Diagnosis of Gout: A Systematic Review in Support of an American College of Physicians Clinical Practice Guideline

Sydne J. Newberry, PhD; John D. FitzGerald, MD, PhD; Aneesah Motala, BA; Marika Booth, MS; Margaret A. Maglione, MPP; Dan Han, MPA; Abdul Tariq, BS; Claire E. O’Hanlon, MPP; Roberta Shanman, MLS; Whitney Dudley, BS; and Paul G. Shekelle, MD, PhD

Background: Alternative strategies exist for diagnosing gout that do not rely solely on the documentation of monosodium urate (MSU) crystals.

Purpose: To summarize evidence regarding the accuracy of clinical tests and classification algorithms compared with that of a reference standard of MSU crystals in joint aspirate for diagnosing gout.

Data Sources: Several electronic databases from inception to 29 February 2016.

Study Selection: 21 prospective cohort, cross-sectional, and case-control studies including participants with joint inflammation and no previous definitive gout diagnosis who had MSU analysis of joint aspirate.

Data Extraction: Data extraction and risk-of-bias assessment by 2 reviewers independently; overall strength of evidence (SOE) judgment by group.

Data Synthesis: Recently developed algorithms including clinical, laboratory, and imaging criteria demonstrated good sensitivity (up to 88%) and fair to good specificity (up to 96%) for diagnosing gout (moderate SOE). Three studies of dual-energy computed tomography (DECT) showed sensitivities of 85% to 100% and specificities of 83% to 92% for diagnosing gout (low SOE). Six studies of ultrasonography showed sensitivities of 37% to 100% and specificities of 68% to 97%, depending on the ultrasonography signs assessed (pooled sensitivity and specificity for the double contour sign: 74% [95% CI, 52% to 88%] and 88% [CI, 68% to 96%], respectively [low SOE]).

Limitation: Important study heterogeneity and selection bias; scant evidence in primary and urgent care settings and in patients with conditions that may be confused with or occur with gout.

Conclusion: Multidimensional algorithms, which must be validated in primary and urgent care settings, may help clinicians make a provisional diagnosis of gout. Although DECT and ultrasonography also show promise for gout diagnosis, accessibility to these methods may be limited.

Primary Funding Source: Agency for Healthcare Research and Quality. (Protocol registration: https://effectivehealthcare.ahrq.gov/ehc/products/564/1937/gout-protocol-140716.pdf)

© 2017 American College of Physicians

Gout is the most common form of inflammatory arthritis (1). In its initial stages, gout is recognized by acute, intermittent episodes of synovitis presenting with joint swelling and pain (referred to as acute gouty arthritis, acute gout attacks, or acute gout flares) that may progress to chronic and persistent symptoms. Gout is the result of excess serum urate crystalizing in the body (as monosodium urate [MSU]). The time between flares is referred to as the intercritical period. Intermittent attacks may progress to more chronic symptoms as the result of either joint damage or chronic synovitis. In some persons, MSU may aggregate in intra- or extra-articular regions (for example, around tendons, in bone or bursa, or in other soft tissues) to form tophi. Monosodium urate crystals may directly stimulate the inflammasome in leukocytes, causing an acute inflammatory attack (2).

Although the presence of MSU crystals in synovial fluid aspirated from the affected joint is sufficient for diagnosing gout, clinicians and researchers debate whether this approach is necessary. Classification and diagnostic algorithms that rely on other signs, symptoms, and laboratory tests without synovial fluid analysis exist. Joint aspiration may be technically difficult to perform and painful for the patient, particularly in smaller joints. Specimen handling and interpretation of synovial fluid analysis may be affected by experience and training. Moreover, the accuracy and utility of MSU assessment are affected by several factors (3-10). This review, conducted to support an American College of Physicians (ACP) clinical practice guideline, addresses the accuracy and safety of using clinical diagnostic or classification algorithms, dual-energy computed tomography (DECT), and ultrasonography for evaluating patients with gout symptoms compared with assessing MSU crystals in joint aspirate, which is considered the gold-standard diagnostic test.

Methods

We developed a protocol, followed PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) guidelines, and detailed our search and selection processes, inclusion criteria, and evidence tables in an evidence report (11-13).
Key Questions

Key questions posed by ACP representatives were revised on the basis of input from a group of key informants, a technical expert panel, and public comments. Questions addressed in this article are as follows: What is the accuracy of clinical signs and symptoms and other diagnostic tests (such as serum urate analysis, ultrasonography, computed tomography, DECT, and plain radiography), alone or in combination, compared with synovial fluid analysis in diagnosing acute gouty arthritis in patients with no previous gout diagnosis? What are the adverse effects (including pain, infection at the aspiration site, and radiation exposure) or harms (related to false-positive, false-negative, or indeterminate results) associated with tests used to diagnose gout? Additional questions and evidence regarding the effects of practitioner type and other factors on successful joint aspiration and accuracy of interpretation are available in the full evidence report (12, 13).

