

## Original article

## International Consensus for ultrasound lesions in gout: results of Delphi process and web-reliability exercise

Marwin Gutierrez<sup>1</sup>, Wolfgang A. Schmidt<sup>2</sup>, Ralf G. Thiele<sup>3,\*</sup>, Helen I. Keen<sup>4,\*</sup>, Gurjit S. Kaeley<sup>5,\*</sup>, Esperanza Naredo<sup>6,\*</sup>, Annamaria Iagnocco<sup>7,\*</sup>, George A. Bruyn<sup>8,\*</sup>, Peter V. Balint<sup>9,\*</sup>, Emilio Filippucci<sup>1,\*</sup>, Peter Mandl<sup>10,\*</sup>, David Kane<sup>11,\*</sup>, Carlos Pineda<sup>12,\*</sup>, Andrea Delle Sedie<sup>13,\*</sup>, Hilde Berner Hammer<sup>14,\*</sup>, Robin Christensen<sup>15,\*</sup>, Maria Antonietta D'Agostino<sup>16,\*</sup> and Lene Terslev<sup>17</sup>, on behalf of the OMERACT Ultrasound Gout Task Force group

## Abstract

**Objective.** To produce consensus-based definitions of the US elementary lesions in gout and to test their reliability in a web-based exercise.

**Methods.** The process consisted of two steps. In the first step a written Delphi questionnaire was developed from a systematic literature review and expert international consensus. This collated information resulted in four statements defining US elementary lesions: double contour (DC), tophus, aggregates and erosion. The Delphi questionnaire was sent to 35 rheumatology experts in US, asking them to rate their level of agreement or disagreement with each statement. The second step tested the reliability by a web-exercise. US images of both normal and gouty elementary lesions were collected by the participants. A facilitator then constructed an electronic database of 110 images. The database was sent to the participants, who evaluated the presence/absence of US elementary lesions. A group of 20 images was displayed twice to evaluate intra-reader reliability.

**Results.** A total of 32 participants responded to the questionnaires. Good agreement (>80%) was obtained for US definitions on DC, tophus, aggregates and erosion in the Delphi exercise after three rounds. The reliability on images showed inter-reader  $\kappa$  values for DC, tophus, aggregates, erosion findings of 0.98, 0.71, 0.54 and 0.85, respectively. The mean intra-reader  $\kappa$  values were also acceptable: 0.93, 0.78, 0.65 and 0.78, respectively.

**Conclusion.** This, the first consensus-based US definition of elementary lesions in gout, demonstrated good reliability overall. It constitutes an essential step in developing a core outcome measurement that permits a higher degree of homogeneity and comparability between multicentre studies.

**Key words:** gout, ultrasound, elementary lesions, Delphi process, web reliability.

<sup>1</sup>Rheumatology Department, Clinica Reumatologica, Università Politecnica delle Marche, Jesi, Ancona, Italy, <sup>2</sup>Rheumatology Department, Immanuel Krankenhaus, Medical Centre for Rheumatology, Berlin, Germany, <sup>3</sup>Department of Medicine, Allergy/Immunology and Rheumatology Division, School of Medicine and Dentistry, University of Rochester, Rochester, NY, USA, <sup>4</sup>School of Medicine and Pharmacology Fiona Stanley Hospital Unit, University of Western Australia, Perth, Australia, <sup>5</sup>Division of Rheumatology, University of Florida College of Medicine, Jacksonville, FL, USA, <sup>6</sup>Rheumatology Department, Hospital General Universitario Gregorio Marañón, Madrid, Spain, <sup>7</sup>Dipartimento di Medicina Interna e Specialità Mediche, Sapienza Università di Roma, Rome, <sup>8</sup>Rheumatology Department, MC Groep Hospitals, Lelystad, Netherlands, <sup>9</sup>3rd Rheumatology Department, National Institute of Rheumatology and Physiotherapy, Budapest, Hungary, <sup>10</sup>Division of Rheumatology, Medical University of Vienna, Vienna, Austria, <sup>11</sup>Rheumatology Department, Trinity College, Dublin, Ireland, <sup>12</sup>Rheumatology Department, Instituto Nacional de Rehabilitación,

Mexico City, Mexico, <sup>13</sup>Rheumatology Unit, University of Pisa, Pisa, Italy, <sup>14</sup>Rheumatology Department, Diakonhjemmet Hospital, Oslo, Norway, <sup>15</sup>Musculoskeletal Statistics Unit, The Parker Institute, Copenhagen University Hospital, Bispebjerg and Frederiksberg, Copenhagen, Denmark, <sup>16</sup>INSERM U1173, Rheumatology Department, Laboratoire d'excellence INFLAMEX, UFR Simone Veil, Versailles-Saint-Quentin University, Saint-Quentin en Yvelines, France and <sup>17</sup>Rheumatology Department, Center for Rheumatology and Spine diseases, Copenhagen University Hospital, Copenhagen, Denmark

Submitted 17 July 2014; revised version accepted 9 March 2015

Correspondence to: Marwin Gutierrez, Clinica Reumatologica, Università Politecnica delle Marche, Ospedale C. Urbani, Via dei Colli, 52, Jesi (Ancona), 60035, Italy. E-mail: dr.gmarwin@gmail.com

\*These authors contributed equally to this study.

