



Published in final edited form as:

Arthritis Rheumatol. 2017 February ; 69(2): 429–438. doi:10.1002/art.39959.

Performance of Ultrasound in the Diagnosis of Gout in a Multi-Center Study: Comparison with Monosodium Urate Crystal Analysis as the Gold Standard

Alexis Ogdie, MD, MSCE¹, William J Taylor, MBChB, PhD, FRACP, FAFRM², Tuhina Neogi³, Jaap Fransen, PhD⁴, Tim L Jansen, MD⁴, H. Ralph Schumacher, MD¹, Worawit Louthrenoo, MD⁵, Janitzia Vazquez-Mellado, MD, PhD⁶, Maxim Eliseev, MD⁷, Geraldine McCarthy, MD, FRCPI^{8,9}, Lisa K. Stamp, MD, FRACP¹⁰, Fernando Perez-Ruiz, MD¹¹, Francisca Sivera, MD¹², Hang-Korng Ea, MD^{13,14}, Martijn Gerritsen, MD¹⁵, Giovanni Cagnotto, MD¹⁶, Lorenzo Cavagna, MD¹⁷, Chingtsai Lin, MD¹⁸, Yin-Yi Chou, MD¹⁹, Anne-Kathrin Tausche, MD²⁰, Manuella Lima Gomes Ochtrop, MD²¹, Matthijs Janssen, MD²², Jiunn-Horng Chen, MD, PhD^{23,24}, Ole Slot, MD²⁵, Juris Lazovskis, MD²⁶, Douglas White, MD²⁷, Marco A. Cimmino, MD²⁸, Till Uhlig, MD²⁹, and Nicola Dalbeth, MD, FRACP³⁰

¹Division of Rheumatology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA ²Department of Medicine, University of Otago, Wellington, New Zealand ³Clinical Epidemiology Research & Training Unit, Boston University School of Medicine, Boston MA, USA ⁴Department of Rheumatology, VieCuri Medical Centre, Venlo, The Netherlands & Scientific IQ HealthCare, Radboud University Medical Center, Nijmegen, The Netherlands ⁵Division of Rheumatology, Department of Internal Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand ⁶Servicio de Reumatología, Hospital General de México, México City, México ⁷Nasonova Research Institute of Rheumatology of Russia, Moscow, Russia ⁸School of Medicine and Medical Science, University College Dublin, Ireland ⁹Mater Misericordiae University Hospital, Eccles St, Dublin, Ireland ¹⁰Department of Medicine, University of Otago Christchurch, PO Box 4345, Christchurch 8140, New Zealand ¹¹Rheumatology Division, Hospital Universitario Cruces & BioCruces Health Research Institute, and Basque Country University, Vizcaya, Spain ¹²Department Reumatología, Hospital General Universitario de Elda, Alicante, Spain ¹³Univ. Paris Diderot, Sorbonne Paris Cité, UFR de Médecine, F-75205 Paris, France ; INSERM, UMR 1132, Hôpital Lariboisière, F-75475 Paris, France ¹⁴Assistance Publique-Hôpitaux

Corresponding Author: Alexis Ogdie, Division of Rheumatology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. Phone: 215-615-4375, Fax: 215-662-4500, alexis.ogdie@uphs.upenn.edu.

COMPETING INTERESTS

AO has received consulting fees from Novartis. ND has received consulting fees, speaker fees or grants from the following companies: Takeda, Menarini, Teijin, Pfizer, Crealta, Cymabay, Fonterra, Ardea Biosciences and AstraZeneca, outside the submitted work. WT has received consulting fees from AstraZeneca and Pfizer. FPR has received consulting fees from AstraZeneca, Menarini and Cymabay; speaker fees from AstraZeneca and Menarini; investigation grants from Spanish Health Ministry, Basque Country Industry and Development Department, Spanish Foundation for Rheumatology, and Cruces Hospital Rheumatology Association. LKS has received consulting fees from AstraZeneca. AKT has received fees from Berlin-Chemie Menarini, Andrea Biosciences/AstraZeneca and Novartis. GM has received consulting and speaker fees from Menarini. MAC has received investigation grants and speakers fees from Menarini.

CONTRIBUTOR STATEMENT

All authors assisted in study conception, design and interpretation. Drs. Ogdie performed the data analysis with assistance from Drs. Dalbeth and Taylor. Dr. Ogdie wrote the first draft of the manuscript and all of the authors reviewed and edited the manuscript.

de Paris (AP-HP), Hôpital Lariboisière, Service de Rhumatologie, Centre Viggo Petersen, Pôle Appareil Locomoteur, 2, Rue Ambroise Paré, F-75010 Paris, France ¹⁵Amsterdam Rheumatology Immunology Center (ARC), Department of Rheumatology, Westfries Gasthuis, Hoorn, the Netherlands ¹⁶Division of Rheumatology, University and IRCCS Policlinico S. Matteo Foundation, Pavia, Italy and Department of Rheumatology, Skane University Hospital Malmö/Lund, Lund University, Sweden ¹⁷Division of Rheumatology, University and IRCCS Policlinico S. Matteo Foundation, Pavia, Italy ¹⁸Division of Allergy, Immunology, and Rheumatology, Taichung Veterans' General Hospital, Taichung, Taiwan ¹⁹Taichung Veterans' General Hospital, 160 Taichung Road Section 3, Taichung, Taiwan 40705, ROC ²⁰Division of Rheumatology, Department of Internal Medicine III, University Hospital Carl Gustav Carus, Fetscherstrasse 74, D-01307 Dresden, Germany ²¹Division of Rheumatology, Department of Internal Medicine, Universidade do Estado do Rio de Janeiro, Brasil ²²Rijnstate Hospital, Arnhem, The Netherlands ²³School of Medicine, China Medical University, Taichung, Taiwan, ROC ²⁴Department of Internal Medicine, China Medical University Hospital, Taichung, Taiwan, ROC ²⁵Centre of Rheumatology and Spine Disorders, Copenhagen University Hospital Glostrup DK-2600, Glostrup, Denmark ²⁶Rheumatology Clinic, Riverside Professional Centre, Sydney, NS, Canada ²⁷Waikato DHB and Waikato Clinical School, University of Auckland, Hamilton, New Zealand ²⁸Research Laboratory and Academic Division of Clinical Rheumatology, Department of Internal Medicine, University of Genova, Italy ²⁹National Advisory Unit on Rehabilitation in Rheumatology, Department of Rheumatology, Diakonhjemmet Hospital, Box 23 Vinderen, 0319 Oslo, Norway ³⁰Department of Medicine, University of Auckland, Auckland, New Zealand

Abstract

Objectives—To examine the performance of ultrasound for the diagnosis of gout using presence of monosodium urate (MSU) crystals as the gold standard.

