Ultrasound as an Outcome Measure in Gout. A Validation Process by the OMERACT Ultrasound Working Group

Lene Terslev, Marwin Gutierrez, Wolfgang A. Schmidt, Helen I. Keen, Emilio Filippucci, David Kane, Ralf Thiele, Gurjit Kaeley, Peter Balint, Peter Mandl, Andrea Delle Sedie, Hilde Berner Hammer, Robin Christensen, Ingrid Möller, Carlos Pineda, Eugene Kissin, George A. Bruyn, Annamaria Iagnocco, Esperanza Naredo and Maria Antonietta D'Agostino, on behalf of OMERACT Ultrasound Working Group

ABSTRACT. ***Objective.*** To summarize the work performed by the Outcome Measures in Rheumatology (OMERACT) Ultrasound (US) Working Group on the validation of US as a potential outcome measure in gout.

Methods. Based on the lack of definitions, highlighted in a recent literature review on US as an outcome tool in gout, a series of iterative exercises were carried out to obtain consensus-based definitions on US elementary components in gout using a Delphi exercise and subsequently testing these definitions in static images and in patients with proven gout. Cohen's κ was used to test agreement, and values of 0–0.20 were considered poor, 0.20–0.40 fair, 0.40–0.60 moderate, 0.60–0.80 good, and 0.80–1 excellent.

Results. With an agreement of > 80%, consensus-based definitions were obtained for the 4 elementary lesions highlighted in the literature review: tophi, aggregates, erosions, and double contour (DC). In static images interobserver reliability ranged from moderate to almost perfect, and similar results were found for the intrareader reliability. In patients the intraobserver agreement was good for all lesions except DC (moderate). The interobserver agreement was poor for aggregates and DC but moderate for the other components.

Conclusion. These first steps in evaluating the validity of US as an outcome measure for gout show that the reliability of the definitions ranged from moderate to excellent in static images and somewhat lower in patients, indicating that a standardized scanning technique may be needed, before testing the responsiveness of those definitions in a composite US score. (First Release September 1 2015; J Rheumatol 2015;42:2177–81; doi:10.3899/jrheum.141294)

Key Indexing Terms:
ULTRASOUND GOUT REPRODUCIBILITY OF RESULTS

From the Centre for Rheumatology and Spine Diseases, Rigshospitalet-Glostrup, Copenhagen, Denmark; Clinica Reumatologica, Università Politecnica delle Marche, Jesi, Italy; Medical Centre for Rheumatology, Immanuel Krankenhaus, Berlin, Germany; Department of Rheumatology, Royal Perth Hospital, Perth, Australia; Department of Rheumatology, Trinity College, Dublin, Ireland; Department of Medicine, Allergy/Immunology and Rheumatology Division, University of Rochester, Rochester, NY, USA; College of Medicine, University of Florida, Jacksonville, FL, USA; National Institute of Rheumatology and Physiotherapy, Budapest, Hungary; Medical University of Vienna, Vienna, Austria; U.O. Reumatologia, Azienda Ospedaliero-Universitaria Pisana (AOUP), Pisa, Italy; Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway; The Parker Institute, Department of Rheumatology, Bispebjerg and Frederiksberg Hospital, The Capital Region of Copenhagen, Denmark; Instituto Poal de Reumatologia, Barcelona, Spain; Instituto Nacional de Rehabilitacion, Mexico City, Mexico; Boston University Medical Center, Boston, MA, USA; Department of Rheumatology, MC Groep Hospitals, Lelystad, the Netherlands; Rheumatology Unit, Sapienza Università di Roma, Rome, Italy; Hospital GU Gregorio Marañón and Universidad Complutense, Madrid, Spain; APHP, Hôpital Ambroise Paré, Rheumatology Department, 92100 Boulogne-Billancourt; INSERM U1173, Laboratoire d'Excellence INFLAMEX, UFR Simone Veil, Versailles-Saint-Quentin University, Saint-Quentin en Yvelines, France.

Dr. Christensen is supported and acknowledges unrestricted grants from the Oak Foundation. Dr. Terslev is supported by the Danish Rheumatism Association.