Data Sources and Searches

We searched, without language restrictions, PubMed, EMBASE, the Cochrane Library, gray literature, and the Web of Science from inception through 29 February 2016 using the word gout combined with the terms for diagnostic methods (MSU crystal analysis, joint aspiration, DECT, ultrasound, and x-ray), clinical signs and symptoms, and outcome measures, without filters specific for the diagnostic tests, as recommended (14). Supplement Table 1 (available at www.annals.org) shows the search methodology. We also obtained relevant references from a search conducted for a simultaneous review on gout management, considered studies suggested by experts, searched ClinicalTrials.gov and the Web of Science for recently completed studies and unpublished or non-peer-reviewed study findings, and contacted manufacturers of equipment and laboratory test kits used to diagnose gout for unpublished data specific to their use for gout diagnosis.

Study Selection

Titles and abstracts identified by the literature searches were screened by 2 reviewers, who independently conducted a full-text review of all selections to exclude articles that reported only on the incidence or prevalence, risk factors, or treatment of gout; included persons younger than 18 years; provided no usable data (sensitivities and specificities or data that could be used to calculate them); reported the same data as another article; enrolled only participants with established gout diagnoses; or did not clearly indicate the use of a recognized diagnostic standard. If necessary, disagreements regarding inclusion at the full-text stage were reconciled with the project leader’s input. We included original prospective and cross-sectional studies that assessed the accuracy (sensitivity and specificity) or safety of tests used to diagnose gout in persons with no prior definitive gout diagnosis who presented with joint inflammation, and in which the reference standard was MSU analysis or a combination of MSU analysis, American Rheumatism Association (ARA) (now the American College of Rheumatology [ACR]) criteria for gout diagnosis, and tests to confirm or rule out other causes of inflammatory arthritis. Studies that enrolled patients with asymptomatic hyperuricemia were excluded. To assess safety, we also included case reports and case series.

Data Extraction and Quality Assessment

Two reviewers abstracted study-level details from articles accepted for inclusion. Outcomes (sensitivity, specificity, and positive and negative predictive value [PPV and NPV]) were singly abstracted and verified by another reviewer. Risk of bias (study quality) of each included study was assessed independently by 2 reviewers using the QUADAS-2 (Revised Quality Assessment of Diagnostic Accuracy Studies) tool (15, 16). Supplement Table 2 (available at www.annals.org) includes the results of the quality assessment. Disagreements regarding study details were reconciled by group discussion, and those related to study quality were mediated by the project leader.

Data Synthesis and Analysis

We organized our narrative descriptions of evidence, which focused on study quality, settings, and findings, according to categories of tests, as well as chronologically. If several studies compared similar tests with the same reference standard, we used bivariate metaregression to pool studies (17). As a group, the reviewers assessed the overall strength of evidence (SOE) for each major comparison and outcome as high, moderate, low, or insufficient using guidance suggested by the Effective Health Care Program (18).

Role of the Funding Source

This topic was nominated by the ACP to the Agency for Healthcare Research and Quality (AHRQ) and funded by AHRQ. Staff at AHRQ and ACP helped to develop and refine the scope of the study and reviewed the draft report but had no role in data extraction, synthesis, or rating of evidence.

RESULTS

The Figure depicts the search and selection process that identified 22 total articles addressing the accuracy (n = 21) or safety (n = 3) of various diagnostic methods. Characteristics and findings of the studies are detailed in Supplement Table 3 (available at www.annals.org).

Validity of Clinical Classification and Diagnostic Criteria

Twelve studies examined diagnostic or classification algorithms for diagnosing gout (19–30). Of these 12 studies, 11 compared the predictions from 10 clinical algorithms (described in Table 1) with assessment of synovial fluid MSU crystals in all enrolled patients (19–23, 25–30). Study quality was good for all but 2 studies (19, 23). All studies were conducted in academic rheumatology departments, although several purposely enrolled patients who were referred by primary care physicians. The ARA’s algorithm included joint fluid culture to rule out the co-occurrence of septic arthritis (19). Sensitivities and specificities for the cli-
cal algorithms varied (compared with MSU analysis as the reference standard) (Table 2).

**Rome and New York Criteria**

Three studies found that the Rome and New York criteria—the first sets of clinical criteria developed to classify gout—had limited sensitivity (70% to 80%) and specificity (79% to 80%) for diagnosing gout compared with identifying MSU crystals in affected joints (20, 27, 30).