Methods—We analyzed data from the Study for Updated Gout Classification Criteria (SUGAR), a large, multi-center observational cross-sectional study of consecutive subjects with at least one swollen joint who conceivably may have gout. All subjects underwent arthrocentesis; cases were subjects with MSU crystal confirmation. Rheumatologists or radiologists, blinded to the results of the MSU crystal analysis, performed ultrasound on one or more clinically affected joints. Ultrasound findings of interest were: double contour sign (DCS), tophus, and 'snowstorm' appearance. Sensitivity, specificity, positive and negative predictive values (PPV and NPV) were calculated. Multivariable logistic regression models were used to examine factors associated with positive ultrasound results among subjects with gout.

Results—Ultrasound was performed in 824 subjects (416 cases and 408 controls). The sensitivity, specificity, PPV and NPV for the presence of any one of the features were 76.9%, 84.3%, 83.3% and 78.1% respectively. Sensitivity was higher among subjects with disease ≥ 2 years duration and among subjects with subcutaneous nodules on exam (suspected tophus). Associations with a positive ultrasound finding included suspected clinical tophus (odds ratio 4.77; 95% CI 2.23–10.21), any abnormal plain film radiograph (4.68; 2.68–8.17) and serum urate (1.31; 1.06–1.62).

Conclusions—Ultrasound features of MSU crystal deposition had high specificity and high positive predictive value but more limited sensitivity for early gout. The specificity remained high in subjects with early disease and without clinical signs of tophi.

Keywords

gout; ultrasound; imaging; sensitivity; specificity; pseudogout

While monosodium urate (MSU) crystals in the synovial fluid are the gold standard for the diagnosis of gout, musculoskeletal ultrasound is increasingly used in studies of gout and for establishing a diagnosis of gout in clinical practice.[1] MSU crystal deposition can be identified on ultrasound by hyperechoic irregular enhancement of the articular surface of the hyaline cartilage (double contour sign, DCS), hyperechoic aggregates suggestive of tophi within the joint or along tendons, and floating hyperechoic foci within the joint space which have the appearance of a snowstorm.[1] Previous studies have suggested reasonable test characteristics of these features, DCS and tophi in particular.[2–8] However, most studies examining the diagnostic properties of ultrasound in gout have been single center studies of relatively small size, with a few exceptions, and most of the subjects enrolled in these studies have had long-standing disease. Additionally, most previous studies have involved centers with musculoskeletal ultrasound world experts, which may not reflect the accuracy of ultrasound in real-life clinical practice. Finally, many existing studies are limited by suboptimal characterization of the control groups (i.e., arthrocentesis was not performed in subjects in the comparator groups).[4]

In the Study for Updated Gout Classification Criteria (SUGAR), a large, international, multi-center cross-sectional study, multivariate logistic regression analysis showed that US findings contributed independently to identification of gout with an odds ratio of 7.2. These results suggested that US provides a useful contribution to discrimination between gout and non-gout and this was at the same level as clinical identification of tophi.[9] The objectives of the current study were to determine the sensitivity and specificity of ultrasound features for the diagnosis of gout in this international cohort and better understand the test characteristics in early versus late disease and among those with versus without clinically suspected tophus on examination. The primary goal was to determine how ultrasound performs in the clinical practice setting. Thus, a scanning protocol and study-specific training were not provided. Additionally, we sought to determine subject and disease characteristics associated with a positive ultrasound scan.

Patients and Methods

Study design and setting

We analyzed data from SUGAR, a large, multi-center observational cross-sectional study designed to inform new gout classification criteria.[9] Standardized ultrasound analysis was included in the data collection, but it was not mandatory.

Study population and gold standard

Consecutive subjects from rheumatology clinics with at least one swollen joint or a subcutaneous nodule who conceivably may have had gout (i.e. gout was on the differential diagnosis for the patient's joint swelling) were enrolled.[9] All subjects underwent arthrocentesis or soft tissue nodule aspiration to establish a crystal-based diagnosis. Twenty-five centers from around the world participated. Within the SUGAR cohort, subjects who underwent an ultrasound examination were included in the present analysis. The gold standard test was the presence of monosodium urate crystals in the joint fluid. Crystal identification was performed by trained observers who were required to pass a certification procedure, which included a web-based crystal recognition test and the examination of 5 vials of synovial fluid. This was described in more detail in Taylor et al.[9] Only sites with participants who completed this certification were able to participate in the study. Cases were subjects with MSU crystal confirmation and controls were subjects with a joint fluid or soft tissue nodule aspirate that was negative for MSU crystals. In a sensitivity analysis, those with clinical suspected gout but MSU negative were also included in the "cases" to better understand the impact of restricting the cases to those with MSU confirmation.

Ultrasound

Ultrasound was performed on one or more clinically affected joints by either rheumatologists or radiologists blinded to the aspiration results. All ultrasonographers had prior US training. Ultrasound DCS was defined as a hyperechoic band on the surface of the articular cartilage.[10] Ultrasound tophus was defined as the presence of a hyperechoic, heterogeneous lesion surrounded by an anechoic rim.[10, 11] Ultrasound snowstorm was defined as a "snowstorm type joint effusion." [12] These definitions were provided in the clinical research form. Ultrasonographer training and equipment used is provided in Supplemental Table 1.

Clinical data and covariates of interest

Data collected at the time of enrollment are reported in detail in the SUGAR study.[9] Covariates included age, gender, ethnicity, number of episodes of joint swelling, use of urate lowering therapy, weeks since first episode, disease duration, and history of podagra (first metatarsophalangeal joint involvement). A physical examination was performed; location of tender and swollen joints, a count of involved joints and presence of suspected tophus on examination were recorded. Current and highest ever (from any available records) serum urate (SUA) and assessment of radiographic abnormalities were also recorded.

