

Responsiveness in Rheumatoid Arthritis. A Report from the OMERACT 11 Ultrasound Workshop

Annamaria Iagnocco, Esperanza Naredo, Richard Wakefield, George A.W. Bruyn, Paz Collado, Sandrine Jousse-Joulin, Stephanie Finzel, Sarah Ohrndorf, Andrea Delle Sedie, Marina Backhaus, Hilde Berner-Hammer, Frederique Gandjbakhch, Gurjit Kaeley, Damien Loeuille, Ingrid Moller, Lene Terslev, Philippe Aegerter, Sibel Aydin, Peter V. Balint, Emilio Filippucci, Peter Mandl, Carlos Pineda, Johannes Roth, Silvia Magni-Manzoni, Niolay Tzaribachev, Wolfgang A. Schmidt, Philip G. Conaghan, and Maria-Antonietta D'Agostino

ABSTRACT. Objective. To summarize the work performed by the Outcome Measures in Rheumatology (OMERACT) Ultrasound (US) Task Force on the validity of different US measures in rheumatoid arthritis (RA) and juvenile idiopathic arthritis (JIA) presented during the OMERACT 11 Workshop.

Methods. The Task Force is an international group aiming to iteratively improve the role of US in arthritis clinical trials. Recently a major focus of the group has been the assessment of responsiveness of a person-level US synovitis score in RA: the US Global Synovitis Score (US-GLOSS) combines synovial hypertrophy and power Doppler signal in a composite score detected at joint level. Work has also commenced examining assessment of tenosynovitis in RA and the role of US in JIA.

Results. The US-GLOSS was tested in a large RA cohort treated with biologic therapy. It showed early signs of improvement in synovitis starting at Day 7 and increasing to Month 6, and demonstrated sensitivity to change of the proposed grading. Subsequent voting questions concerning the application of the US-GLOSS were endorsed by > 80% of OMERACT delegates. A standardized US scoring system for detecting and grading severity of RA tenosynovitis and tendon damage has been developed, and acceptable reliability data were presented from a series of exercises. A preliminary consensus definition of US synovitis in pediatric arthritis has been developed and requires further testing.

Conclusion. At OMERACT 11, consensus was achieved on the application of the US-GLOSS for evaluating synovitis in RA; and work continues on development of RA tenosynovitis scales as well as in JIA synovitis. (First Release Nov 15 2013; J Rheumatol 2014;41:379–82; doi:10.3899/jrheum.131084)

Key Indexing Terms:

ULTRASOUND RESPONSIVENESS RHEUMATOID ARTHRITIS OMERACT

From the Rheumatology Unit, Sapienza Università di Roma, Rome, Italy; Rheumatology Department, Hospital General Universitario Gregorio Marañón and Universidad Complutense, Madrid, Spain; University of Leeds and UK National Institute for Health Research (NIHR) Leeds Musculoskeletal Biomedical Research Unit, Leeds, UK; Department of Rheumatology, MC Groep Hospitals, Lelystad, The Netherlands; Hospital Universitario Severo Ochoa, Madrid, Spain; Rheumatology Unit, Cavale Blanche Hospital, Brest, France; University of Erlangen-Nuremberg, Department of Internal Medicine 3, Erlangen; Department of Rheumatology and Clinical Immunology, Charité-Universitätsmedizin Berlin, Charitéplatz, Berlin, Germany; Rheumatology Unit, Azienda Ospedaliero Universitaria Pisana, Pisa, Italy; Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway; Université Pierre et Marie Curie, La Pitié Salpêtrière Hospital, APHP, Paris, France; University of Florida College of Medicine, Jacksonville, Florida, USA; Division of Rheumatology, CHU Pitié-Salpêtrière, Paris, France; Instituto Pòal, Barcelona, Spain; Department of Rheumatology, Copenhagen University Hospital at Glostrup, Copenhagen, Denmark; Department of Epidemiology and Biostatistics, Université Paris Ouest-Versailles-Saint Quentin en Yvelines, Hôpital Ambroise Paré, APHP, Boulogne-Billancourt, France; Istanbul Medeniyet University, Goztepe Training and Research Hospital, Istanbul, Turkey; National Institute of Rheumatology and Physiotherapy, Budapest, Hungary; Clinica Reumatologica, Università Politecnica delle Marche, Jesi (Ancona), Italy; Division of