L. Terslev, MD, PhD, Centre for Rheumatology and Spine Diseases, Rigshospitalet-Glostrup; M. Gutierrez, MD, Clinica Reumatologica, Università Politecnica delle Marche; W.A. Schmidt, MD, PhD, Medical Centre for Rheumatology, Immanuel Krankenhaus; H.I. Keen, MD, Department of Rheumatology, Royal Perth Hospital; E. Filippucci, Clinica Reumatologica, Università Politecnica delle Marche; D. Kane, MD, PhD, Department of Rheumatology, Trinity College; R. Thiele, MD, PhD, Department of Medicine, Allergy/Immunology and Rheumatology Division, University of Rochester; G. Kaeley, MD, College of Medicine, University of Florida; P. Balint, MD, PhD, National Institute of Rheumatology and Physiotherapy; P. Mandl, MD, PhD, Medical University of Vienna; A. Delle Sedie, MD, U.O. Reumatologia, AOUP; H.B. Hammer, MD, PhD, Department of Rheumatology, Diakonhjemmet Hospital; R. Christensen, PhD, The Parker Institute, Department of Rheumatology, Bispebjerg and Frederiksberg Hospital; E. Naredo, MD, Hospital GU Gregorio Marañón and Universidad Complutense; I. Möller, MD, Instituto Poal de Reumatologia; C. Pineda, MD, Instituto Nacional de Rehabilitacion; E. Kissin, MD, Boston University Medical Center; G.A. Bruyn, MD, PhD, Department of Rheumatology, MC Groep Hospitals; A. Iagnocco, MD, Rheumatology Unit, Sapienza Università di Roma; M-A. D'Agostino, MD, PhD, APHP, Hôpital Ambroise Paré, Rheumatology Department; INSERM U1173, Laboratoire d'Excellence INFLAMEX, UFR Simone Veil, Versailles-Saint-Quentin University.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2015. All rights reserved.

Downloaded on November 12, 2026 from www.jrheum.org

Address correspondence to Prof. D'Agostino, Rheumatology Department, Ambroise Paré Hospital, APHP, 9 avenue Charles de Gaulle, 92100 Boulogne-Billancourt, France.
E-mail: maria-antonietta.dagostino@apr.aphp.fr

Gout is the most common inflammatory joint disease[1,2] and is caused by the formation and deposition of monosodium urate (MSU) crystals in joints, and/or soft tissues. The conventional way to diagnose gout is based on clinical history and physical examination in combination with serum urate levels — but the key element in the diagnostic process is still joint or tophus aspiration and subsequent polarized microscopy for demonstration of MSU crystals[3]. Uncontrolled hyperuricemia has been associated with renal failure, and cardiovascular disease, with subsequent increased morbidity and mortality[4,5], stressing the importance of early and accurate diagnosis. This need, in combination with the development of new therapeutic options, has focused attention on new imaging modalities and their potential to improve diagnosis and disease management, possibly rendering joint puncture unnecessary. A recent systematic literature review, as part of the 3e Initiative on Diagnosis and Management of Gout, highlighted dual energy computerized tomography (DECT) and ultrasound (US) as promising imaging tools that have shown markedly better performance than clinical findings, although further testing is warranted[6]. In comparison to DECT, US appears to be more feasible because it is patient-friendly, safe and non-invasive, free of ionizing radiation, allows multiple target assessment in real time, and may aid the aspiration process for relevant material for microscopy[7]. Several published studies have highlighted the ability of US to detect crystal deposits and have addressed its possible role in very early diagnosis, based on their US findings indicative of gout in patients with asymptomatic hyperuricemia[7,8,9].

It is therefore critical to evaluate the US definitions proposed for gout, and their reliability; and to pinpoint areas in which further research is needed before US may be considered a potential imaging technique in this area.

The Outcome Measures in Rheumatology (OMERACT) US Working Group (Appendix 1) developed a gout subgroup with the purpose of validating US as an imaging tool for gout. If this objective is achieved, US may be implemented as an outcome measure in gout.

Is Ultrasound a Validated Outcome Measure in Gout?

In 2013, a systematic literature review was published evaluating US as an outcome tool in gout and asymptomatic hyperuricemia[10]. The report found 18 out of 67 articles published since 1975 to be eligible for review. Described in the literature were 4 main pathologies related solely to gout: tophi, double contour sign (DC), soft tissue abnormalities, and bony lesions.