**ARA Criteria**

In 1977, in a high-quality study using clinical characteristics of 706 patients seen in 38 rheumatology clinics across the United States, an ARA subcommittee developed an algorithm to classify gout for research purposes (19). The final algorithm included 12 clinical signs (not including MSU analysis or presence of tophi, which were considered diagnostic by themselves), with the presence of 6 or more of these signs required to classify a patient as having gout. The initial assessment reported sensitivities of 96%, 88%, and 74%, and specificities of 73% to 93%, 89% to 99%, and 97% to 100% for 5 or more, 6 or more, and 7 or more positive criteria, respectively. Only 47% of participants underwent MSU testing, with physician judgment serving as the reference standard for those not tested.
Five subsequent tests of the ARA criteria were done in groups of patients with suspected gout, all of whom underwent MSU analysis (19, 20, 22, 27, 30). These studies, which enrolled 82 to 983 patients, reported sensitivities of 70% to 92% and specificities of 53% to 92% (Table 2). Positive predictive value ranged from 66% to 92%, and NPV from 69% to 86% (study quality was high for 2 studies and moderate for the other 2). In 1 study, among the patients with false-positive results, 50% had deposits of calcium pyrophosphate crystals (20).

Two studies compared the sensitivity and specificity of the algorithm between patients with recent onset of initial symptoms and those with more longstanding flares. For patients with a more recent onset of the first episode, the sensitivity and specificity

| Table 1. Summary of Components of Clinical Algorithms for Diagnosing Gout* |
|-------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Components                  | Rome Criteria (20) | New York Criteria, (20)† | Wallace et al, 1977 (ARA/ACR) | EULAR, 2006 (31) | Janssens Diagnostic Rule, 2010 (21) |
| Clinical characteristics | ARA/ACR          | ARA/ACR          | ARA/ACR          | ARA/ACR          | ARA/ACR          |
| >1 attack of acute arthritis | ✓              | ✓              | ✓              | ✓              | ✓              |
| Maximum inflammation developed within 1 d | ✓              | ✓              | ✓              | ✓              | ✓              |
| Painful joint swelling, abrupt onset, clearing within 1-2 wk | ✓              | ✓              | ✓              | ✓              | ✓              |
| Monoarthritis/oligoarthritis attack | ✓              | ✓              | ✓              | ✓              | ✓              |
| Redness observed over joints | ✓              | ✓              | ✓              | ✓              | ✓              |
| First MTP joint painful or swollen | ✓              | ✓              | ✓              | ✓              | ✓              |
| Unilateral first MTP joint attack | ✓              | ✓              | ✓              | ✓              | ✓              |
| Unilateral tarsal joint attack | ✓              | ✓              | ✓              | ✓              | ✓              |
| Abrupt onset and remission in 1-2 wk initially | ✓              | ✓              | ✓              | ✓              | ✓              |
| Response to colchicine—major reduction in inflammation within 48 h | ✓              | ✓              | ✓              | ✓              | ✓              |
| Pain intensity ≥9/10 | ✓              | ✓              | ✓              | ✓              | ✓              |
| Involvement of toes, foot, or ankle | ✓              | ✓              | ✓              | ✓              | ✓              |
| Treatment with corticosteroids | ✓              | ✓              | ✓              | ✓              | ✓              |
| Treatment with nonsteroidal anti-inflammatory drugs | ✓              | ✓              | ✓              | ✓              | ✓              |
| Resolution of pain <15 d after onset | ✓              | ✓              | ✓              | ✓              | ✓              |
| Tophi (proven or suspected) | ✓              | ✓              | ✓              | ✓              | ✓              |
| First MTP joint ever involved | ✓              | ✓              | ✓              | ✓              | ✓              |
| Location of currently tender joints: other foot/ankle | ✓              | ✓              | ✓              | ✓              | ✓              |
| Imaging | ✓              | ✓              | ✓              | ✓              | ✓              |
| Asymmetrical swelling within a joint on radiography | ✓              | ✓              | ✓              | ✓              | ✓              |
| Subcortical cysts without erosions on radiography | ✓              | ✓              | ✓              | ✓              | ✓              |
| Radiographic erosion or cyst | ✓              | ✓              | ✓              | ✓              | ✓              |
| Double contour sign on ultrasonography | ✓              | ✓              | ✓              | ✓              | ✓              |
| DECT urate deposition on symptomatic joint or bursa | ✓              | ✓              | ✓              | ✓              | ✓              |
| Joint fluid | ✓              | ✓              | ✓              | ✓              | ✓              |
| Negative results on joint fluid culture | ✓              | ✓              | ✓              | ✓              | ✓              |
| MSU crystals in synovial fluid or tissues†† | ✓              | ✓              | ✓              | ✓              | ✓              |
| Comorbid conditions or risk factors | ✓              | ✓              | ✓              | ✓              | ✓              |
| Hyperuricemia | ✓              | ✓              | ✓              | ✓              | ✓              |
| Male sex | ✓              | ✓              | ✓              | ✓              | ✓              |
| Hypertension or ≥1 cardiovascular disease | ✓              | ✓              | ✓              | ✓              | ✓              |
| Hypertriglyceridemia | ✓              | ✓              | ✓              | ✓              | ✓              |

3e = Evidence, Expertise, Exchange; ACR = American College of Rheumatology; ARA = American Rheumatology Association; DECT = dual-energy computed tomography; EULAR = European League Against Rheumatism; MSU = monosodium urate; MTP = metatarsophalangeal; SUGAR = Study for Updated Gout Classification Criteria.

