and Physiotherapy, Budapest, Hungary; Medical Centre Leeuwarden, Leeuwarden, The Netherlands; Rheumatology Department, CHU Brest, Brest, France; Adelaide and Meath Hospital, Dublin, Ireland; Institute of Rheumatology, Belgrade, Serbia; Department of Radiology, Chapel Allerton Hospital, Leeds, UK; and Mikkeli Central Hospital, Mikkeli, Finland.

Dr. F. Joshua is supported by an OMERACT Fellowship.

R.J. Wakefield, BM, MD, MRCP, Senior Lecturer in Rheumatology, Academic Unit of Musculoskeletal Disease, University of Leeds; M-A. D'Agostino, MD, Assistant Professor, Ambroise Pare Hospital; A. Iagnocco, MD, Consultant Rheumatologist, University La Sapienza; E. Filippucci, MD, PhD, Consultant in Rheumatology, University of Ancona; M. Backhaus, MD, PhD, Associate Professor of Rheumatology, Charite University Hospital, Humboldt University; A.K. Scheel, MD, PhD, Associate Professor of Rheumatology, Georg-August University Gottingen; F. Joshua, MB BS, FRACP, Rheumatology Research Fellow, Department of Rheumatology, St. George Hospital, University of NSW; E. Naredo, MD, Consultant Rheumatologist, Severo Ochoa Hospital; W.A. Schmidt, MD, Local Director, Medical Center for Rheumatology Berlin-Buch; W. Grassi, MD, Professor of Rheumatology, University of Ancona; I. Moller, MD, Consultant Rheumatologist, Department of Rheumatology, Hospital de Bellvitge; C. Pineda, MD, Consultant Rheumatologist, Instituto Nacional de Cardiologia Ignacio Chavez; A. Klauser, MD, Consultant Skeletal Radiologist, University Hospital Innsbruck; M. Szkudlarek, MD, PhD, Consultant Rheumatologist, University of Copenhagen Hvidovre Hospital; L. Terslev, MD, PhD, Rheumatologist, The Parker Institute, Frederiksberg Hospital; P.V. Balint, MD, PhD, Consultant Rheumatologist, Third Rheumatology Department, National Institute of Rheumatology and Physiotherapy, Budapest; G.A.W. Bruyn, MD, PhD, Consultant Rheumatologist, Medical Centre Leeuwarden; W.A.A. Swen, MD, PhD, Consultant Rheumatologist, Alkmaar, The Netherlands; S. Jousse-Joulin, MD, Rheumatology Department, CHU Brest; D. Kane, MRCPI, PhD, Consultant Rheumatologist, Adelaide and Meath Hospital; S. Milutinovic, MD, Institute of Rheumatology, Belgrade; P.J. O'Connor, MBBS, MRCP, FRCR, Consultant Skeletal Radiologist, Department of Radiology, Chapel Allerton Hospital; J.M. Koski, MD, Consultant Rheumatologist, Mikkeli Central Hospital; P.G. Conaghan, MB BS, PhD, FRACP, FRCP, Professor of Musculoskeletal Medicine, Academic Unit of Musculoskeletal Disease, University of Leeds.

Address reprint requests to Dr. R.J. Wakefield, Academic Unit of Musculoskeletal Disease, Chapel Allerton Hospital, Chapeltown Road, Leeds LS7 4SA, United Kingdom. E-mail: medrjw@leeds.ac.uk

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2007. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org

State of the field — US in rheumatoid arthritis (RA)

The first report describing the use of US in RA was by Cooperberg, *et al*[2] in 1978 for the assessment of synovitis in the knee. The first description of the use of US in the hand in RA was a decade later by De Flaviis, *et al*[3]. This report highlighted the spectrum of pathology that could be visualized including synovitis, tenosynovitis, and erosions. It was not until the 1990s, however, that detailed assessment of inflammatory joint diseases became a reality, largely through the development of smaller high-frequency transducers that were better suited for superficial structures such as the small joints.