Rheumatology, Medical University of Vienna, Vienna, Austria; Biomedical Research, Instituto Nacional de Rehabilitación, Mexico City, Mexico; University of Ottawa, Children's Hospital of Eastern Ontario, Ottawa, Canada; Division of Rheumatology IRCCS Ospedale Pediatrico Bambino Gesù, Rome, Italy; Department of Pediatric Rheumatology, Center for Rheumatic Diseases, Bad Bramstedt; Immanuel Krankenhaus — Medical Center for Rheumatology Berlin, Berlin, Germany; Department of Rheumatology, Université Versailles-Saint Quentin en Yvelines, Ambroise Paré Hospital, APHP, Boulogne-Billancourt, France.

A. Iagnocco, MD, Rheumatology Unit, Sapienza Università di Roma; E. Naredo, MD, Rheumatology Department, Hospital General Universitario Gregorio Marañón and Universidad Complutense; R. Wakefield, MD; P.G. Conaghan, MB BS, PhD, FRACP, FRCP, University of Leeds and NIHR Leeds Musculoskeletal Biomedical Research Unit; G.A.W. Bruyn, MD, PhD, Department of Rheumatology, MC Groep Hospitals; P. Collado, MD, Hospital Universitario Severo Ochoa; S. Jousse-Joulin, MD, Rheumatology Unit, Cavale Blanche Hospital; S. Finzel, MD, Department of Internal Medicine, University of Erlangen-Nuremberg; S. Ohrndorf, MD; Marina Backhaus, MD, Department of Rheumatology and Clinical Immunology, Charité-Universitätsmedizin Berlin; A. Delle Sedie, MD, Rheumatology Unit, Azienda Ospedaliero Universitaria Pisana; H. Berner-Hammer, MD, Department of Rheumatology, Diakonhjemmet Hospital; F. Gandjbakhch, MD, Université Pierre et Marie Curie, La Pitié

Salpêtrière Hospital, APHP; G. Kaeley, MD, University of Florida College of Medicine; D. Loeuille, MD, PhD, Division of Rheumatology, CHU Pitié-Salpêtrière; I. Moller, MD, Instituto Poal; L. Terslev, MD, PhD, Department of Rheumatology, Copenhagen University Hospital at Glostrup; P. Aegerter, MD, PhD, Department of Epidemiology and Biostatistics, Université Paris Ouest-Versailles-Saint Quentin en Yvelines, Hôpital Ambroise Paré, APHP; S. Aydin, MD, Istanbul Medeniyet University, Goztepe Training and Research Hospital; P.V. Balint, MD, PhD, National Institute of Rheumatology and Physiotherapy; E. Filippucci, MD, PhD, Clinica Reumatologica, Università Politecnica delle Marche; P. Mandl, MD, PhD, Division of Rheumatology, Medical University of Vienna; C. Pineda, MD, Biomedical Research, Instituto Nacional de Rehabilitación; J. Roth, MD, University of Ottawa, Children's Hospital of Eastern Ontario; S. Magni-Manzoni, MD, Division of Rheumatology IRCCS Ospedale Pediatrico Bambino Gesù; N. Tzaribachev, MD, Department of Pediatric Rheumatology, Center for Rheumatic Diseases; W.A. Schmidt, MD, Immanuel Krankenhaus, Medical Center for Rheumatology Berlin; M-A. D'Agostino, MD, PhD, Department of Rheumatology, Université Versailles-Saint Quentin en Yvelines, Ambroise Paré Hospital, APHP.

Address correspondence to Prof. D'Agostino, Versailles Saint Quentin en Yvelines University, Rheumatology Department, Ambroise Paré Hospital, APHP, 9 avenue Charles de Gaulle, 92100 Boulogne-Billancourt, France; E-mail: maria-antonieta.dagostino@apr.aphp.fr

The OMERACT Ultrasound Task Force met for the first time at OMERACT 7 in 2004 with the aim of investigating the role of ultrasound (US) in rheumatology, and developing its metric properties, particularly in the field of inflammatory arthritis. The first step was to develop and publish standardized definitions of US pathologies¹. The group then performed a systematic review of the metric properties of US for the detection of synovitis at the individual joint level (for multiple peripheral joints) that highlighted gaps in the literature, including a lack of reliability data with respect to intraoccasion, intrareader, and interreader reliability². Following this work, through a series of iterative intra-observer and interobserver reliability exercises, the group developed and tested definitions and scoring of joint pathologies related either to inflammation or to structural damage in rheumatoid arthritis (RA). This subsequently resulted in the ongoing development of a US GLObal Synovitis Score (US-GLOSS), at the whole patient level, that would be proposed for application in multicenter therapeutic trials, providing a responsive US scoring system in patients with RA³.