The review highlighted that US was able to detect tophi using magnetic resonance imaging (MRI) as a gold standard, and this measure was found sensitive to change. The DC is an articular cartilage abnormality related to the deposition of crystals on the surface of the hyaline cartilage, which seemed specific to gout, with excellent inter-reader reliability and sensitive to change (the latter only in a very small patient population). Soft tissue pathology such as intrasynovial hyperechogenicity may be indicative of gout. US was less sensitive than MRI for diagnosing erosions (bony lesions) but more sensitive than conventional radiography, as is also known from rheumatoid arthritis studies[11,12].

Criterion and construct validity were assessed only for tophi, and overall there was a lack of consensus on the definitions of the 4 elementary lesions and their validity according to the OMERACT filter[13].

Current Limitations of US in Gout Assessment

Despite clear interest in this imaging technique for the management of gout, the literature review clearly pointed to a lack of clear US definitions for the main 4 elementary lesions identified: tophi, DC, soft tissue hyperechogenicity (punctuate crystal aggregates), and bony lesions (erosions). This lack of consensus-based definitions impairs the ability to validate US according to the OMERACT filter and hampers widespread use of US in therapeutic clinical trials, due to the difficulty to measure the same phenomenon.

In order to implement US in the management of patients with established or suspected gout the "gout subgroup of the OMERACT US Working Group" initiated a validation process. The first step was to obtain consensus-based definitions for the US elementary lesion as indicated by the literature review. This was accomplished by performing a Delphi exercise[14]. Thirty-five rheumatologists performing US and with an interest in gout were invited to participate, and 32 responded positively. After 3 Delphi rounds, > 80% agreement was obtained for each definition (Figure 1), as follows:

Double contour. "Abnormal hyperechoic band over the superficial margin of the articular hyaline cartilage, independent of the angle of insonation and which may be either irregular or regular, continuous, or intermittent and can be distinguished from the cartilage interface sign."

Tophus [independent of location (e.g. extraarticular/intraarticular/intratendinous)]. "A circumscribed, inhomogeneous, hyperechoic and/or hypoechoic aggregation (which may or may not generate posterior acoustic shadow), which may be surrounded by a small anechoic rim."

Aggregates [independent of location (intraarticular/intratendinous)]. "Heterogeneous hyperechoic foci that maintain their high degree of reflectivity even when the gain setting is minimized or the insonation angle is changed, and which occasionally may generate posterior acoustic shadow."

Erosions. "An intra- and/or extraarticular discontinuity of the bone surface (visible in 2 perpendicular planes)."

Agreement was obtained to use the existing definitions for both synovitis and tenosynovitis[15] because these may be

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2015. All rights reserved.