### Rheumatology key messages

- The results provide OMERACT consensus-based definitions of US gout lesions.
- The web-exercise reliability of US gout definitions showed overall good  $\kappa$  values.
- This is the first step to ensure homogeneity and comparability between US multicentre studies in gout.

## Introduction

Gout is an inflammatory disease induced by the precipitation of MSU crystals in a variety of tissues, both inside and around joints, where they can lead to acute or chronic arthritis [1]. Its poor control can lead to renal failure, cardiovascular disease, increased morbidity–mortality and poorer quality of life [2, 3].

Recent data suggest that gout affects up to 1–2.5% of adults; however, its prevalence increases with age to rates of up to 7% in men aged over 65 and 3% in women aged over 85 [4, 5]. The treat-to-target approach is to lower serum uric acid levels, allowing MSU crystals to dissolve, leading to both the elimination of acute gouty attacks of inflammation and the reduction of tophi formation [4]. However, despite new guidelines and recommendations for diagnosis and management, and evolving new therapeutic options, suboptimal management of the disease is still reported [4, 6–9]. Early detection and accurate diagnosis of gout play a crucial role in improving the outcome of the disease, which is still a challenge in clinical daily practice due to the great heterogeneity of the disease and the non-specificity of symptoms in its early phases.

Imaging may play an important role, especially US, due to its utility in both clinical practice and research activity [10–24]. It offers relevant properties such as direct visualization of crystal deposits in the tissues, high sensitivity in identifying early anatomical changes, and the ability to determine therapy efficacy [21, 23–26]. Additionally it is patient friendly, safe and non-invasive, free of ionizing radiation, less expensive, allows multiple-target assessment in real time and can be performed at the bedside.

To date, several studies supporting its accuracy, and construct and criterion validity as well reliability in gout have been published [15, 16, 21, 23, 24]. Interesting data have also emerged from studies speculating about its possible role in the very early diagnosis of the disease, because US findings indicative of gout have also been demonstrated in patients with asymptomatic hyperuricaemia [27–30].

Despite the growing body of evidence supporting the utility of US in gout, standardized definitions of the elementary morphostructural changes and their reliability have not been thoroughly established. Several authors have developed and published definitions for US gout lesions, but the use of these definitions has been limited to local clinical settings. In order to improve the use of US in the evaluation of patients with gout and strengthen its capability as an outcome instrument, an OMERACT US Gout Task Force was formed.

The first aim of this study was to produce consensus-based definitions of the US elementary lesions in gout using a Delphi process among an international panel of

experts. The second aim was to test the inter- and intra-reader reliability of the US definitions of each elementary lesion in gout using a web-based exercise.

## Methods

### Study design

The OMERACT US Gout Task Force was formed and held their first international meeting at the ACR congress in 2012. A systematic review of the literature was presented, and relevant questions regarding the definition of gout elementary lesions were discussed in order to plan the Delphi exercise. A validation process was started and the first two steps were defined for carrying out a Delphi exercise on the US definitions of elementary lesions in gout and for testing these definitions in an inter- and intra-reader reliability web exercise. The institutional ethics committee (Comitato Etico della Azienda Sanitaria Unica Regionale di Ancona) approved the study, and informed consent was obtained from patients scanned for the inter- and intra-reader exercise.

### First step: Delphi consensus on definition of gout elementary US lesions

Thirty-five rheumatologists (experienced in both US and gout) from 15 countries of Europe, America and Australasia (Australia, Austria, Denmark, France, Germany, Hungary, Ireland, Italy, Mexico, Netherlands, Norway, Turkey, Spain, UK and USA) were identified. The 35 experts were then emailed invitations to participate in the Delphi consensus-building exercise.

A written Delphi questionnaire was constructed on the basis of data collected from both the literature on US in gout and discussion among OMERACT US Gout Task Force members. The criteria presented for committee scrutiny were assembled from PubMed and Medline literature searches as well as from highly cited manuscripts on US in gout [10–26]. Abstracts presented at the 2012 ACR and EULAR scientific meetings were also included. The previous definitions of US elementary lesions described in the literature were used as a basis to formulate the Delphi statements presented to the panel.