* Meets 2 of the criteria.
† MSU crystals in joint fluid or tophus or tissue or meets 2 of the criteria.
‡ ≥4/8 of the criteria checked.
§ Guideline 1 states that MSU is required for a definitive diagnosis, but, in its absence, such clinical criteria as those checked can be used or characteristic imaging findings may substitute.
¶ Designed to be administered by telephone by nonphysicians to assess the prevalence of gout via patient self-report; treatment questions refer to the most prominent episode.
** Involvement of joints or bursae other than the first MTP only as polyarticular presentation.
†† Several algorithms specified the presence of these crystals as definitive in lieu of other signs.
were, respectively, 71% and 84% in 1 study and 58% and 92% (PPV, 71%; NPV, 86%) in the other. For patients with a longer history of symptoms, the sensitivity and specificity were, respectively, 92% and 53% in 1 study and 89% and 65% (PPV, 92%; NPV, 59%) in the other (25, 30).

Janssens Diagnostic Rule

The Janssens diagnostic rule (also known as the Netherlands criteria) is a 7-item algorithm developed for use in clinical diagnosis (Table 1). Creation of the diagnostic rule was based on 328 Dutch primary care patients with suspected gout, who constituted the population Janssens used to assess the accuracy of the ARA classification criteria (study quality, high) (21).

The algorithm, which incorporated clinical signs and symptoms used by primary care physicians, was assessed in 3 high- or moderate–high-quality studies ranging in size from 233 to 983 patients with suspected gout in the United States, the Netherlands, and Thailand (24, 27, 30). In the initial validation study, each of the 7 items was assigned a point value from 0.5 to 3.5, for a total of 13 possible points; the PPV of scores greater than 8 was 87%, and the NPV of scores less than 4 was 95% compared with MSU crystal assessment (24). In 2 subsequent studies using the same threshold values as the original, the sensitivities were 73% to 96% and the specificities 47% to 76%, depending on the recency of initial symptoms (27, 30). Sensitivities ranged from 73% to 88% and specificities from 75% to 86% for patients with recent-onset symptoms; among patients whose symptoms began at least 2 years earlier, sensitivities were 92% to 96% and specificities 47% to 50% (Table 2) (27, 30).

Clinical Gout Diagnosis Criteria

The Clinical Gout Diagnosis (CGD) is an 8-item algorithm developed in 2010 from a subset of the ARA/ACR classification criteria for diagnosing chronic gout in the primary care setting without reliance on joint fluid analysis. It was validated in 167 consecutive patients with arthritis who were referred by primary care physicians to 2 academic rheumatology clinics in Mexico City (study quality, low) (Table 1) (23). Using a cutoff of 4 or more positive items and the reference standard of MSU crystal analysis, the algorithm had a sensitivity of 97% and specificity of 96% (3% of patients with rheumatoid arthritis, 7% with septic arthritis, and 3% with osteoarthritis fulfilled 4 or more of the 8 criteria, compared with 97% of those with gout). Ninety percent of patients who had a CGD score of 4 or higher scored positive for gout on the basis of the Janssens criteria. All patients with early gout and 94% of women with gout also fulfilled 4 or more criteria.

The CGD was assessed further by 2 other groups (27, 30). One group, using the data set of Taylor and colleagues (27), found that the CGD had 87% sensitivity and 66% specificity in patients whose symptoms began less than 2 years earlier and 99% sensitivity and 34% specificity in those whose symptoms began more than 2 years earlier (27). In an analysis using the Thai patient data set, the CGD showed sensitivity and specificity of 97% and 68% overall, 89% and 82% in patients whose symptoms began up to 2 years earlier, and 99% and 31% in those whose symptoms began more than 2 years previously (30).

Monoarthritis of the First Metatarsophalangeal Joint

Monoarthritis of the first metatarsophalangeal joint, as the sole diagnostic criterion for gout, was evaluated in a high-quality study in 159 Dutch patients seen by
Review