Over this relatively short period there have been an increasing number of reports demonstrating a greater sensitivity of US over clinical examination and radiography for the detection of synovitis, erosions, and tendon disease, particularly in the hand[4,5]. More recently, power Doppler has started to replace gray-scale US as an indicator of inflammatory joint disease[6,7].

Clinical problem

Despite increasing interest in US, widespread application has been impeded by a perception that its use is unproven and unreliable. This view has been compounded by a paucity of data regarding its metric properties, making interpretation and comparison of studies difficult. In particular, there are limited data describing standardized scanning methodology and standardized definitions of US pathologies in addition to how to quantify these abnormalities. As a result, the EULAR Working Party on Ultrasound recently acknowledged that the application of the OMERACT filter — truth, discrimination, and feasibility — would further advance the application of US. The first OMERACT/EULAR group was thus formed at the OMERACT 7 meeting in 2004.

The aim of the OMERACT 7 meeting was to bring together international groups that were expert in musculoskeletal US, in the form of a Special Interest Group (SIG). Member groups presented the results of systematic literature reviews performed prior to the meeting for a range of US pathologies at specific anatomical sites (shoulder, hand, wrist, elbow, knee, forefoot, ankle, and hindfoot). Sizable gaps in the literature on the metric properties of US were revealed; this research helped inform the future research agendas. Highlighted deficiencies in the literature included a paucity of both reliability and validity data. Interobserver reliability was the most studied area, with relatively few data relating to intraobserver or to intermachine reliability. It was noted that data on normal joint structures were scarce.

Importantly, the US SIG at OMERACT 7 enabled consensus to be reached for preliminary US definitions for common pathologies seen in the inflammatory joint diseases, including synovitis, erosions, tenosynovitis, and enthesopathy[8]. The conclusion from the OMERACT 7 meeting was that a number of exercises would be undertaken to test and further develop the pathological definitions and focus on reliability issues. The results of these would be reported back to the OMERACT 8 meeting.

At OMERACT 8. The US SIG meeting enabled presentations of an iterative series of studies undertaken within the EULAR/OMERACT network at different centers in order to address the reliability issues above and to test the preliminary pathological definitions. A summary of these studies is given below.

Study 1: Interobserver reliability between 14 experts (presented by Dr. M. Backhaus). The first study by the group was conducted in Berlin in June 2004 during a "Train the Trainers" course. Its aims were 2-fold — first, to assess the interobserver reliability among 14 rheumatology experts in musculoskeletal US, and second, to compare this consensus agreement with MRI, which served as the imaging gold standard. Four joint regions (shoulder, knee, hand, and foot) from 4 patients with different inflammatory joint diseases were scanned independently by all the experts using the same machine. A range of pathologies was documented including synovitis, erosions, bursitis, and tendon disease (tenosynovitis and tendinopathy). Taking a majority agreement in US examination of 10 out of 14 experts into account, a modified kappa index was calculated for interobserver agreement between different joint regions and pathologies. The results of this study have been published[9]. Calculations for each joint region and pathology were presented. Considering all pathologies, the modified kappa showed high values for the knee (1.0), moderate values for the shoulder (0.76) and hand/finger (0.59), but low agreement for ankle/toe joints (0.28). Relatively good agreement was found for most US findings compared with MRI (Table 1) for the shoulder (sensitivity 76%, specificity 89%, overall agreement 81%) and knee joint (sensitivity 91%, specificity 88%, overall agreement 88%). Sensitivities were lower for wrist/finger (sensitivity 66%, specificity 88%, overall agreement 73%) and ankle/toe joints (sensitivity 61%, specificity 92%, overall agreement 82%). Interobserver reliabilities, sensitivities, and specificities in comparison with MRI were moderate to good. This study, which did not use the OMERACT pathology definitions, confirmed our previous observation that further standardization of US scanning techniques was necessary even among experts.