The first online survey that was sent to the participating committee members consisted of 23 statements/items, including definitions for: double contour (DC), synovitis, tophus (intra-articular and intra-tendinous) and erosions (supplementary Table S1, available at *Rheumatology* Online). The names of the panellists were kept confidential and all responses were re-identified prior to releasing them to the group. This allowed each member to answer questions without being influenced by the opinions of the other panellists. The panel was asked to rate each item

using a level of agreement or disagreement for each statement according to a 1–5 Likert scale [31].

The answers from each Delphi questionnaire were summarized with mean scores by a facilitator and re-sent with a revised questionnaire to the panel for the next round. The successive rounds consisted of new surveys and included only the items that had achieved significant agreement among the participants. An additional criterion suggested by the panel during the first interaction, as well as items that required rewording for definitions, were included in the new survey.

### Second step: inter- and intra-reader reliability web exercise

The participants collected representative US images of both normal and gouty elementary lesions agreed upon during the preliminary Delphi process. Each participant was asked to collect at least one US image of each individual elementary lesion presented in the longitudinal and transverse scans (Fig. 1). The anatomical sites requested were as follows: knee (hyaline cartilage of the femoral condyles; patellar tendon, including both proximal and distal insertion; femoral bone profile), ankle (Achilles tendon) and foot (first MTP joint for hyaline cartilage, bone profile, periarticular tissue). These anatomical areas were selected because of their accessibility by US and their frequent involvement in patients with gout.

After a collection period of 1 month, the images were sent by e-mail to a facilitator, who constructed an electronic database of 110 images based on five findings (1, DC; 2, tophus; 3, aggregates; 4, erosion; 5, no gouty findings). It was re-sent to the participants, asking them to read each image and determine the presence/absence of the lesion by applying the definitions agreed in the first step of the validation process. A randomly selected group of 20 images was displayed twice in order to assess intra-reader reliability.

### Statistical analysis

In the Delphi process, the agreement was assessed using a 1–5 Likert scale, which was graded as follows: 1, strongly disagree; 2, disagree; 3, neither agree nor disagree; 4, agree; 5, strongly agree. Group agreement with the issue under consideration was defined as total cumulative agreement >80%. Only when sentences achieved a score >80%, did we consider that the group had reached a consensus and that the category was defined as appropriate. Only the statements satisfying these requirements were used for defining the definitive elementary lesions.

Inter- and intra-reader reliability between the participants were estimated (by unweighted Cohen's *k*-statistic) in terms of dichotomous assessment (i.e. presence or absence) and interpreted as follows: values of: 0–0.20 represent slight; 0.21–0.40 fair; 0.41–0.60 moderate; 0.61–0.80 substantial and >0.80 almost perfect reliability [32]. Additionally, the 95% CIs were calculated. Statistical analysis was performed using PROC GLIMMIX, applying SAS software (v. 9.2; SAS Institute Inc., Cary, NC, USA).

## Results

### Delphi exercise

Thirty-two of 35 participants responded to the first Delphi questionnaire (91% response rate). All 32 participants also responded to the second and third rounds of the Delphi questionnaire (100% response rate).

As previously mentioned, the first round of the Delphi exercise included 23 statements regarding US elementary lesions: DC, synovitis, tophi (intra-articular and intra-tendinous) and erosions. Of the 23 statements with the highest agreement, 11 were re-worded according to the comments suggested in the first interaction and presented in the second round to the participants. The panel agreed not to include synovitis among the elementary lesions of gout. Moreover, they suggested the term aggregates rather than intra-tendinous tophus. Thus, the elementary lesions included in the second Delphi exercise were: DC, aggregates independent of location (e.g. extra-articular/intra-articular/intra-tendinous), tophus independent of location (e.g. extra-articular/intra-articular/intra-tendinous) and erosions. Table 1 shows the total cumulative agreement (percentage) for each statement of the second round. After this, consensus for aggregates and erosions (90% and 87%, respectively) was reached, but no consensus was found for DC or tophus. A third and final survey interaction was presented to the panel participants. It included a rewording of the definitions that did not reach the 80% of consensus in the second round (DC and tophus). Finally, a definitive consensus on four US elementary lesions was reached after the third round, with agreement for DC, aggregates, tophus and erosions (91%, 90%, 81%, and 87%, respectively). The final definitions were as follows:

DC: 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.