Test Characteristics for Gout compared with CPPD

Because a recent study suggested difficulty differentiating gout from calcium pyrophosphate deposition (CPPD) disease on ultrasound [13], we additionally examined specificity when ultrasound features were compared among subjects with MSU-positive gout and those with CPPD (sensitivity would be the same as in the previous analyses). CPPD was defined by the presence of CPP crystals in the synovial fluid. Those with clinically suspected gout, but who were MSU-negative, were not included in this analysis.

Statistical analyses

Demographic data were descriptively reported. Differences between those with and without an ultrasound examination in the SUGAR cohort were examined using the Chi-square test or the Wilcoxon Rank Sum test as appropriate. Among subjects with ultrasound, we examined sensitivity, specificity, positive and negative predictive values of the three features individually (DCS, tophus, ‘snowstorm’) or the presence versus the absence of any one of the features in all subjects with gout versus controls and then separately within subgroups defined by two factors: early disease (defined as less than 2 years since first symptom onset/gout attack) and late disease, subjects with and without clinically suspected tophus on examination.[14] We additionally examined the test characteristics of having one, two or all three of these features present. Sensitivity was calculated as [true positives] divided by [true positives plus false negatives]. Specificity was calculated by dividing [true negatives] by [false positives plus true negatives]. Positive predictive value (PPV) was calculated as [true positives] divided by [true positives plus false positives] and negative predictive value (NPV) was calculated as [true negatives] divided by [true negatives plus false negatives].

Finally, we examined factors associated with having a positive ultrasound finding among cases. The goal of this analysis was to gain insight into what factors may influence the test results. We first performed univariable logistic regression models to examine association between single factors and the presence of a positive test (among cases with gout only). Age, sex, and all variables that were significant at the univariable level ($p < 0.1$) were placed into a multivariable model. Covariates with the highest p-value were then removed one by one until only significant ($p < 0.05$) covariates remained in the model. We then retested addition of each variable and left it in the model had a significant p-value (< 0.05). Age and sex were retained in all models given their biologic significance.

Sensitivity Analyses—We examined the impact of changing the definition of gout to a physicians’ diagnosis of gout rather than MSU positive gout. Next, we examined stratified analysis by years of experience using MSK US (dichotomized as less than 10 years or more than 10 years). Finally, we examined the test characteristics among patients with or without a tender or swollen first MTP joint as the presence of MTP symptoms may have influenced the ultrasonographer. Finally, we examined subgroups of men with hyperuricemia to determine whether subclinical hyperuricemia may have led to more “false positives” in patients without clinical gout.

Results

Baseline characteristics

Among the 982 subjects enrolled within SUGAR, ultrasound was performed among 824 (416 cases and 408 controls). The baseline characteristics (Table 1) of the ultrasound group were reflective of the larger SUGAR population with a few exceptions. Subject characteristics among those with and without an ultrasound examination are shown in Supplemental Table 2. Relatively fewer subjects of Caucasian and Black or African ethnicity had an ultrasound performed (whereas the majority of Hispanics, South Asians and East Asians had an ultrasound performed). The mean SUA was slightly lower in the ultrasound

group (0.41 versus 0.45, $p=0.001$) and more patients who received an ultrasound had higher prevalence of tender (72% vs 63%, $p=0.02$) and swollen (73% vs 63%, $p=0.01$) joints proximal to the ankle.

Among subjects included in the ultrasound analysis, mean age was 60.2 (SD 14.6) and 59.5 (SD 16.0) for cases and controls respectively and 87% and 54% were male respectively (Table 1). Twenty-six percent of MSU-positive gout subjects had <2 years disease duration whereas 48% of controls had <2 years of disease duration. The majority of subjects had more than five episodes of joint “flares” (69% and 49% of cases and controls respectively). The diagnoses of the control subjects were CPPD (N=98), clinically suspected gout (N=41), osteoarthritis (N=63), rheumatoid arthritis (N=59), spondyloarthropathies (N=60), undifferentiated inflammatory arthritis (N=47), septic arthritis (N=7), systemic lupus erythematosus (N=5), and other (N=28). US was performed for a single joint for most patients, however, it was performed for more than one joint in 16% of patients. The most commonly examined joints were the knees, metatarsophalangeal joints and ankles (Supplemental Table 3).

Sensitivity, Specificity, Positive and Negative Predictive Values of Ultrasound in MSU-positive Gout

In Table 2, we report the sensitivities, specificities and positive and negative predictive values for ultrasound features. Among all subjects, sensitivity for DCS was 60.1% (55.2–64.9), specificity was 91.4% (88.3–94.0), PPV 87.7% (83.3–91.3), and NPV 69.3% (65.2–73.2). Sensitivity for tophus on ultrasound was 46.0% (41.1–50.9), specificity was 94.9% (92.2–96.8), PPV 90.0% (85.1–93.7) and NPV 65.6% (59.6–67.4). The sensitivity for snowstorm was 30.3% (25.9–35.0), specificity was 90.9% (87.7–93.5), and PPV 77.2% (69.9–83.4), and NPV 56.3% (52.4–60.1). When the presence of any one of the three features was considered a “positive” test, the sensitivity was 76.9% (72.6–80.9), specificity was 84.3% (80.4–87.7), PPV was 83.3% (79.2–86.9), and NPV was 78.2 (74.0–82.0). We additionally examined the test characteristics of the ultrasound features among the subsets of subjects with early disease (defined as 2 years or fewer of symptom duration: 109 cases and 195 controls) versus late disease (304 cases and 209 controls), and of subjects in whom there were no suspected tophus (or subcutaneous nodules; 265 cases and 389 controls) present on examination versus those with suspected tophus on examination (150 cases and 19 controls). Overall, sensitivity and specificity were not substantially different in the subgroups except among patients with suspected tophus on examination. Sensitivity and PPV were higher and specificity and NPV were lower among subjects with suspected tophus (see Table 2). In a sensitivity analysis, we included subjects in the “gout” group that were clinically diagnosed with gout but had an aspirate that was negative for MSU crystals (N=41); this did not substantially change the results except among patients with clinical suspected tophus on examination, the specificity improved for all three features and the composite outcome (Supplemental Table 4). We additionally examined the test characteristics of having one, two or three features of gout on US (Table 3 and Figure 1). Sensitivity and NPV decreased as the number of features of gout increased and specificity for gout increased as the number of features increased. The PPV remained relatively high.