Following the decision taken during the OMERACT 10 meeting, the group expanded the validation of US in RA by working on tenosynovitis. Thus, starting from a systematic literature review of US definitions, scoring systems, and validity for tendon involvement in RA, a series of reliability exercises was performed that focused on the detection and scoring of tenosynovitis and tendon damage in patients with RA^{4,5,6}. It was also agreed at OMERACT 10 to work on US as an outcome measure in pediatric arthritis. Based on a systematic literature review, the group started work on the definition of a normal joint that will lead to definitions of pediatric joint synovitis. At the same time future areas of standardization were agreed on⁷.

The aim of the present article is to summarize the recent

work performed by the OMERACT US Group in the assessment of responsiveness for a number of measures in RA, as presented during the OMERACT 11 Workshop.

US-GLOSS

The concept of developing a US-based scoring system for synovitis in RA at the patient level was agreed on through group discussions and feedback sessions at OMERACT 10 in 2010⁸. This decision was based on the need for integrating the components of synovitis (i.e., hypoechoic synovial hypertrophy and synovial Doppler signal) in a unique global score of joint inflammatory activity in RA. This scoring system, applied at the patient level by means of a multijoint US assessment, should allow an objective followup of patients under treatment and should provide a feasible and objective instrument for evaluating disease activity in patients with RA⁸. The work developed by the group in the assessment of reliability in the detection and scoring of RA synovitis demonstrated positive results at the joint level^{8,9,10,11,12,13}. In addition, recent longitudinal studies applying either extensive or reduced US joint counts have shown the feasibility of US for following patients under treatment. However, many differences have been reported in the scoring system used (i.e., greyscale synovitis, Doppler, both modalities) as well as in the number of joints assessed and in the correlations (construct validity) between US scores and clinical and laboratory findings^{8,14,15,16,17,18,19}. Thus, development of a US-GLOSS required testing of sensitivity to change of the OMERACT synovitis scoring system in a multicenter study.

During the OMERACT 11 Workshop, results were presented of the application of the US-GLOSS in patients with RA. The study represented the first international, multicenter, prospective, open-label therapeutic trial that used the OMERACT US-GLOSS and tested the response to treatment with abatacept and methotrexate (MTX) in patients with RA and inadequate response to MTX (manuscript in preparation). In this study, US responsiveness was used as the primary objective. US-GLOSS, by combining synovial hypertrophy and power Doppler (PD) signal in a composite score applied at joint level, was able to demonstrate, at patient level, early signs of improvement in synovitis from days 7, 15, and 29, increasing up to Month 6³. A reduced subset of joints (GLOSS) that best represented the global PDUS score for paired 22 joints of all patients over 3 timepoints was identified at baseline (BL) and days 85 and 169, based on principal component analysis. The sensitivity to change was assessed for GLOSS and the existing 12-joint and 7-joint sets using standardized response mean (SRM). When the GLOSS was compared with the 2 previously published reduced joint sets (assessed bilaterally), the mean change from BL improved up to Day 169 for all 3 measures, with similar SRM.

Based on these results, at the plenary session 4 voting

questions for US synovitis were presented to participants for potential endorsement. The first voting question: “Do you support continuing to evaluate US in assessing synovitis by using greyscale and Doppler modalities separately?” was endorsed by > 80% of the OMERACT delegates to have appropriate validation data. The second question: “In addition, do you support to continue development and subsequent validation of a combined score?” received 70% of the vote. The third question: “Do you support the idea to improve the feasibility of US assessment by developing and validating a reduced joint count?” was endorsed by > 80% of delegates. Finally, the fourth question: “Do you support implementation of US as a secondary/exploratory outcome in therapeutic clinical trials?” received 80% of the vote.