Downloaded on November 12, 2026 from www.jrheum.org

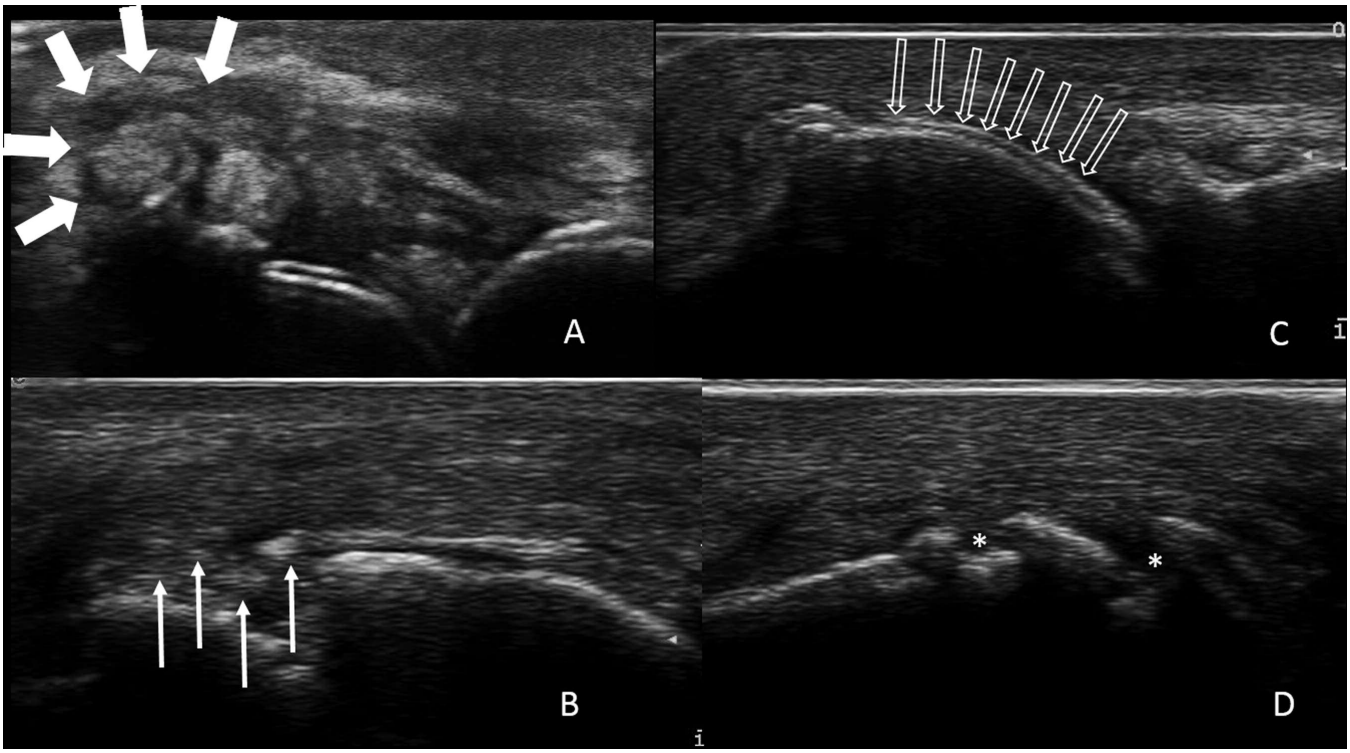

Figure 1. Ultrasound images assessed in the Delphi exercise to evaluate proposed definitions of elementary lesions. A. Bold white arrows: tophus. B. Narrow white arrows: aggregates. C. Open arrows: double contour. D. Asterisks: ultrasound examples of elementary lesions evaluated in the Delphi exercise.

co-components in gout disease. Agreement could not be obtained to include synovitis (including Doppler activity) as an elementary lesion indicative of gout, because the presence of synovitis alone was not considered specific enough to define gout disease because it is a key component in other inflammatory arthropathies as well[14]. On the other hand, even if erosions may also be seen in other arthropathy conditions, since they may also be found extraarticularly in gout and may possibly have a slightly different appearance, it was decided to test, as part of the Delphi exercise, whether the existing definition worked also in gout. Perfect agreement was obtained to keep the definition close to the definition used for erosions in general.

The second step was to test the reliability of the obtained definitions in a Web exercise consisting of static images of the elementary lesions. The Web exercise included 110 US images of the 4 lesions obtained from feet and knees and 20 of these images were shown twice in order to test both inter- and intrareader reliability. Twenty-seven of the 35 rheumatologists participating in the Delphi exercise participated in the reliability study. Cohen's κ was used to evaluate inter- and intrareader reliability. K values 0–0.20 were considered poor; 0.20–0.40 fair; 0.40–0.60 moderate; 0.60–0.80 good; and 0.80–1 excellent[16]. The unweighted κ values for all 4 lesions when reading the static images ranged from moderate to almost perfect for interobserver reliability (0.54–0.98), and similar results were found for the intrareader reliability (0.65–0.93). The US lesion with a higher level of reliability among the participants was the DC (intrareader: 0.98 and inter-reader: 0.93) whereas the lowest level of agreement was seen for the soft tissue aggregates (intrareader: 0.54 and inter-reader: 0.65)[14].

The third step was to test the agreement and reliability of the elementary lesions in a cohort of patients with gout. Sixteen of the rheumatologists previously involved in the first and second step participated in a workshop with 8 patients with crystal-proven gout. Both intra- and inter-reader reliability was assessed by scanning the patients twice within the same day. The areas of attention were the intercondylar region of the knee, the 1st metatarsophalangeal (MTP) joint, and the patellar tendon.