In sensitivity analyses, we examined whether patient or sonographer characteristics affected the test characteristics for DCS, tophus, snowstorm or having any positive feature (Supplemental Table 5). There were not substantial differences among the test characteristics when stratifying by years of experience of the sonographer (<10 years compared to 10 years) except for a slightly lower NPV in patients examined by sonographers with more training. An actively swollen or tender first MTP was associated with a lower sensitivity and lower NPV for DCS, Tophus on US, and Snowstorm. This was particularly true for snowstorm. Similarly, having an elevated SUA was associated with lower specificity and lower NPV. Finally, we examined the test characteristics after stratifying by disease durations (split into 5 categories; Supplemental Figure 1). Overall, specificity was relatively unchanged. Sensitivity of DCS and tophus on US increased with increasing duration of disease.

Performance of ultrasound features among subjects with gout and CPPD

We examined the specificity of the same ultrasound features for gout in subjects with CPPD-positive aspirates. (Sensitivity is the same as this is calculated from the cases). When compared with subjects with CPPD, ultrasound features still had a relatively high specificity (88–96%) and positive predictive value (93–97%) for gout (Table 4). Of note, subjects with CPPD were more likely to have less than two years of symptoms (52.8% of those with CPPD, 26.4% of those with gout, $p<0.001$).

Associations with a positive test among cases

Finally, we sought to determine which patients with gout were most likely to have a positive ultrasound. Associations with positive ultrasound findings among cases are shown in Table 5. After adding age and sex, the subject characteristics most strongly associated with a DCS were the highest serum SUA ever, joint tenderness proximal to the ankle, and any radiographic feature of gout (erosion, asymmetric swelling, or cyst). Suspected tophus on examination, current SUA, and asymmetrical swelling on plain film radiographs were significantly associated with identification of tophus on ultrasound. Suspected tophus on examination and cysts on plain film radiographs were positively associated with snowstorm appearance on ultrasound whereas the number of episodes was negatively associated with snowstorm. Finally, suspected tophus on examination, current SUA and having any radiographic abnormality were strongly associated with having any ultrasound feature of gout.

Discussion

We found high specificity (>90% for individual features) and positive predictive value for ultrasound of clinically symptomatic joints when using MSU-positivity as the gold standard. However, sensitivity was better for those with long standing disease (> 2 years of symptoms or tophi) than those with early disease or without suspected tophi on examination. Thus, the absence of one of these features does not rule out the disease. Additionally, the ultrasound features of DCS and tophi had high specificity for gout even when compared with those with CPPD, suggesting that these features still allow for differentiation between gout and CPPD. A positive ultrasound finding indicative of gout among subjects with MSU-positive gout was

associated with suspected tophi on examination, higher SUA, and abnormal radiographic findings. Thus, a positive ultrasound may be associated with a higher urate burden in general. Overall, these results suggest that, while US performs best in patients who would likely be more easily diagnosed without imaging (e.g., patients with tophi), US is still useful in patients without obvious gout from examination (e.g., without tophi, without MTP tenderness or swelling) and specificity remains high (~90%) in these subgroups.

In a previous meta-analysis, we summarized nine manuscripts and two meeting abstracts examining the sensitivity and specificity of ultrasound for the diagnosis of gout compared with MSU crystals as the gold standard.[4] In the meta-analysis, the ultrasound features of DCS and tophi had pooled sensitivity of 83% and 65%, respectively, and pooled specificity of 76% and 80%, respectively (there were insufficient data on snowstorm to generate a pooled sensitivity and specificity). Compared with the pooled test characteristics in the meta-analysis, sensitivity was lower and specificity was higher in the current study. This may be related to differences in the subject characteristics in our study compared to previously reported studies. Sensitivity and specificity, while test characteristics, are tied to the population studied and the reference standard used. The current study included more subjects with early disease relative to previous studies. Additionally, our study enrolled subjects who had a clinical presentation that could be consistent with gout. Arguably, this is the population in which ultrasound is most useful as a diagnostic test. Many of the studies included in this meta-analysis used comparator groups that were less likely to have a diagnosis of gout and may have been less likely to have active joint disease. A recent single-center examination of the test characteristics of ultrasound in a similar population to those enrolled in our study reported sensitivity of 74% and specificity of 82% for any ultrasound feature.[15] This was similar to our findings for any ultrasound feature (sensitivity 77% and specificity 84%).

Because sensitivity and specificity are tied to the populations studied, one of the strengths of our study was the ability to look at these test characteristics among subgroups. We found a higher sensitivity for those with disease duration >2 years but similar specificity among these groups. Ultrasound had a higher sensitivity among subjects with suspected tophus on examination (compared with those without), but much lower specificity. There were a relatively larger number of “positive” ultrasounds among subjects with suspected tophus on examination but aspirations negative for MSU crystals (8 of 19 subjects) and specificity is calculated as the number of true negatives over the number of subjects without gout. These subjects may in fact have had gout but the “active” joint was related to a different diagnosis or the MSU crystals were missed among the certified assessors. As noted in the SUGAR study, suspected tophus on examination may be a stronger predictor of existing gout than ultrasound features.[9] Furthermore, when we reclassified patients according to clinician diagnosis, the specificity increased substantially in the subgroup of patients with suspected tophus on examination. In addition to the two previous mentioned subgroups, we also compared the test characteristics of ultrasound among those with MSU-positive gout with the subgroup of controls with CPPD. CPPD is associated with similar ultrasound findings. [12, 13, 16–19] In our study, the specificity of US for gout remained >80%. We found relatively few subjects with CPPD and ultrasound changes attributed to urate deposition. A recent meta-analysis by Gamon, et al. found that sensitivity and specificity of ultrasound in

CPPD varies by the joint or tendon analyzed.[17] It may be that joints specific for gout (in general lower extremities) have lower sensitivity for CPPD. Alternatively, it may be that the subjects with CPPD in this study may have had fewer positive tests related to shorter disease duration.