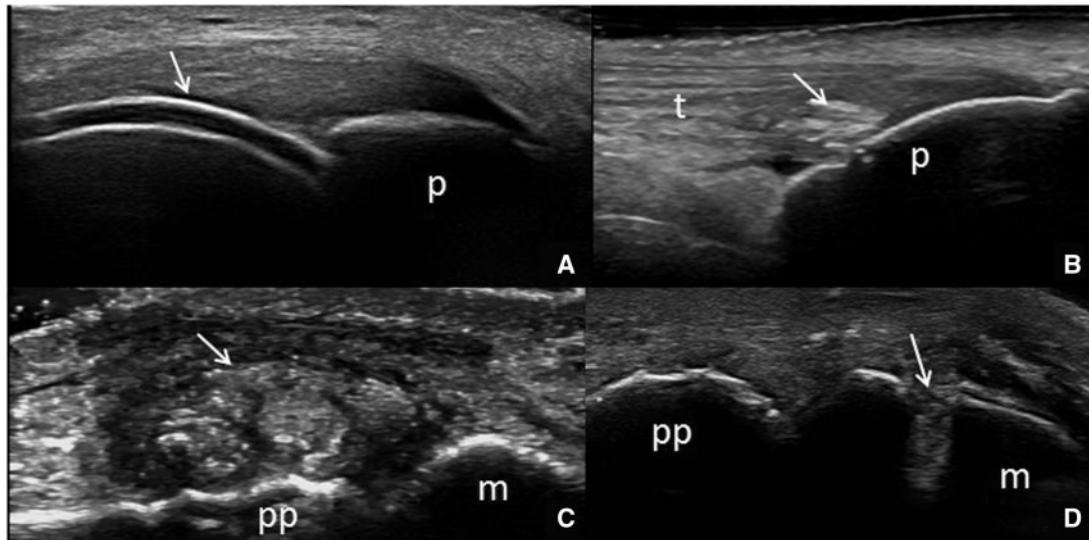
Tophus [independent of location (e.g. extra-articular/intra-articular/intra-tendinous)]: a circumscribed, inhomogeneous, hyperechoic and/or hypoechoic aggregation (which may or may not generate posterior acoustic shadow), which may be surrounded by a small anechoic rim.

Aggregates [independent of location (intra-articular/intra-tendinous)]: 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.

Erosion: an intra- and/or extra-articular discontinuity of the bone surface (visible in two perpendicular planes).

### Reliability of reading images

Of the 32 participants, 27 (84%) collected a total of 110 US images of the definitive elementary lesions including the anatomical areas previously described. Twenty of those 110 images were repeated twice for the assessment

**Fig. 1** US findings of MSU deposits in a range of tissues

(A) Longitudinal suprapatellar scan with maximum knee flexion showing the double contour sign characterized by the hyperechoic band over the superficial margin of the articular hyaline cartilage of the femoral condyle (arrow). (B) Knee. Patellar tendon (t). Note the derangement of the normal fibrillar echotexture of the tendon caused by the aggregates of MSU crystals. (C) Foot. First MTP joint showing an intra-articular tophus. Note the circumscribed, inhomogeneous, hyperechoic aggregation, without acoustic shadow, surrounded by a small anechoic rim (arrow). (D) Foot. First MTP joint, showing an intra-articular discontinuity of the bone surface (arrow). Note that B, C, D images show proximal on the right and distal on the left. p: patella; pp: proximal phalanx; m: metatarsal head.

of intra-reader reliability. Ten additional normal US images were included in the electronic system. The US equipment used for image collections differed from institution to institution and comprised of the following: General Electric Logiq 9 and E9, Esaote XVG, 60, 70 and Twice, Siemens Antares, Phillips HDI 5000.

The observed agreement and the kappa coefficient concerning inter and intra-reader reliability are listed in Table 2. Both  $\kappa$ -values and 95% CI showed moderate to almost perfect reliability between the investigators for all the US elementary lesions. The lowest  $\kappa$  values were obtained for aggregates (0.54), whereas the best rate was for DC (0.98). Similar results were obtained for the intra-reader reliability, with the kappa coefficients ranging from 0.65 to 0.93 (Table 2).

## Discussion

A standardized US approach for patients with gout has not yet been proposed, and there is still no international consensus on the definitions of elementary lesions. This was also underlined in a recent systematic literature review, which showed non-uniformity in the definitions being applied, and raised the necessity for a standardized definition of US findings in gout [30]. Although important advances in the care of gout have been achieved in recent years, its management is far from optimal in either primary care or rheumatology practice [6–9]; therefore, the quality of assessment provided to gout patients needs to improve.

US plays a key role in the management of these patients as some of the elementary lesions are demonstrated to be highly specific for the diagnosis [10, 14, 21–25]. Its role was also highlighted in the EULAR evidence-based recommendations for gout as a potential tool for diagnosis and monitoring of gout [6, 7].

Taking into account the growing number of rheumatologists who are incorporating US into their research and clinical assessment of gout, the development of a standardized approach is imperative. Therefore, the OMERACT US Gout Task Force conducted a Delphi consensus process and web-based reliability exercise in order to obtain expert agreement on those elementary lesions that should be included in the US definition of gout and subsequently tested agreement on these definitions on US images.