Study 2: Interobserver reliability between 23 experts (presented by Dr. E. Naredo). The second study was undertaken at a further "Teach the Teachers" course in Sitges, near Barcelona,

Table 1. Overall levels of agreement between US and MRI for different joints.

Joint	Sensitivity	Specificity	Overall Agreement
Shoulder	0.76	0.89	0.81
Knee	0.91	0.88	0.88
Wrist and finger	0.66	0.88	0.73
Ankle and toe	0.61	0.92	0.82

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2007. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org

Spain in September 2004. This exercise involved 23 expert sonographers blinded to the clincal details from independent performance of US of one of 4 anatomical regions (shoulder, hand/wrist, knee, foot/ankle) in 24 patients with mixed inflammatory and noninflammatory disorders. The results[10] revealed that while reliability was generally good, for example, kappa = 0.73 for detection of synovitis in the wrist, again there was variability related to individual joints and the type of pathology examined (kappa = 0.50 for shoulder tendon lesions). It particular, during our subsequent discussions, it was apparent that while the OMERACT definitions appeared to work well, differences occurred with regard to grading of pathologies and the technique of acquiring images; for example, some operators used dynamic image acquisition more than others, and some may have imaged joints from only one position. The difficulty of differentiating normal physiological fluid from pathological fluid was also highlighted — a normal scan to one sonographer may represent low grade pathology to another.

Due to the complexity of analyzing differences between multiple joints and pathologies it was decided that the subsequent exercises would be more focused on a specific disease, joint region, and joint pathology.

Study 3: Intra- and interobserver reliability in RA (presented by Dr. M-A. D'Agostino). The final study consisted of several substudies. These were conducted in Paris between December 2004 and December 2005[11]. For these studies it had been decided to focus on one disease (RA), one joint group (small joints), and one pathology (synovitis), as this was considered the most validated area in inflammatory disease. The objective therefore was to evaluate the intra- and interobserver reproducibility of several observers for detecting and scoring synovitis in the small joints of patients with RA. The first exercise was divided into 2 parts. In the first, 17 experts were asked to evaluate a sequence of 86 US images of MCP, PIP, MTP, and wrist joints for the presence or absence of gray-scale and power Doppler synovitis. The images were also scored using a 0–3 semiquantitative scale for each modality. In order to assess intrareader reliability, 24 US images were randomly presented twice. In the second part of the study, the acquisition of images was tested by each expert scanning 32 MCP (2nd to 5th) and 32 PIP (2nd to 5th) joints of the same 8 patients with RA. Each patient was scanned using the same type of machine. Sixteen joints were scanned twice on the same day using the same US equipment. All the joints were examined on both dorsal and palmar aspects. The kappa values for intra- and interobserver reliability are given in Table 2. No difference in the reliability was observed using dorsal or palmar scans.

It was concluded that the intra- and interreader reliability of interpreting static images for detecting and scoring gray-scale and power Doppler synovitis in patients with RA was very good. It should be noted that the kappa value of 1 for binary scoring of power Doppler arose because the scorer was only expected to comment on the presence or absence of color signal. The semiquantitative scores, however, yielded a lower value. As well, when both image acquisition and interpretation were tested simultaneously, the observed intra- and interreader reliability was both variable and too low for regular use across centers in trials. The reasons for this were felt to be multifactorial, including a lack of familiarity with the US equipment for some experts and differences in scanning techniques.

The second study (manuscript in preparation) involved developing a consensus among 14 experts on semiquantitative grading criteria (0–3 scale) for both gray-scale and power Doppler synovitis in patients with RA. After deciding on agreed criteria, an exercise was conducted on patients with RA in order to retest the intra- and interobserver reliability of the detection of synovitis in the MCP joints. Preliminary data from this ongoing study appear promising, and suggest that use of the rules for scoring synovitis can be achieved. A further exercise comparing the use of a freehand technique versus a single-position technique for acquiring images found that using a single position improved the reliability of acquisition, although it would be expected that the sensitivity for detecting pathology would be lower.