Cohen's κ was used to evaluate inter- and intrareader reliability. K values of 0–0.20 were considered poor; 0.20–0.40 fair; 0.40–0.60 moderate; 0.60–0.80 good; and 0.80–1 excellent[16]. The results showed that the intraobserver agreement was good for all lesions (0.61–0.73) except DC, which was found to be only moderate (0.53). The interobserver agreement was surprisingly poor for aggregates and DC (0.21 and 0.47, respectively) but moderate for the other components (tophus 0.69 and erosions 0.74)[17]. "Aggregates" was the lesion with the lowest reliability in the static image exercise[14] but even lower when assessed in patients, especially between observers, raising the issue as to whether or not the definition covers the US lesions it is supposed to

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2015. All rights reserved.

Downloaded on November 12, 2026 from www.jrheum.org

describe (soft tissue hyperechogenicity believed to be punctuate crystal deposits). Because a tophus is also a collection of crystals, there may be overlap that impairs agreement between observers. The US definition for "aggregates" must be improved before it may be considered a US elementary lesion in gout.

At this point in the validation process no attempt was made to define which components define acute gout and which define chronic gout; nor was any attempt made to define a core set of components to constitute a scoring system.

Future Steps

Tophus burden is proposed as part of the clinical core domain in chronic gout, and joint inflammation and joint damage imaging are proposed as a discretionary domain. Joint swelling and joint tenderness are proposed as part of the clinical core domain for acute gout[18]. For both chronic and acute gout the suggested parameters in the core domain set may be evaluated by US. The first steps, namely, to validate further use of US in gout have been taken. The elementary US components in gout that have been identified by the literature review are: DC, tophi, aggregates (all aspects of the urate burden), and erosions (component of structural damage).

Development of US consensus-based definitions of gout lesions is the first initiative towards achieving a higher degree of homogeneity and comparability of results between studies and in daily clinical practice. The first steps in evaluating the reliability of the definitions show moderate-to-perfect reliability in static images and somewhat lower results in patients, indicating that a standardized scanning technique and a core set of joints are needed. Existing studies indicate that such a core set of joints may include 1st MTP joints and knees, although ankles and some tendons also appear to be frequently involved[19,20]. The next step is to validate the definitions following all steps covered by the OMERACT filter. Further studies are already planned that aim to: (1) investigate whether aggregates is an elementary lesion in gout (because of poor reliability in static images and when assessing patients) by conducting a Web-exercise to identify what is perceived as aggregates; (2) determine responsiveness of lesions during urate-lowering therapy in a multicenter prospective study; and (3) determine the discriminative ability of US gout lesions, in comparison to those of other arthropathies, in a patient workshop. The development of a scoring system with possible inclusion of synovitis, tenosynovitis, and Doppler activity will be a future step to ensure US as an outcome measure in clinical trials.

ACKNOWLEDGMENT

The following authors contributed equally to this report: Marwin Gutierrez, Wolfgang A. Schmidt, Helen I. Keen, Emilio Filippucci, David Kane, Ralf Thiele, Gurjit Kaeley, Peter Balint, Peter Mandl, Andrea Delle Sedie, Hilde Berner Hammer, Robin Christensen, Ingrid Möller, Carlos Pineda, George A. Bruyn, Annamaria Iagnocco, and Esperanza Naredo.

APPENDIX 1

List of study collaborators. OMERACT Ultrasound Working Group: Philippe Aegerter, Sibel Aydin, Marina Backhaus, David Bong, Isabelle Chary-Valckenaere, Paz Collado, Eugenio De Miguel, Christian Dejaco, Oscar Epis, Jane E. Freeston, Frederique Gandjbakhch, Walter Grassi, Petra Hanova, Sandrine Jousse-Joulin, Fredrick Joshua, Juhani Koski, Damien Loeuille, Viviana Ravagnani, Anthony Reginato, Veronica Sharp, Nanno Swen, Marcin Szkudlarek, Richard J. Wakefield, and Hans-Rudolf Ziswiler.