Three rounds of Delphi exercise were necessary to reach an agreement on the definitive four US gout elementary lesions highlighted in the literature review: DC, aggregates, tophus and erosions. No agreement was obtained for the inclusion of synovitis (including power Doppler), which has been suggested by other authors in previous papers. The main reasons for the disagreement may include (i) the low specificity of these findings for gout [33] since they are frequent in other inflammatory or degenerative chronic arthropathies and (ii) the availability of standardized definitions for this lesion, which were taken as the reference for this exercise [34, 35]. However, we are aware that specific definitions for synovitis and tenosynovitis for gout are needed, so this aspect will be part of a



**TABLE 1** Total cumulative agreement (percentage) for each statement of the second and third rounds

Statements round 2	Cumulative agreement, %	Category appropriated	Statements round 3	Cumulative agreement, %	Category appropriated
Double contour (DC)					
Abnormal hyperechoic band over the an-echoic superficial margin of the articular hyaline cartilage, which can range from focal to homogeneous thickening, and whose reflectivity is independent of the angle of insonation.	59.3	No	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.	91	Yes
Abnormal hyperechogenicity of the superficial margin of the articular hyaline cartilage, which is of irregular thickness, independent of angle of insonation and can be either continuous or intermittent (focal).	59.3	No			
Abnormal hyperechoic band over the an-echoic superficial margin of the articular hyaline cartilage, which can range from focal to homogeneous thickening and whose reflectivity is independent of the angle of insonation and which can be distinguished from the cartilage interface sign.	72	No			
Soft tissue aggregates					
Heterogeneous substance composed by hyperechoic and/or hypoechoic material (which may generate snowstorm appearance when probe pressure is applied) that may generate posterior acoustic shadow.	47	No			
Heterogeneous substance containing hyperechoic foci (bright stippled foci) > 1 mm, without acoustic shadowing and independent of gain.	31	No			
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.	87.5	Yes			
Tophus					
Inhomogeneous hyperechoic/hypoechoic material (which may or may not generate posterior acoustic shadows) surrounded by a small anechoic rim.	62.5	No	A circumscribed, inhomogeneous, hyperechoic and/or hypoechoic aggregation (which may or may not generate a posterior acoustic shadow), which may be surrounded by a small anechoic rim.	81	Yes

(continued)

TABLE 1 Continued

Statements round 2	Cumulative agreement, %	Category appropriated	Statements round 3	Cumulative agreement, %	Category appropriated
Inhomogeneous hyperechoic/hypoechoic material, which may generate posterior acoustic shadows.	46.8	No			
Erosion					
An intra- and/or extra-articular discontinuity of the bone surface (visible in two perpendicular planes).	84	Yes			
An intra- and/or extra-articular discontinuity of the bone surface (visible in two perpendicular planes) that may be associated with aggregates.	37	No			
Discontinuity of the bone surface (visible in two perpendicular planes) that may be accompanied by a cortical extra-articular defect and may associated with hyperechoic cloudy material within the erosion.	40.6	No			

forthcoming international meeting of the OMERACT US Gout Task Force meeting and exercises on scoring and sensitivity to change.

After the Delphi exercise, the web exercise tested inter- and intra-reader reliability, showing good  $\kappa$  values in all US definitions. The range of  $\kappa$  values for inter-observer reliability was moderate to almost perfect (0.98, 0.71, 0.54, 0.85). Similar results were also detected for intra-reader reliability. The US lesion with the highest level of agreement among the participants was DC, whereas the lowest level of agreement was found for aggregates, which is in line with previous studies that tested the reliability of US gout findings [22, 36]. One reason for the overall good reliability for elementary lesions in gout could be related to the type of image used in the web exercise, which demonstrated typical and clear US findings. On the other hand, there are possible explanations for the relatively low level of agreement on aggregates. The panel of US images used for the exercise included aggregates with a range of levels of difficulty of detection (as in daily practice); while large MSU aggregates could be found easily, small crystal aggregates could be more difficult to identify. Another reason could be related to the use of only static images for the web exercise, rather than video-dynamic clips, which may limit the global visualization of the area of interest and the optimization of the US images. Future exercises testing the US definitions on patients are needed to establish whether this will hold for clinical scanning, taking image acquisition into account. In particular, the utility of a separate definition of aggregates will be analysed on the basis of the results of the reliability exercise on gouty patients (real-life scanning). If aggregates are confirmed to have the lowest reliability, the possibility of merging the two definitions (tophus and aggregates) will be considered. Another crucial issue to solve is the use of the terminology tophus because it is a histological concept as opposed to an imaging description. Although several imaging studies of US, MRI and CT [10–15, 17, 19, 21–23, 25, 37, 38] have commonly adopted this term, we believe that further work is needed in order to improve this lack of consensus in the wider international community. Thus, the OMERACT US Gout Task Force is planning studies involving close collaboration with histologists in order to delineate a common understanding of the terminology.