Our study is a unique contribution to the literature in that it is an examination of a real-life, multi-center experience of ultrasound in the assessment of gout features. Additionally, this study examined the performance characteristics of DCS, tophus, and snowstorm. Most previous studies have mainly focused on DCS and tophus with very few examining test characteristics of snowstorm. Finally, we used MSU crystal analysis as the gold standard when calculating sensitivity and specificity, which included confirming absence of MSU crystals in controls and we required certification for MSU crystal identified among all investigators. The use of a rigorous gold standard is of critical importance when determining the sensitivity and specificity of a diagnostic test.

Limitations of this study include possible selection bias, variation in ultrasonographer training and ultrasound machine use, and possible test interpretation bias. Ultrasound was not performed among all subjects in the SUGAR study due to availability of ultrasound and trained ultrasonographers at the enrolling sites. However, there were few differences in the subjects who did versus did not undergo ultrasonography. This suggests that there was not significant selection bias in which subjects underwent ultrasound. Next, a variety of machines were used and many different ultrasonographers performed the ultrasounds. Ultrasonographers were mainly rheumatologists who used ultrasound in clinical practice although not necessarily certified or radiologists. Although definitions of US features were provided to all ultrasonographers, a standardized scanning protocol was not required. Inter-rater reliability was not assessed; this has been reported [6, 20, 21] and this was not the primary goal of the study. There was some variability in the false positive and false negative rates at the individual sites (data not shown). We did not have the ability to centrally re-read ultrasound images. However, this reflects “real world” use of ultrasound in clinical practice, increasing the external validity of the results. Understanding ultrasound performance in the “real world” was the primary objective of this study (acknowledging that this is still under “study” condition). Next, only clinically affected joints were scanned. Inclusion of additional asymptomatic joints may have increased the sensitivity. However, our primary goal was to assess the ability of ultrasound to assist in diagnosing gout in the symptomatic joint. Test interpretation bias is possible, although should not have a significant influence given that ultrasound was performed blinded to synovial fluid analysis. Importantly, ultrasonographers may not have been blinded to all clinical features, for example, it is possible that the presence of visible tophi or other clinically apparent characteristics influenced the interpretation of the ultrasound results. However, this also reflects real-life clinical practice, in which ultrasound is used as an additive test to available clinical data. Finally, it is important to recognize that these data are relevant for patients with symptomatic joint swelling. These results cannot be applied to diagnosis of gout in patients with asymptomatic hyperuricemia.

In conclusion, musculoskeletal ultrasound had high specificity for the diagnosis of gout among subjects with at least one swollen joint and a clinical presentation that could be

consistent with gout, but sensitivity was modest. The specificity remained high for early disease and without clinical signs of tophi, the population in which ultrasound may be the most useful in establishing a diagnosis of gout.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

FUNDING STATEMENT

This study was supported by the American College of Rheumatology, European League against Rheumatism, Arthritis New Zealand, Association Rhumatisme et Travail, and Asociación de Reumatólogos del Hospital de Cruces. Dr. Ogdie is supported by NIH K23AR063764. Dr Dalbeth is supported by the Health Research Council of New Zealand. Dr. Tuhina Neogi is supported by NIH P60 AR47785 and R01 AR062506. Dr. William Taylor is supported by Arthritis New Zealand. Dr. Jasvinder Singh is supported by NIAMS, NIA, NCI, and AHRQ CERTs.

We gratefully acknowledge the help of Joung-Liang Lan, Chien-Chung Huang, Po-Hao Huang, Hui-Ju Lin and Su-Ting Chang (China Medical University Hospital, Taiwan), Anne Madigan (Dublin, Ireland), Yi-hsing Chen (Taichung, Taiwan), Carina Soto Fajardo and Lucio Ventura Ríos (México City, México), Viktoria Fana (Copenhagen, Denmark), Panomkorn Lhakum and Kanon Jatuworapruk (Chiang Mai, Thailand), Dianne Berendsen (Nijmegen, Netherlands), Femke Lamers-Karnebeek (Amsterdam, Netherlands), Olivier Peyr (Paris, France), Heidi Lunøe and Anne Katrine Kongtorp (Oslo, Norway), Geraldo da Rocha Castelar-Pinheiro (Rio de Janeiro, Brasil), Fatima Kudaeva (Moscow, Russia), Angelo Gaffo (Birmingham AL), Douglas White (Hamilton, New Zealand), Carlo Scirè (Milan, Italy), Maria Winzer (Dresden, Germany) and Juris Lazovskis (Sydney, Canada) with data collection, crystal examination or patient referral. We are grateful to Eliseo Pascual (Alicante, Spain) for help with MSU observer certification. Finally, we thank Yihui Connie Jiang, BA, and Adam Marc, BA, for administrative assistance.

References

1. Grassi W, Okano T, Filippucci E. Use of ultrasound for diagnosis and monitoring of outcomes in crystal arthropathies. *Current opinion in rheumatology*. 2015; 27(2):147–155. [PubMed: 25633243]
2. Chowalloor PV, Keen HI. A systematic review of ultrasonography in gout and asymptomatic hyperuricaemia. *Annals of the rheumatic diseases*. 2013; 72(5):638–645. [PubMed: 23291387]
3. Mathieu S, Pereira B, Couderc M, Soubrier M. Usefulness of ultrasonography in the diagnosis of gout: a meta-analysis. *Annals of the rheumatic diseases*. 2013; 72(10):e23. [PubMed: 23852689]
4. Ogdie A, Taylor WJ, Weatherall M, Fransen J, Jansen TL, Neogi T, Schumacher HR, Dalbeth N. Imaging modalities for the classification of gout: systematic literature review and meta-analysis. *Annals of the rheumatic diseases*. 2015; 74(10):1868–1874. [PubMed: 24915980]
5. Lamers-Karnebeek FB, Van Riel PL, Jansen TL. Additive value for ultrasonographic signal in a screening algorithm for patients presenting with acute mono-/oligoarthritis in whom gout is suspected. *Clinical rheumatology*. 2014; 33(4):555–559. [PubMed: 24510062]
6. Naredo E, Uson J, Jimenez-Palop M, Martinez A, Vicente E, Brito E, Rodriguez A, Cornejo FJ, Castaneda S, Martinez MJ, et al. Ultrasound-detected musculoskeletal urate crystal deposition: which joints and what findings should be assessed for diagnosing gout? *Annals of the rheumatic diseases*. 2014; 73(8):1522–1528. [PubMed: 23709244]
7. Ottaviani S, Richette P, Allard A, Ora J, Bardin T. Ultrasonography in gout: a case-control study. *Clinical and experimental rheumatology*. 2012; 30(4):499–504. [PubMed: 22512867]
8. Thiele RG, Schlesinger N. Diagnosis of gout by ultrasound. *Rheumatology (Oxford, England)*. 2007; 46(7):1116–1121.
9. Taylor WJ, Fransen J, Jansen TL, Dalbeth N, Schumacher HR, Brown M, Louthrenoo W, Vazquez-Mellado J, Eliseev M, McCarthy G, et al. Study for Updated Gout Classification Criteria (SUGAR): identification of features to classify gout. *Arthritis care & research*. 2015