REFERENCES

1. Mikuls TR, Farrar JT, Bilker WB, Fernandes S, Schumacher HR Jr, Saag KG. Gout epidemiology: Results from the UK General Practice Research Database, 1990-1999. Ann Rheum Dis 2005;64:267-72.
2. Zhu Y, Pandya BJ, Choi HK. Prevalence of gout and hyperuricemia in the US general population: the National Health and Nutrition Examination Survey 2007-2008. Arthritis Rheum 2011;63:3136-41.
3. Hamburger M, Baraf HS, Adamson TC 3rd, Basile J, Bass L, Cole B, et al. 2011 recommendations for the diagnosis and management of gout and hyperuricemia. Postgrad Med 2011;123:3-36.
4. Choi HK, Curhan G. Independent impact of gout on mortality and risk for coronary heart disease. Circulation 2007;116:894-900.
5. Krishnan E, Baker JF, Furst DE, Schumacher HR. Gout and the risk of acute myocardial infarction. Arthritis Rheum 2006;54:2688-96.
6. Sivera F, Andrès M, Falzon L, van der Heijde DM, Carmona L. Diagnostic value of clinical, laboratory, and imaging findings in patients with a clinical suspicion of gout: a systematic literature review. J Rheumatol Suppl 2014;92:3-8.
7. De Miguel E, Puig JG, Castillo C, Peiteado D, Torres RJ, Martín-Mola E. Diagnosis of gout in patients with asymptomatic hyperuricaemia: a pilot ultrasound study. Ann Rheum Dis 2012;71:157-8.
8. Pineda C, Amezcua-Guerra LM, Solano C, Rodriguez-Henríquez P, Hernández-Díaz C, Vargas A, et al. Joint and tendon subclinical involvement suggestive of gouty arthritis in asymptomatic hyperuricemia: an ultrasound controlled study. Arthritis Res Ther 2011;13:R4.
9. Puig JG, de Miguel E, Castillo MC, Rocha AL, Martínez MA, Torres RJ. Asymptomatic hyperuricemia: impact of ultrasonography. Nucleosides Nucleotides Nucleic Acids 2008;27:592-5.
10. Chowalloor PV, Keen HI. A systematic review of ultrasonography in gout and asymptomatic hyperuricaemia. Ann Rheum Dis 2013;72:638-45.
11. Wakefield RJ, Gibbon WW, Conaghan PG, O'Connor P, McGonagle D, Pease C, et al. The value of sonography in the detection of bone erosions in patients with rheumatoid arthritis: a comparison with conventional radiography. Arthritis Rheum 2000;43:2762-70.
12. Szkudlarek M, Narvestad E, Klarlund M, Court-Payen M, Thomsen HS, Østergaard M. Ultrasonography of the metatarsophalangeal joints in rheumatoid arthritis: comparison with magnetic resonance imaging, conventional radiography, and clinical examination. Arthritis Rheum 2004;50:2103-12.
13. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino MA, et al. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. J Clin Epidemiol 2014;67:745-53.
14. Gutierrez M, Schmidt WA, Thiele R, Keen H, Kaeley G, Naredo E, et al. International consensus for ultrasound lesions in gout. Results of Delphi process and Web-reliability exercise. Rheumatology 2015; May 13 (E-pub ahead of print).
15. 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.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2015. All rights reserved.

Downloaded on November 12, 2026 from www.jrheum.org

16. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33:159-74.
17. Terslev L, Gutierrez M, Schmidt WA, Keen HI, Filippucci E, Kane D, et al. Ultrasound as an outcome measure in gout. A validation process by the OMERACT Ultrasound Working Group. J Rheumatol 2015;42:2177-81.
18. Schumacher HR, Taylor W, Edwards L, Grainger R, Schlesinger N, Dalbeth N, et al. Outcome domains for studies of acute and chronic gout. J Rheumatol 2009;36:2342-5.
19. Peiteado D, De Miguel E, Villalba A, Ordóñez MC, Castillo C, Martín-Mola E. Value of a short four-joint ultrasound test for gout diagnosis: a pilot study. Clin Exp Rheumatol 2012;30:830-7.
20. Naredo E, Uson J, Jiménez-Palop M, Martínez A, Vicente E, Brito E, et al. Ultrasound-detected musculoskeletal urate crystal deposition: which joints and what findings should be assessed for diagnosing gout? Ann Rheum Dis 2014;73:1522-8.

Articles presented at the OMERACT 12 Conference, Budapest, Hungary, May 7-11, 2014

- Part 1: Pre-conference Meeting
- Part 2: Working Groups
- Part 3: Special Interest Groups I
- Part 4: Special Interest Groups II

Part 3 will appear in the December issue.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2015. All rights reserved.

Downloaded on November 12, 2026 from www.jrheum.org