Systematic Review in Support of ACP Guideline on Diagnosis of Gout

Table 2. Summary of Findings

| Study, Year (Reference) | Sensitivity, % | Specificity, % |
|-------------------------|---------------|---------------|
| **Clinical algorithm**  |               |               |
| New York criteria (2 of 4 criteria) | Malik et al, 2009 (20) | 70 | 79 |
| Taylor et al, 2016 (27) | 80 | 80 |
| Jatuworapruk et al, 2016 (30) | 46 (≤2 y) | 85 (≤2 y) |
| 87 (>2 y) | 65 (>2 y) |
| Rome criteria (2 of 3 criteria) | Malik et al, 2009 (20) | 67 | 89 |
| Taylor et al, 2016 (27) | 76 | 81 |
| Jatuworapruk et al, 2016 (30) | 73 (≤2 y) | 87 (≤2 y) |
| 76 (>2 y) | 65 (>2 y) |
| ARA/ACR (≥6 of 12 criteria) | Wallace et al, 1977 (19) | 88 | 89-99 |
| Malik et al, 2009 (20) | 70 | 79 |
| Janssens et al, 2010 (22) | 80 | 64 |
| Taylor et al, 2016 (27) | 71, 56 (≤2 y) | 84, 92 (≤2 y) |
| 92, 89 (>2 y) | 53, 65 (>2 y) |
| Jatuworapruk et al, 2016 (30) | 58 (≤2 y) | 92 (≤2 y) |
| 89 (>2 y) | 65 (>2 y) |
| Janssens diagnostic rule (≥8 of 13 points) (the Netherlands) | Taylor et al, 2016 (27) | 88 (≤2 y) | 75 (≤2 y) |
| 96 (>2 y) | 47 (>2 y) |
| Jatuworapruk et al, 2016 (30) | 73 (≤2 y) | 86 (≤2 y) |
| 92 (>2 y) | 50 (>2 y) |
| Clinical Gout Diagnosis (≥4 of 8 criteria) | Vázquez-Mellado et al, 2012 (23) | 97, 97 | 96, 68 |
| Taylor et al, 2016 (27) | 87 (≤2 y) | 66 (≤2 y) |
| 99 (>2 y) | 34 (>2 y) |
| Jatuworapruk et al, 2016 (30) | 89 (≤2 y) | 82 (≤2 y) |
| 99 (>2 y) | 31 (>2 y) |
| **SUGAR** (28) | 84 with ultrasonography and radiography | 81 with ultrasonography and radiography |
| 88 without imaging | 72 without imaging |
| Richette Survey (26) | 88 | 90-93 |
| 2016 ACR/EULAR (29) | 92 with imaging | 89 with imaging |
| 85 without imaging | 78 without imaging |
| **Monoarthritis of the first MTP joint** (22) | 99 | 8 |
| **DECT** | Bongartz et al, 2015 (34); Huppertz et al, 2014 (33); and Glazebrook et al, 2011 (35) | 85-100 | 83-92 |
| **Ultrasonography** | Pooled results for 4 trials of double contour sign (36-39) | 74 (95% CI, 52-88) | 88 (95% CI, 68-96) |

ACR = American College of Rheumatology; ARA = American Rheumatology Association; DECT = dual-energy computed tomography; EULAR = European League Against Rheumatism; MTP = metatarsophalangeal; SUGAR = Study for Updated Gout Classification Criteria.

Across all criteria sets, the sensitivity was lower and specificity higher in patients with disease of shorter duration (Table 2). In patients who had symptoms for 2 years or less, the sensitivity ranged from 58% (for the New York criteria) to 88% (for the Janssens diagnostic rule), whereas in those who had symptoms for more than 2 years, it was 84% (Rome criteria) to 99% (CGD). Among patients whose symptoms were of shorter duration, the specificity ranged from 66% (CGD) to 88% (New York criteria), whereas among those with more long-standing symptoms, it was 84% (CGD) to 70% (New York criteria). Analysis of MSU crystals was the reference standard.

Researchers assessed the sensitivity and specificity of 4 algorithms that allowed presence of synovial joint MSU to be sufficient for a gout diagnosis. Among patients with recent-onset symptoms, sensitivity ranged from 93% (for the Rome criteria) to 100% (for the New York, ARA, and CGD criteria), whereas among those with more long-standing symptoms, it was 99% to 100% for those 4 criterion sets. For patients with symptoms of shorter duration, specificity ranged from 66% (CGD) to 88% (New York), whereas among those with longer-duration disease, it was 34% (CGD) to 70% (New York). Thus, all criterion sets had somewhat poor specificity for more long-standing symptoms.

The investigators then devised their own 10-item algorithm from 56 criteria selected by an appropriate-ness panel (2 of the items were imaging characteristics: 1 ultrasonography and 1 plain radiography) (28). The items with the strongest predictive value were the presence of tophi (odds ratio [OR], 7.29 [95% CI, 2.42 to 21.99]), resolution of pain by 2 weeks (OR, 3.58 [CI, 1.85 to 6.95]), and a double contour sign on ultrasonography and radiography.
DECT studies assessed the sole use of computed tomography compared with ultrasonography with DECT) (33, 36–40). No Imaging Studies clinical and laboratory criteria (study quality, high).

### Ultrasonography

Six studies evaluated the use of ultrasonography for assessing symptomatic joints by comparing the predictions based on ultrasonography signs with identification of synovial fluid MSU crystals (4 studies) or with a combination of a validated clinical algorithm and MSU in patients with suspected gout or an undefined crystalline arthropathy. All studies were conducted in academic rheumatology departments. Three studies assessed symptomatic as well as asymptomatic joints in the same patients (33, 36, 40). The overall quality was good for 4 studies (36, 37, 39, 40) and moderate for 2 (33, 38).