10. Terslev L, Gutierrez M, Schmidt WA, Keen HI, Filippucci E, Kane D, Thiele R, Kaeley G, Balint P, Mandl P, et al. Ultrasound as an Outcome Measure in Gout. A Validation Process by the OMERACT Ultrasound Working Group. *The Journal of rheumatology*. 2015; 42(11):2177–2181. [PubMed: 26329333]
11. Thiele RG. Role of ultrasound and other advanced imaging in the diagnosis and management of gout. *Current rheumatology reports*. 2011; 13(2):146–153. [PubMed: 21210257]
12. Grassi W, Meenagh G, Pascual E, Filippucci E. “Crystal clear”-sonographic assessment of gout and calcium pyrophosphate deposition disease. *Seminars in arthritis and rheumatism*. 2006; 36(3): 197–202. [PubMed: 17011611]
13. Loffler C, Sattler H, Peters L, Loffler U, Uppenkamp M, Bergner R. Distinguishing gouty arthritis from calcium pyrophosphate disease and other arthritides. *The Journal of rheumatology*. 2015; 42(3):513–520. [PubMed: 25399385]
14. Taylor WJ, Franssen J, Dalbeth N, Neogi T, Schumacher HR, Brown M, Louthrenoo W, Vazquez-Mellado J, Eliseev M, McCarthy G, et al. Performance of classification criteria for gout in early and established disease. *Annals of the rheumatic diseases*. 2016; 75(1):178–182. [PubMed: 25351521]
15. Zufferey P, Valcov R, Fabreguet I, Dumusc A, Omoumi P, So A. A prospective evaluation of ultrasound as a diagnostic tool in acute microcrystalline arthritis. *Arthritis research & therapy*. 2015; 17:188. [PubMed: 26198435]
16. Frediani B, Filippou G, Falsetti P, Lorenzini S, Baldi F, Acciai C, Siagkri C, Marotto D, Galeazzi M, Marcolongo R. Diagnosis of calcium pyrophosphate dihydrate crystal deposition disease: ultrasonographic criteria proposed. *Annals of the rheumatic diseases*. 2005; 64(4):638–640. [PubMed: 15769921]
17. Gamon E, Combe B, Barnette T, Mouterde G. Diagnostic value of ultrasound in calcium pyrophosphate deposition disease: a systematic review and meta-analysis. *RMD open*. 2015; 1(1):e000118. [PubMed: 26535143]
18. Ottaviani S, Juge PA, Aubrun A, Palazzo E, Dieude P. Sensitivity and Reproducibility of Ultrasonography in Calcium Pyrophosphate Crystal Deposition in Knee Cartilage: A Cross-sectional Study. *The Journal of rheumatology*. 2015; 42(8):1511–1513. [PubMed: 26077407]
19. Ruta S, Catay E, Marin J, Rosa J, Garcia-Monaco R, Soriano ER. Knee effusion: ultrasound as a useful tool for the detection of calcium pyrophosphate crystals. *Clinical rheumatology*. 2015
20. Gutierrez M, Schmidt WA, Thiele RG, Keen HI, Kaeley GS, Naredo E, Iagnocco A, Bruyn GA, Balint PV, Filippucci E, et al. International Consensus for ultrasound lesions in gout: results of Delphi process and web-reliability exercise. *Rheumatology (Oxford, England)*. 2015; 54(10): 1797–1805.
21. Terslev L, Gutierrez M, Christensen R, Balint PV, Bruyn GA, Delle Sedie A, Filippucci E, Garrido J, Hammer HB, Iagnocco A, et al. Assessing Elementary Lesions in Gout by Ultrasound: Results of an OMERACT Patient-based Agreement and Reliability Exercise. *The Journal of rheumatology*. 2015; 42(11):2149–2154. [PubMed: 26472419]