The present consensus for US lesions in gout requires additional considerations. First, it represents the first step of the OMERACT validation process of US definitions in gout and is based on an international panel of US gout experts with broad daily clinical and sonographic practice. Second, the available evidence was appraised and summarized following a rigorous approach and was combined with the experience of several rheumatologists. Third, the OMERACT gout US Task Force discussed largely the details relating to the Delphi process and the way in which propositions were developed. The Delphi approach has several advantages: reduction in individual bias afforded by anonymity, equal weighting afforded to all members, good external validity and ready identification

**TABLE 2** Inter- and intra-reader reliability results for the four US elementary lesions

US lesion	Inter-reader reliability: unweighted, <i>k</i> -values (95% CI)	Intra-reader reliability: unweighted, <i>k</i> -values (95% CI)
Double contour	0.98 (0.97, 0.98)	0.93 (0.821, 1.05)
Tophus	0.71 (0.60, 0.79)	0.78 (0.375, 1.19)
Aggregates	0.54 (0.39, 0.36)	0.65 (0.312, 1.32)
Erosion	0.85 (0.79, 0.90)	0.78 (0.375, 1.19)

CI: confidence interval.

of areas where more research data are required [39]. Fourth, the US definitions were corroborated by a reliability exercise, which strengthens the reproducibility of the agreed definitions. Finally, the high level of agreement on the final US gout definitions and the multinational participation increases their utility and will hopefully facilitate their international dissemination and implementation.

We are aware that some limitations are present in this study. These are mainly related to the practicalities common to these types of exercise, such as: whether the aim should be the comprehensive coverage of all options or rather a selective highlight of only key issues; how best to edit and amalgamate statements submitted in the first round; subsequent modifications to wording or content of propositions to consider. An additional point to consider is related to the fact that in order to improve clarity, the Task Force undertook minor modifications to the wording of some propositions at the end of the process (after they had been researched, voted on and fully discussed in a meeting held in Berlin), but no change was made to the key content of the propositions.

The next step in the validation process is to test the definitions during scanning of gout patients and subsequently their diagnostic value in patients with gout and undiagnosed gout in multicentre studies. Another limitation is represented by the fact that only US images were used to test the reliability. This type of modality, generally, produces more reliable results with respect to US scanning of patients. However the OMERACT US Gout Task Force is currently conducting an exercise on patients in order to obtain more accurate and reliable results for US elementary lesions.

In conclusion, the results of this study provide preliminary consensus-based definitions of US gout lesions and the first step in ensuring a higher degree of homogeneity and comparability of results between studies and in daily practice. Ongoing investigations testing their validity and the responsiveness may strengthen these preliminary definitions.

## Acknowledgements

M.G., L.T. and W.S. participated in the design of the study and the acquisition and interpretation of data, performed the statistical analysis together with R.C., the drafting of

the manuscript and gave final approval of the version of the paper to be published. R.T., H.K., G.K., G.B., P.B., E.F., P.M., D.K., C.P., A.D.S., H.B.H., M.A.D. was involved in the Delphi process and in the web-reliability exercise, made substantial contributions to the manuscript preparation and was involved in revising the manuscript for important intellectual content. E.N., A.I., M.A.D. participated in the study development and gave substantial input to the data evaluation and manuscript preparation. All authors read and approved the final version of manuscript. OMERACT Ultrasound Gout Task Force group members and contributions: David Bong, Eugenio de Miguel, Christian Dejaco, Eugene Kissin, Frederique Gandjbakhch, Jane Freeston, Juhani Koski, Ingrid Möller, Marcin Szkudlarek, Nanno Swen, Anthony Reginato, Oscar Epis, Sibel Zehra Aydin, Viviana Ravagnani, Richard Wakefield, Petra Hanova, Ghada Mardenly and Veronica Sharp. D.B., Ed.M., C.D., E.K., F.G., J.F., J.K., I.M., M.S., N.S., A.R., O.E., S.Z.A., V.R., R.W., P.H., G.M. and V.S. were involved in the Delphi process and in the web-reliability exercise, made substantial contributions to the manuscript preparation and was involved in revising the manuscript for important intellectual content. All members approved the final version of manuscript.

**Funding:** No specific funding was received from any funding bodies in the public, commercial or not-for-profit sectors to carry out the work described in this manuscript.