Tenosynovitis

The work on tenosynovitis was essentially focused on the assessment of discrimination of US for detecting and scoring tendon involvement in RA. The starting point was a systematic literature review of US definitions, scoring systems, and metric properties for tenosynovitis and tendon damage in RA⁴. Then, based on the limited available data on reliability, the group started a series of exercises focused initially on testing the reliability of US in detecting lesions, and subsequently aimed at assessing the reliability of US in scoring tenosynovitis and tendon damage^{5,6}. A Delphi process was previously performed on defining the scoring systems for tenosynovitis⁶. During the US Workshop at OMERACT 11, results of the literature review were presented with focus on the different aspects of validity⁴. The results of the intraobserver and interobserver reliability exercises were also reported: moderate-to-good intra-observer and interobserver reliability for tenosynovitis and tendon damage were demonstrated for both greyscale and PD. Then, following the above Delphi consensus process (which examined agreement of a group of experts on US-defined tenosynovitis and US scoring systems of tenosynovitis in RA), a second patient-based reliability exercise was performed. The results showed good intra-observer reliability and moderate to good interobserver reliability for B-mode and PD modalities⁶. A formal research agenda was also developed including a third patient-based reliability exercise on tendon damage in RA, performed in June 2012 in Amsterdam, with the aim of developing an OMERACT atlas on scoring tenosynovitis and tendon damage in RA.

Pediatric Arthritis

Several publications have highlighted the capability of US to detect joint involvement in pediatric arthritis revealing the high frequency of asymptomatic findings. At OMERACT 11 the group reported results of a systematic literature review demonstrating that validity of US in pediatric arthritis had not yet been established and

standardized⁷. Using results of an international survey on the use of US in pediatric rheumatology as the basis for future work, the first identified area for standardization was the definition and detection of US synovitis in juvenile idiopathic arthritis. Such pathology does not differ from that observed in adults; however, the appearance of the joint can vary according to the age and maturation of the child, with the presence of variable amounts of vascularized epiphyseal cartilage. This makes the distinction of normal findings from synovitis more challenging. The work of the group has also been focused on developing definitions of normal joint components for different age groups through a Delphi consensus process and by testing them in a reliability exercise involving healthy children, performed in Madrid in March, 2012. The preliminary results were presented during the session. The group is now developing preliminary definitions for common US-detected pathologies in pediatric rheumatology diseases.

At OMERACT 11 all aspects presented on synovitis were strongly endorsed by OMERACT delegates, further supporting US as a good outcome measure for evaluating activity and responsiveness in RA. Consensus was achieved on the application of a US-GLOSS scoring system for synovitis in RA and its use in therapeutic clinical trials. Further randomized controlled studies are needed to confirm these results and explore predictive validity in terms of structural damage and stratification of “patients at risk” for severe disease.

REFERENCES

1. Wakefield RJ, Balint PV, Szkudlarek M, Filippucci E, Backhaus M, D'Agostino MA, et al. Musculoskeletal ultrasound including definitions for ultrasonographic pathology. *J Rheumatol* 2005;32:2485-7.
2. Joshua F, Lassere M, Bruyn G, Szkudlarek M, Naredo E, Schmida WA, et al. Summary findings of a systematic review of the ultrasound assessment of synovitis. *J Rheumatol* 2007;34:839-47.
3. D'Agostino MA, Wakefield R, Berner Hammer H, Vittecoq O, Galeazzi M, Balint P, et al. Early response to abatacept plus MTX in MTX-IR RA patients using power Doppler ultrasonography: an open-label study [abstract]. *Ann Rheum Dis* 2012;71 Suppl 3:186.
4. Alcalde M, D'Agostino MA, Bruyn G, Moller I, Iagnocco A, Wakefield R, et al. A systematic literature review of US definitions, scoring system and validity according to the OMERACT filter for tendon lesion in RA and other inflammatory joint diseases. *Rheumatology* 2012;51:1246-60.
5. Bruyn G, Moller I, Garrido J, Bong D, D'Agostino MA, Iagnocco A, et al. Reliability testing of tendon disease using two different scanning methods in patients with rheumatoid arthritis. *Rheumatology* 2012;51:1655-61.
6. Naredo E, D'Agostino MA, Wakefield R, Möller I, Balint PV, Filippucci E, et al. Reliability of a consensus-based ultrasound score for tenosynovitis in rheumatoid arthritis. *Ann Rheum Dis* 2013;72:1328-34.
7. Collado P, Jousse-Joulin S, Alcalde M, Naredo E, D'Agostino MA. Is ultrasound a validated imaging tool for the diagnosis and management of synovitis in juvenile idiopathic arthritis? A systematic literature review. *Arthritis Care Res* 2012;64:1011-9.
8. Naredo E, Wakefield R, Iagnocco A, Terslev L, Filippucci E,