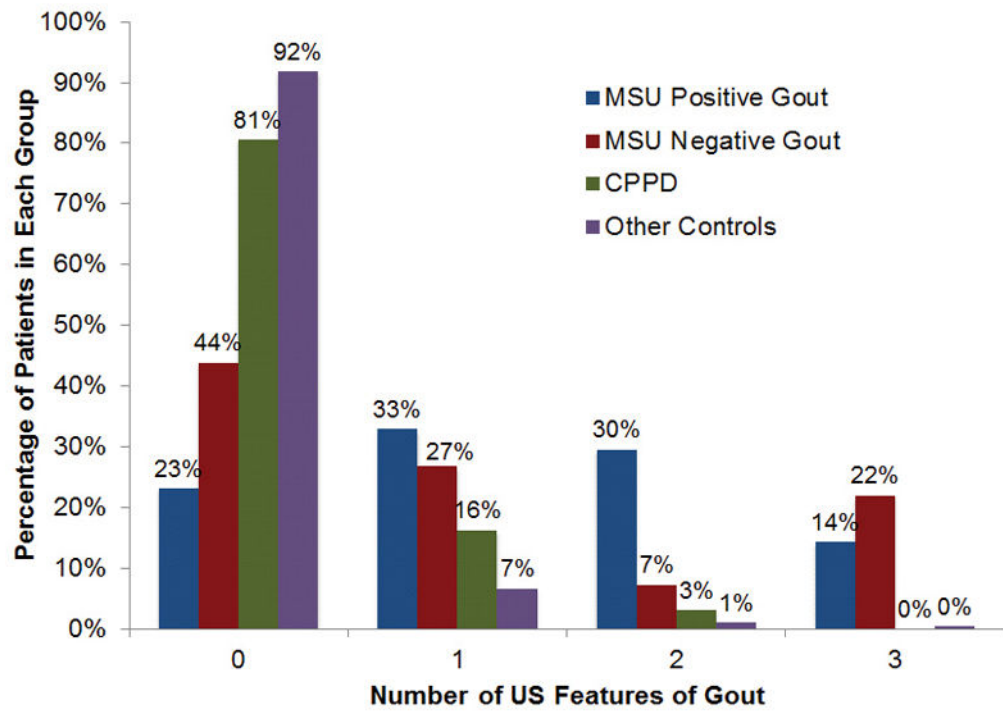


Figure 1. Percentage of patients with one or more ultrasound features of gout. The majority of patients with MSU positive gout (77%) had at least one positive ultrasound feature of gout whereas the majority of controls without CPPD or a clinical diagnosis of gout (92%) had a negative ultrasound.

Table 1

Subject Demographics

	Cases (N=416)	Controls* (N=408)
Age (yrs) Mean (SD)	60.2 (14.6)	59.5 (16.0)
Male sex	363 (87%)	222 (54%)
Ethnicity		
White/European/Caucasian	272 (65%)	221 (54%)
African/Black	6 (1%)	8 (2%)
Hispanic	21 (5%)	19 (5%)
South Asian	40 (10%)	35 (9%)
East Asian	65 (16%)	111 (27%)
Pacific Island	3 (0.7%)	1 (0.3%)
Other indigenous	3 (0.7%)	4 (1%)
Other	6 (1%)	9 (2%)
Number of episodes		
1	36 (9%)	93 (23%)
2 to 5	92 (22%)	116 (28%)
>5	288 (69%)	199 (49%)
Previous diagnosis of gout	345 (83%)	115(28%)
Current urate-lowering therapy	147 (35%)	35 (9%)
Early disease (<2 yrs symptoms)	109 (26%)	195 (48%)
MTP1 ever involved	302 (73%)	109 (27%)
MTP1 currently swollen or tender	150 (36%)	48 (12%)
Suspected clinical tophus	150 (36%)	19 (5%)
Tender joint count	3.27 (4.73)	2.58 (3.99)
Swollen joint count	2.87 (3.66)	2.12 (2.96)
Mean current SUA (mmol/l) Mean (SD)	0.466 (0.138)	0.335 (0.122)
Mean highest ever SUA (mmol/l) Mean (SD)	0.563 (0.140)	0.380 (0.143)

*SUA = serum urate; Highest ever SUA was the highest SUA value from available records.

Table 2
Sensitivity, Specificity, Positive and Negative Predictive Values of Ultrasound Features for Gout

	Cases	Controls	Sensitivity (95% CI*)	p-value	Specificity (95% CI*)	p-value	PPV (95% CI*)	p-value	NPV (95% CI*)	p-value
Ultrasound: DCS										
All subjects	249/414 (60%)	35/408 (9%)	60.1% (55.2–64.9)	N/A	91.4% (88.3–94.0)	N/A	87.7% (83.3–91.3)	N/A	69.3% (65.2–73.2)	N/A
Early disease (<2 yrs)	55/108 (51%)	15/195 (8%)	50.9% (41.1–60.7)	0.02	92.3% (87.6–95.6)	0.73	78.6% (67.1–87.5)	0.004	77.3% (71.3–82.5)	<0.001
Late disease (2 yrs)	192/303 (63%)	18/209 (9%)	63.4% (57.7–68.8)		91.4% (86.7–94.8)		91.4% (86.8–94.8)		63.3% (57.5–68.7)	
No suspected clinical tophus**	141/264 (53%)	27/389 (7%)	53.4% (47.2–59.5)	<0.001	93.1% (90.1–95.4)	<0.001	83.9% (77.5–89.1)	0.022	74.6% (70.5–78.5)	<0.001
Suspected clinical tophus**	107/149 (72%)	8/19 (42%)	71.8% (63.9–78.9)		57.9% (33.5–79.7)		93.0% (86.8–96.9)		20.8% (10.8–34.1)	
Ultrasound: Tophus										
All subjects	189/411 (46%)	21/408 (5%)	46.0% (41.1–50.9)	N/A	94.9% (92.2–96.8)	N/A	90.0% (85.1–93.7)	N/A	65.6% (59.6–67.4)	N/A
Early disease (<2 yrs)	36/107 (34%)	9/195 (5%)	33.6% (25.0–43.4)	0.003	95.4% (91.4–97.9)	0.94	80.0% (65.4–90.4)	0.004	72.4% (66.5–77.7)	<0.001
Late disease (2 yrs)	152/301 (51%)	10/209 (5%)	50.5% (44.7–56.3)		95.2% (91.4–97.7)		93.8% (88.9–97.0)		57.2% (51.8–62.4)	
No suspected clinical tophus**	77/262 (29%)	13/389 (3%)	29.4% (23.9–35.3)	<0.001	96.7% (94.4–98.2)	<0.001	85.6% (76.6–92.1)	0.063	67.0% (63.0–70.9)	<0.001
Suspected clinical tophus**	112/148 (76%)	8/19 (42%)	75.7% (67.9–82.3)		57.9% (33.5–79.7)		93.3% (87.3–97.1)		23.4% (12.3–38.0)	
Ultrasound: Snowstorm										
All subjects	125/412 (30%)	37/407 (9%)	30.3% (25.9–35.0)	N/A	90.9% (87.7–93.5)	N/A	77.2% (69.9–83.4)	N/A	56.3% (52.4–60.1)	N/A
Early disease (<2 yrs)	35/108 (32%)	15/195 (8%)	32.4% (23.7–42.1)	0.58	92.3% (87.6–95.6)	0.49	70.0% (55.4–82.1)	0.100	71.2% (65.1–76.6)	<0.001
Late disease (2 yrs)	89/301 (30%)	20/208 (10%)	29.7% (24.5–35.1)		90.4% (85.5–94.2)		81.7% (73.1–88.4)		47.0% (42.0–52.0)	
No suspected clinical tophus**	63/262 (24%)	29/388 (7%)	24.1% (19.0–29.7)	<0.001	92.5% (89.4–94.9)	<0.001	68.5% (58.0–77.8)	0.003	64.3% (60.2–68.3)	<0.001
Suspected clinical tophus**	62/142 (42%)	8/19 (42%)	41.6% (33.6–50.0)		57.9% (33.5–79.7)		88.6% (78.7–94.9)		11.2% (5.7–19.2)	
Ultrasound: Any feature										
All subjects	320/416 (77%)	64/408 (16%)	76.9% (72.6–80.9)	N/A	84.3% (80.4–87.7)	N/A	83.3% (79.2–86.9)	N/A	78.2% (74.0–82.0)	N/A
Early disease (<2 yrs)	78/109 (72%)	31/195 (16%)	71.6% (62.1–79.8)	0.13	84.1% (78.2–88.9)	0.77	71.6% (62.1–79.8)	<0.001	84.1% (78.2–88.9)	0.006
Late disease (2 yrs)	239/304 (79%)	31/209 (15%)	78.6% (73.6–83.1)		85.2% (79.6–89.7)		88.5% (84.1–92.1)		73.3% (67.2–78.7)	
No suspected clinical tophus**	182/265 (69%)	55/389 (14%)	68.7% (62.7–74.2)	<0.001	85.9% (82.0–89.2)	<0.001	76.8% (70.9–82.0)	<0.001	80.1% (75.9–83.8)	<0.001