**Disclosure statement:** W.A.S. is a member the speakers bureau, a consultant and has received grants/research support from Berlin-Chemie, Novartis, Esaote, General Electric and Siemens. M.G. has attended advisory board meetings, scientific consultancies and has obtained speaking fees from Abbott Immunology, AbbVie, UCB Pharma, Esaote S.p.a, and Bristol-Myers Squibb and Merck Sharp & Dohme. P.M. has obtained speaker fees from AbbVie, Pfizer, Janssen-Cilag and Bristol-Myers Squibb. H.B.H. has received honoraria from AbbVie, Bristol-Myers Squibb, Pfizer and Roche as well as research support from AbbVie, Roche and Pfizer. A.I. has attended advisory board meetings, scientific consultancies and has obtained speaking fees from Abbvie, General Electric, MSD and Bristol-Myers Squibb. E.N. has received speaker fees from Abbvie, Roche Farma, Bristol-Myers Squibb, Pfizer, UCB, General Electric Healthcare, and Esaote, and has received research

funding from MSD. E.F. has given expert testimony for General Electric Medical System and has obtained speaking fees from Abbott, UCB Pharma, Esaote S.p.a, General Electric Medical System, Bristol-Myers Squibb, Schering Plough and Merck Sharp & Dohme. P.V.B. has received speaker fees, consulting research grants and scientific consultancies from Abbvie, Bristol-Myers Squibb, General Electric Healthcare, Esaote S.p.a, Janssen-Cilag, MSD, Pfizer, Philips, Richter, Roche, SonoSite and UCB. All other authors have declared no conflicts of interest.

## Supplementary data

Supplementary data are available at *Rheumatology* Online.

## References

- 1 Tausche AK, Jansen TL, Schröder HE *et al.* Gout—current diagnosis and treatment. *Dtsch Arztebl Int* 2009;106:549–55.
- 2 Singh JA, Strand V. Gout is associated with more comorbidities, poorer health-related quality of life and higher health care utilization in US veterans. *Ann Rheum Dis* 2008;67:1310–6.
- 3 Krishnan E, Svendsen K, Neaton JD *et al.* Long-term cardiovascular mortality among middle-aged men with gout. *Arch Intern Med* 2008;168:1104–10.
- 4 Kuo CF, Grainge MJ, Mallen C, Zhang W, Doherty M. Rising burden of gout in the UK but continuing suboptimal management: a nationwide population study. *Ann Rheum Dis* 2015;74:661–7.
- 5 Lawrence RC, Helmick CG, Arnett FC *et al.* Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. *Arthritis Rheum* 1998;41:778–99.
- 6 Zhang W, Doherty M, Pascual E *et al.* EULAR evidence based recommendations for gout. Part I: Diagnosis. Report of a task force of the Standing Committee for International Clinical Studies Including Therapeutics (ESCI-SIT). *Ann Rheum Dis* 2006;65:1301–11.
- 7 Zhang W, Doherty M, Bardin T *et al.* EULAR evidence based recommendations for gout. Part II: Management. Report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCI-SIT). *Ann Rheum Dis* 2006;65:1312–24.
- 8 Hamburger M, Baraf HS, Adamson TC *et al.* 2011 recommendations for the diagnosis and management of gout and hyperuricemia. *Phys Sportsmed* 2011;39:98–123.
- 9 Sivera F, Andrés M, Carmona L *et al.* Multinational evidence-based recommendations for the diagnosis and management of gout: integrating systematic literature review and expert opinion of a broad panel of rheumatologists in the 3e initiative. *Ann Rheum Dis* 2014;73:328–35.
- 10 Grassi W, Meenagh G, Pascual E, Filippucci E. “Crystal clear”: sonographic assessment of gout and calcium pyrophosphate deposition disease. *Semin Arthritis Rheum* 2006;36:197–202.
- 11 Thiele RG, Schlesinger N. Diagnosis of gout by ultrasound. *Rheumatology* 2007;46:1116–21.
- 12 Wright SA, Filippucci E, McVeigh C *et al.* High-resolution ultrasonography of the first metatarsal phalangeal joint in gout: a controlled study. *Ann Rheum Dis* 2007;66:859–64.
- 13 Filippucci E, Meenagh G, Delle Sedie A *et al.* Ultrasound imaging for the rheumatologist XXXVI. Sonographic assessment of the foot in gout patients. *Clin Exp Rheumatol* 2011;29:901–5.
- 14 Rettenbacher T, Ennemoser S, Weirich H *et al.* Diagnostic imaging of gout: comparison of high-resolution US versus conventional X-ray. *Eur Radiol* 2008;18:621–30.
- 15 Filippucci E, Scirè CA, Delle Sedie A *et al.* Ultrasound imaging for the rheumatologist. XXV. Sonographic assessment of the knee in patients with gout and calcium pyrophosphate deposition disease. *Clin Exp Rheumatol* 2010;28:2–5.
- 16 Ottaviani S, Allard A, Bardin T, Richette P. An exploratory ultrasound study of early gout. *Clin Exp Rheumatol* 2011;29:816–21.
- 17 Thiele RG. Role of ultrasound and other advanced imaging in the diagnosis and management of gout. *Curr Rheumatol Rep* 2011;13:146–53.
- 18 Huppertz A, Hermann KG, Diekhoff T *et al.* Systemic staging for urate crystal deposits with dual-energy CT and ultrasound in patients with suspected gout. *Rheumatol Int* 2014;34:763–71.
- 19 Perez-Ruiz F, Martin I, Canteli B. Ultrasonographic measurement of tophi as an outcome measure for chronic gout. *J Rheumatol* 2007;34:1888–93.
- 20 Mathieu S, Pereira B, Couderc M, Soubrier M. Usefulness of ultrasonography in the diagnosis of gout: a meta-analysis. *Ann Rheum Dis* 2013;72:e23.
- 21 Naredo E, Uson J, Jiménez-Palop M *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.
- 22 Gutiérrez M, Di Geso L, Rovisco J *et al.* Ultrasound learning curve in gout: a disease-oriented training program. *Arthritis Care Res* 2013;65:1265–74.
- 23 Peiteado D, De Miguel E, Villalba A *et al.* Value of a short four-joint ultrasound test for gout diagnosis: a pilot study. *Clin Exp Rheumatol* 2012;30:830–7.
- 24 Ottaviani S, Richette P, Allard A, Ora J, Bardin T. Ultrasonography in gout: a case-control study. *Clin Exp Rheumatol* 2012;30:499–504.
- 25 Filippucci E, Gutierrez MG, Georgescu D, Salaffi F, Grassi W. Hyaline cartilage involvement in patients with gout and calcium pyrophosphate deposition disease. An ultrasound study. *Osteoarthritis Cartilage* 2009;17:178–81.
- 26 Thiele RG, Schlesinger N. Ultrasonography shows disappearance of monosodium urate crystal deposition on hyaline cartilage after sustained normouricemia is achieved. *Rheumatol Int* 2010;30:495–503.
- 27 Pineda C, Amezcua-Guerra LM, Solano C *et al.* Joint and tendon subclinical involvement suggestive of gouty arthritis in asymptomatic hyperuricemia: an ultrasound controlled study. *Arthritis Res Ther* 2011;13:R4.