Ultrasonography showed overall sensitivities ranging from 37% to 100% and specificities from 68% to 97%, depending on the signs assessed (33, 36–40). The sensitivity and specificity of the double contour sign, a hyperechoic, irregular enhancement over the hyaline cartilage surface, ranged from 37% (attributed to very early-onset gout) to 78% and 64% to 97%, respectively. The pooled sensitivity and specificity for the double contour sign from 4 studies (464 joints) that contributed data were 74% (CI, 52 to 88) and 88% (CI, 68 to 96), respectively (36–39).

### Adverse Events Associated With Gout Diagnostic Tests

Two studies assessed adverse effects associated specifically with tests used to diagnose gout (34, 41). One small, single-site, good-quality study reported no adverse events associated with synovial fluid aspiration for MSU analysis or with DECT (34). An assessment of adverse events associated with synovial fluid aspiration for MSU analysis was conducted in conjunction with the SUGAR project (41). This good-quality study, conducted at 25 centers in 16 countries among 887 patients, identified 1 serious adverse event (septic arthritis 11 days after arthrocentesis; culture positive for *Staphylococcus aureus*; event rate, 0.1% [CI, 0% to 0.34%]) and 11 nonserious adverse events (mostly mild pain after the procedure; event rate, 1.4% [CI, 0.6% to

2015 ACR/European League Against Rheumatism Gout Classification Criteria

A joint committee of the ACR and European League Against Rheumatism (EULAR) recently developed a new set of criteria primarily intended for classification (29). Based on the SUGAR criteria and further validated with a representative set of test patient profiles, this criterion set takes into account the progressive and varied nature of the disease and the likelihood that patients who do not meet a particular criterion when first classified may fulfill it later. The entry criterion that must be met is at least 1 episode of swelling, pain, or tenderness in a peripheral joint. Identification of MSU crystals in the synovial fluid is considered diagnostic for gout; however, if this criterion is not met, a hierarchical set of 7 criteria, arranged in clinical, laboratory, and imaging domains, is applied (Table 1). Each criterion contributes points (weighted for its putative importance), ranging from −4 to 4, to the total score, which may range from 6 to 23. Some criteria, such as serum urate, may contribute more or fewer points (including negative points) in proportion to severity. A score of 8 or more is considered diagnostic for gout. With MSU applied as the reference standard, the classification criteria had a sensitivity of 92% and specificity of 89% (excluding clinical and imaging domains) or 85% and 78% (excluding imaging) (Table 2) (29). Thus, imaging findings improved both the sensitivity and specificity of clinical and laboratory criteria (study quality, high).

## Imaging Studies

Three studies assessed the use of DECT (33–35), and 6 evaluated the use of ultrasonography (1 study compared ultrasonography with DECT) (33, 36–40). No studies assessed the sole use of computed tomography or plain radiography.

### DECT

The DECT studies compared the predictions based on this imaging method with synovial fluid analysis for MSU crystals by using a validated clinical algorithm or some combination of the 2 reference standards. All the studies were conducted in academic rheumatology departments. Quality was high for 1 study (34) and moderate for 2 studies (33, 35).

One study included 94 consecutive patients seen in an academic rheumatology clinic. Two rheumatologists interpreted the DECT images. Of the 94 patients, 31 had successful joint aspiration. Sensitivity was 100% (CI, 74% to 100%) for both readers, and specificity was 89% (CI, 67% to 99%) and 79% (CI, 54% to 94%) for the 2 readers (35). In another study, 60 consecutive patients with suspected gout were seen in a German rheumatology clinic, where they had synovial fluid aspiration and DECT. The sensitivity and specificity for diagnosing gout were 85% and 86%, respectively (PPV, 92%; NPV, 75%). All false-negative results found were in patients with recent-onset acute gout (33). However, a study in 81 consecutive patients with suspected gout who were seen in the rheumatology department at Mayo Clinic reported 90% sensitivity (CI, 76% to 97%) and 83% specificity (CI, 68% to 93%) with MSU as the reference standard. Sensitivity was lower in patients with recent-onset acute gout; all false-positive results were in patients with advanced osteoarthritis of the knee (34).

Across the 3 studies, the sensitivity of DECT for predicting gout ranged from 85% to 100% and the specificity from 83% to 92% compared with MSU analysis and a validated clinical algorithm (33–35). Two of the studies showed that DECT is less sensitive in patients with a shorter history of flares than in those with a longer one.

### Ultrasonography

Six studies evaluated the use of ultrasonography for assessing symptomatic joints by comparing the predictions based on ultrasonography signs with identification of synovial fluid MSU crystals (4 studies) or with a combination of a validated clinical algorithm and MSU in patients with suspected gout or an undefined crystalline arthropathy. All studies were conducted in academic rheumatology departments. Three studies assessed symptomatic as well as asymptomatic joints in the same patients (33, 36, 40). The overall quality was good for 4 studies (36, 37, 39, 40) and moderate for 2 (33, 38).