	Cases	Controls	Sensitivity (95% CI)*	p-value	Specificity (95% CI)*	p-value	PPV (95% CI)*	p-value	NPV (95% CI)*	p-value
Suspected clinical tophus**	137/150 (91%)	9/19 (48%)	91.3% (85.6–95.3)		52.6% (28.8–75.6)		93.8% (88.6–97.1)		43.5% (23.2–65.5)	

The number of available test results for each imaging feature may differ as some features were not reported for all subjects.

* 95% Confidence intervals are presented for sensitivity, specificity, positive and negative predictive values

** Suspected clinical tophus refers to assessment on physical examination

Table 3

Test characteristics for the presence of one or more ultrasound features of gout.

	Cases	Controls	Sensitivity	Specificity	PPV	NPV
One US feature of gout*	320/416 (77%)	64/408 (5%)	76.9% (72.6–80.9)	84.3% (80.4–87.7)	83.3% (79.2–86.9)	78.1% (74.0–82.0)
Two US features of gout	183/416 (44%)	19/408 (5%)	44.0% (39.2–48.9)	95.3% (92.8–97.2)	90.6% (85.7–94.2)	62.5% (58.6–66.4)
Three US features of gout	60/416 (14%)	10/408 (2%)	14.4% (11.2–18.2)	97.6% (95.6–98.8)	85.7% (75.3–92.9)	52.8% (49.2–56.4)

* Double contour sign, or tophus or snowstorm appearance.

Table 4

Test Characteristics of Ultrasound Features compared to controls with CPPD

	Cases with MSU-positive Gout N=414	CPPD positive controls N=98	Specificity	PPV	NPV
Gout vs CPPD*					
DCS	249 (60%)	7 (7%)	92.9% (85.8–97.1)	97.3% (94.4–98.9)	35.6% (29.7–41.7)
Tophi	189 (46%)	5 (5%)	94.9% (88.5–98.3)	97.4% (94.1–99.2)	29.5% (24.5–34.9)
Snowstorm	125 (30%)	10 (10%)	89.7% (81.9–94.9)	92.6% (86.8–96.4)	23.3% (19.1–27.9)
Any ultrasound feature	320 (77%)	19 (19%)	80.6% (71.4–87.9)	94.4% (91.4–96.6)	45.1% (37.6–52.8)

* Five subjects with both CPPD and MSU crystals were classified as gout rather than CPPD.

Table 5

Associations with positive ultrasounds among subjects with MSU-positive gout.

Factor	Univariable OR	Multivariable* OR
Double Contour Sign on Ultrasound		
Highest SUA	1.31 (1.13–1.53)	1.24 (1.05–1.46)
Tender proximal joint	2.02 (1.34–3.04)	1.79 (1.12–2.84)
Any radiographic feature [†]	3.04 (1.95–4.75)	3.19 (1.97–5.15)
Tophus on Ultrasound		
Suspected clinical tophus [‡]	7.47 (4.72–11.84)	7.05 (4.14–11.99)
Current SUA	1.22 (1.05–1.41)	1.23 (1.01–1.47)
Asymmetric swelling on radiograph	6.14 (3.93–9.60)	5.18 (3.14–8.56)
Snowstorm on Ultrasound		
Suspected clinical tophus [‡]	2.25 (1.46–3.47)	2.29 (1.40–3.76)
Cyst on radiograph	1.87 (1.20–2.93)	1.70 (1.06–2.72)
Number of episodes		
1	REF	REF
2–5	0.30 (0.13–0.69)	0.20 (0.08–0.51)
>5	0.55 (0.27–1.12)	0.34 (0.15–0.76)
Any Ultrasound Feature		
Suspected clinical tophus	4.81 (2.57–8.98)	4.77 (2.23–10.21)
Abnormal radiographic feature [†]	5.57 (3.34–9.29)	4.68 (2.68–8.17)
Current SUA	1.31 (1.09–1.56)	1.31 (1.06–1.62)

* Multivariable logistic regression models are adjusted for age, sex, and remainder of factors in the table.

[†] Abnormal radiographic feature refers to the presence of either cystic changes, asymmetry or erosion.

[‡] Suspected clinical tophus refers to assessment on physical examination

Highest ever and current SUA are continuous variables where a change in one unit reflects a change in serum urate of 0.1 mmol/l.

MTP1 involvement ever refers to symptomatic pain and/or swelling of the first metatarsophalangeal (MTP) joint during any episode.