- 28 De Miguel E, Puig JG, Castillo C *et al.* Diagnosis of gout in patients with asymptomatic hyperuricaemia: a pilot ultrasound study. *Ann Rheum Dis* 2012;71:157-8.
- 29 Puig JG, de Miguel E, Castillo MC *et al.* Asymptomatic hyperuricemia: impact of ultrasonography. *Nucleosides Nucleotides Nucleic Acids* 2008;27:592-5.
- 30 Chowalloor PV, Keen HI. A systematic review of ultrasonography in gout and asymptomatic hyperuricemia. *Ann Rheum Dis* 2013;72:638-45.
- 31 Wuensch, KL. "What is a Likert Scale? and How Do You Pronounce 'Likert?'" . East Carolina University, 2009. <http://core.ecu.edu/psyc/wuenschk/StatHelp/Likert.htm> (4 April 2015, date last accessed).
- 32 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
- 33 Gerster JC, Landry M, Dufresne L, Meuwly JY. Imaging of tophaceous gout: computed tomography provides specific images compared with magnetic resonance imaging and ultrasonography. *Ann Rheum Dis* 2002;61:52-4.
- 34 Wakefield RJ, Balint PV, Szkudlarek M *et al.* OMERACT 7 Special Interest Group. Musculoskeletal ultrasound including definitions for ultrasonographic pathology. *J Rheumatol* 2005;32:2485-7.
- 35 Naredo E, D'Agostino MA, Wakefield RJ *et al.* Reliability of a consensus-based ultrasound score for tenosynovitis in rheumatoid arthritis. *Ann Rheum Dis* 2013;72:1328-34.
- 36 Howard RG, Pillinger MH, Gyftopoulos S *et al.* Reproducibility of musculoskeletal ultrasound for determining monosodium urate deposition: concordance between readers. *Arthritis Care Res* 2011;63:1456-62.
- 37 McQueen FM, Reeves Q, Dalbeth N. New insights into an old disease: advances imaging in the diagnosis and management of gout. *Postgrad Med J* 2013;89:87-93.
- 38 Dalbeth N, Doyle A, McQueen FM. Imaging in gout: insights into the pathological features of disease. *Curr Opin Rheumatol* 2012;24:132-8.
- 39 Raine R, Sanderson C, Black N. Developing clinical guidelines: a challenge to current methods. *BMJ* 2005;331:631-3.