Ultrasonography showed overall sensitivities ranging from 37% to 100% and specificities from 68% to 97%, depending on the signs assessed (33, 36–40). The sensitivity and specificity of the double contour sign, a hyperechoic, irregular enhancement over the hyaline cartilage surface, ranged from 37% (attributed to very early-onset gout) to 78% and 64% to 97%, respectively. The pooled sensitivity and specificity for the double contour sign from 4 studies (464 joints) that contributed data were 74% (CI, 52 to 88) and 88% (CI, 68 to 96), respectively (36–39).

### Adverse Events Associated With Gout Diagnostic Tests

Two studies assessed adverse effects associated specifically with tests used to diagnose gout (34, 41). One small, single-site, good-quality study reported no adverse events associated with synovial fluid aspiration for MSU analysis or with DECT (34). An assessment of adverse events associated with synovial fluid aspiration for MSU analysis was conducted in conjunction with the SUGAR project (41). This good-quality study, conducted at 25 centers in 16 countries among 887 patients, identified 1 serious adverse event (septic arthritis 11 days after arthrocentesis; culture positive for *Staphylococcus aureus*; event rate, 0.1% [CI, 0% to 0.34%]) and 11 nonserious adverse events (mostly mild pain after the procedure; event rate, 1.4% [CI, 0.6% to
2.1%). No studies reported on adverse events associated with ultrasonography or clinical examination.

A third study assessed the outcomes of misdiagnosis of acute gout in 2 academic medical centers in South Korea (42). This moderate-quality study found that missed or delayed diagnosis was associated with a longer interval between the onset of attack and joint aspiration and resulted in longer hospitalizations.

**DISCUSSION**

Our principal findings are that several algorithms based solely on clinical signs and symptoms showed sensitivities and specificities greater than 80% for diagnosing early-onset gout compared with the gold standard of synovial fluid aspiration and MSU crystal analysis (moderate SOE). Ultrasonography and DECT also showed some evidence of diagnostic validity on their own or in conjunction with algorithms to improve their diagnostic accuracy (low SOE). Although detection of synovial fluid MSU crystals is considered the diagnostic benchmark, questions linger about the sensitivity of MSU analysis in early gout and the marginal value of aspiration or advanced imaging methods in routine clinical practice. Aspiration of smaller joints is technically difficult and painful. Misdiagnosis of gout has been associated with the use of joint aspiration and MSU assessment as well as with use of diagnostic algorithms and imaging techniques. The recency of the first attack (more or less than 2 years before the current attack) affects the sensitivity and specificity of diagnostic tests, as longer disease duration improves the performance of both algorithms and imaging tests. Evidence from imaging studies suggests that the location of the affected site also may be a factor in misdiagnosis; for example, both ultrasonography (37) and DECT (34) were less likely to detect signs of gout-related crystal deposition in the knee than in the first metatarsopalangeal joint.

The most important issue for gout diagnosis in the primary, urgent, or emergency care setting is whether using some combination of clinical signs and symptoms to diagnose gout, without relying on synovial fluid analysis of MSU crystals or imaging, is sensitive enough to detect or confirm suspicion of gout (particularly recent-onset disease). Several algorithms (including the CGD criteria and Janssens diagnostic rule) had high sensitivity in patients whose attacks were of recent onset—the patients most likely to be seen in the primary, urgent, or emergency care setting (27, 30)–without reliance on MSU analysis. However, the relatively low specificity of these algorithms raises the risk for a missed diagnosis of a condition with a similar initial clinical presentation, notably septic arthritis. None of the criteria or algorithms has been tested specifically to rule out septic arthritis. Therefore, if septic arthritis is in the differential, aspiration for Gram stain and culture remains an essential part of the diagnostic evaluation.

This review had several limitations. Selection bias was likely, because few studies enrolled only patients with suspected gout but without a confirmed diagnosis, and many were conducted in referral settings. Only 1 study was designed specifically to evaluate whether a particular test or algorithm could rule out conditions that might be confused with gout or that might co-occur in a patient with gout, such as calcium pyrophosphate deposition or septic arthritis (19). A limitation specific to the algorithm studies is that their objective typically was to develop or validate algorithms to classify patients for research and clinical trials and not to develop criteria for making a diagnosis in individual patients. Specifically, the aim of classification criteria often is to allow relatively homogeneous groups of study participants with clearly defined signs and symptoms to be identified and enrolled in trials. Diagnosis involves a global judgment on the physician’s part, taking all aspects of a patient’s clinical presentation into account (43). Also, whereas higher specificity may be preferred for trial classification, higher sensitivity may be needed for accurate diagnosis and therefore treatment (44). Nevertheless, for diseases with clear causes and presentation, classification and diagnostic criteria may be somewhat interchangeable; whether MSU analysis of joint aspirate fulfills these requirements remains unclear.