- Gandjbakhch F, et al. The OMERACT Ultrasound Task Force – status and perspectives. *J Rheumatol* 2011;38:2063-7.
9. D'Agostino MA, Conaghan PG, Naredo E, Aegerter P, Iagnocco A, Freeston JE, et al. The OMERACT ultrasound task force – Advances and priorities. *J Rheumatol* 2009;36:1829-32.
 10. Wakefield R, D'Agostino MA, Backhaus M, Balint PV, Bruyn GAW, Filippucci E, et al. Combined evaluation of influence of sonographer and machine type on the reliability of power Doppler ultrasonography (PDUS) for detecting, scoring and scanning synovitis in rheumatoid arthritis (RA) patients: results of an intermachine reliability exercise of the EULAR/OMERACT US group [abstract]. *Ann Rheum Dis* 2008;67 Suppl II:421.
 11. D'Agostino MA, Wakefield R, Backhaus M, Balint PV, Bruyn GAW, Filippucci E, et al. Evaluation of the intra- and inter-reader reliability of power Doppler ultrasonography (PDUS) for detecting and scoring synovitis of several joints in rheumatoid arthritis (RA) by using a validated scoring system: a report of the EULAR/OMERACT US group [abstract]. *Ann Rheum Dis* 2008;67 Suppl II:422.
 12. Bruyn GA, Naredo E, Moller I, Moragues C, Garrido J, de Bock GH, et al. Reliability of ultrasonography in detecting shoulder disease in patients with rheumatoid arthritis. *Ann Rheum Dis* 2009;68:357-61.
 13. Bruyn GA, Pineda C, Diaz CH, Ventura-Rios L, Moya C, Garrido J, et al. Validity of ultrasonography and measures of adult shoulder function and reliability of ultrasonography in detecting shoulder synovitis in patients with rheumatoid arthritis using MRI as a gold standard. *Arthritis Care Res* 2010;62:1079-86.
 14. Mandl P, Naredo E, Wakefield RJ, Conaghan PG, D'Agostino MA. A systematic literature review analysis of ultrasound joint count and scoring systems to assess synovitis in rheumatoid arthritis according to the OMERACT filter. *J Rheumatol* 2011;38:2055-62.
 15. Naredo E, Möller I, Cruz A, Carmona L, Garrido J. Power Doppler ultrasonographic monitoring of response to anti-tumor necrosis factor therapy in patients with rheumatoid arthritis. *Arthritis Rheum* 2008;58:2248-56.
 16. Naredo E, Rodríguez M, Campos C, Rodríguez-Heredia JM, Medina JA, Giner E, et al. Validity, reproducibility, and responsiveness of a twelve-joint simplified power Doppler ultrasonographic assessment of joint inflammation in rheumatoid arthritis. *Arthritis Rheum* 2008;59:515-22.
 17. Backhaus M, Ohrndorf S, Kellner H, Strunk J, Backhaus TM, Hartung W, et al. Evaluation of a novel 7-joint ultrasound score in daily rheumatologic practice: A pilot project. *Arthritis Rheum* 2009;61:1194-201.
 18. Iagnocco A, Filippucci E, Perella C, Ceccarelli F, Cassarà E, Alessandri C, et al. Clinical and ultrasonographic monitoring of the response to adalimumab treatment in rheumatoid arthritis. *J Rheumatol* 2008;35:35-40.
 19. Perricone C, Ceccarelli F, Modesti M, Vavala C, Di Franco M, Valesini G, et al. The 6-joint ultrasonographic assessment: a valid, sensitive-to-change and feasible method for evaluating joint inflammation in RA. *Rheumatology* 2012;51:866-73.