Conclusions and future research agenda

Ultrasound is an exciting new imaging modality that offers a number of potential benefits over other techniques such as radiography and MRI. It remains largely untested, however, in the clinical trials setting, with a paucity of data relating to its metric properties. The work reported here presents the results of a number of iterative studies that have culminated in improved intra- and interobserver reliability for detecting and

Table 2. Intra-and interobservers' kappa values (ranges and means) for interpreting and acquiring US for small joints of the hand.

Imaging Modality	Interpretation of Static Images		Acquisition and Interpretation Using Patients	
	Intraobserver (range of k)	Interobserver (mean kappa)	Intraobserver (range of k)	interobserver (mean kappa)
Gray-scale (Yes/No)	0.5–1.0	0.71	0.2–0.94	0.31
Power Doppler (Yes/No)	1.0	0.98	0.1–0.83	0.42
Gray-scale (grade 0–3)	0.68–0.91	0.76	0.2–0.91	0.43
Power Doppler (grade 0–3)	0.89–1.0	0.94	0.2–0.89	0.42

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2007. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org

scoring static images for synovitis in hand joints of patients with RA, as well as improving real-time acquisition reliability. Future studies are planned to refine and validate these definitions before testing intermachine reliability, which will be particularly important in the context of multicenter studies. The stage will then be set to evaluate the role of US in longitudinal clinical trials, an area where data are only starting to emerge[12].

REFERENCES

1. Grassi W, Salaffi F, Filippucci E. Ultrasound in rheumatology. Best Pract Res Clin Rheumatol 2005;19:467-85.
2. Cooperberg PL, Tsang I, Truelove L, Knickerbocker WJ. Gray scale ultrasound in the evaluation of rheumatoid arthritis of the knee. Radiology 1978;126:759-63.
3. De Flaviis L, Scaglione P, Nessi R, Ventura R, Calori G. Ultrasonography of the hand in rheumatoid arthritis. Acta Radiol 1988;29:457-60.
4. Kane D, Grassi W, Sturrock R, Balint PV. Musculoskeletal ultrasound — a state of the art review in rheumatology. Part 2: Clinical indications for musculoskeletal ultrasound in rheumatology. Rheumatology Oxford 2004;43:829-38.
5. Ostergaard M, Szkudlarek M. Ultrasonography: a valid method for assessing rheumatoid arthritis? Arthritis Rheum 2005;52:681-6.
6. Wakefield RJ, Brown AK, Emery P. Power Doppler sonography: improving disease activity assessment in inflammatory joint disease. Arthritis Rheum 2003;48:285-8.
7. Schmidt WA. Doppler sonography in rheumatology. Best Pract Res Clin Rheumatol 2004;18:827-46.
8. Wakefield RJ, Balint PV, Szkudlarek M, et al. Musculoskeletal ultrasound including definitions for ultrasonographic pathology. J Rheumatol 2005;32:2485-7.
9. Scheel AK, Schmidt WA, Hermann KG, et al. Interobserver reliability of rheumatologists performing musculoskeletal ultrasonography: results from a EULAR "Train the trainers" course. Ann Rheum Dis 2005;64:1043-9.
10. Naredo E, Moller I, Moragues C, et al. Interobserver reliability in musculoskeletal ultrasonography: results from a "Teach the Teachers" rheumatologist course. Ann Rheum Dis 2006;65:14-9.
11. D'Agostino MA, Wakefield RJ, Filippucci E, et al. Intra- and inter-observer reliability of ultrasonography for detecting and scoring synovitis in rheumatoid arthritis: a report of a EULAR ESCISIT Task Force. Ann Rheum Dis 2005;64 Suppl III:62.
12. Taylor PC, Steuer A, Gruber J, et al. Comparison of ultrasonographic assessment of synovitis and joint vascularity with radiographic evaluation in a randomised controlled study of infliximab therapy in early rheumatoid arthritis. Arthritis Rheum 2004;50:1107-16.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2007. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org