In summary, the results of studies assessing the sensitivity and specificity of various clinical algorithms for gout diagnosis or classification suggest that these instruments show promise for providing at least a provisional gout diagnosis in patients who present with signs of early-stage disease, who are most likely to be seen in primary, urgent, and emergency care settings. Imaging methods, such as ultrasonography and DECT, also seem favorable for gout diagnosis, but primary and urgent care settings may lack access to such equipment or the expertise necessary to use these techniques.

From RAND Corporation, Santa Monica, and David Geffen School of Medicine at the University of California and Veterans Affairs Greater Los Angeles Healthcare System, Los Angeles, California.

**Disclaimer:** The authors of this manuscript are responsible for its content. Statements in the manuscript should not be construed as endorsement by the AHRQ or U.S. Department of Health and Human Services. The AHRQ retains a license to display, reproduce, and distribute the data and the report from which this manuscript was derived under the terms of the agency’s contract with the author.

**Acknowledgment:** The authors acknowledge Patricia Smith for her administrative assistance on the project. In addition, the authors acknowledge the guidance provided by the key informants and technical expert panel members on the evidence report: Hyon K. Choi, MD, PhD; Nicola Dalbeth, MD; Russell Harris, MD, MPH; David Hoelting, MD; Sanford S. Kaplan, DDS; Gerald D. Levy, MD; Esther Myers, PhD, RD; Savvas Nicolaou, MD; Ralf G. Thiele, MD; Richard Treger MD; Daniel Waxman, MD, PhD; Lee Ann Weintraub, MS, RD; and Neil Wenger, MD.
Financial Support: The evidence report was funded under contract 290-2012-000061 from the AHRQ, U.S. Department of Health and Human Services.

Disclosures: Ms. Booth and Ms. Shanmugam report grants from AHRQ during the conduct of the study. Ms. O’Hanlon reports a contract with AHRQ during the conduct of the study. Dr. Shekelle reports grants from AHRQ, personal fees from ECRI Institute, and royalties from UpToDate. Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M16-0462.

Reproducible Research Statement: Study protocol: Available at https://effectivehealthcare.ahrq.gov/ehr/products/564/1937/gout-protocol-140716.pdf. Statistical code and data set: Not available.

Requests for Single Reprints: Sydne J. Newberry, PhD, RAND Corporation, 1776 Main Street, Santa Monica, CA 90401; e-mail, sydnen@rand.org.

Current author addresses and author contributions are available at www.annals.org.

References

1. 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. [PMID: 21800283] doi:10.1002/art.30520
2. Kingsbury SR, Conaghan PG, McDermott MF. The role of the NLRP3 inflammasome in gout. J Inflamm Res. 2011;4:39-49. [PMID: 22096368] doi:10.2147/JIR.S11330
3. Harrold LR, Mazor KM, Negron A, Ogarek J, Firneno C, Yood RA. Primary care providers’ knowledge, beliefs and treatment practices for gout: results of a physician questionnaire. Rheumatology (Oxford). 2013;52:1623-9. [PMID: 23620554] doi:10.1093/rheumatology/ ket158
4. Owens D, Whelan B, McCarthy G. A survey of the management of gout in primary care. Br Med J. 2008;336:147-9. [PMID: 18624262]
5. Khosa S, Thiele R, Baumbauer JF. Ultrasound guidance for intra-articular injection of the foot and ankle. Foot Ankle Int. 2009;30:886-90. [PMID: 19755074] doi:10.3113/FAL.2009.0886
6. Hasselbalch P. Variation in synovial fluid analysis by hospital laboratories. Arthritis Rheum. 1987;30:637-42. [PMID: 3606682]
7. Schumacher HR Jr, Sieck MS, Rothfuss S, Claman CN, Baumrind AH. Synovial fluid crystal analysis. J Clin Rheumatol. 2009;15:22-4. [PMID: 19125136] doi:10.1097/RHU.0b013e3181945b79
8. Janssens HJ, Fransen J, van de Lisdonk EH, van Riel PL, van Weel C, Janssens M. A diagnostic rule for acute gouty arthritis in primary care without joint fluid analysis. Arch Intern Med. 2010;170:1120-6. [PMID: 20625017] doi:10.1001/archinternmed.2010.196
9. Janssens JH, Janssen M, van de Lisdonk EH, Fransen J, van Riel PL, van Weel C. Limited validity of the American College of Rheumatology criteria for classifying patients with gout in primary care (Letter). Ann Rheum Dis. 2010;69:1255-6. [PMID: 19910298] doi:10.1136/ard.2009.123687
10. Vázquez-Mellado J, Hernández-Cuevas CB, Alvarez-Hernández E, Ventura-Ríos L, Peláez-Ballestas I, Casasola-Vargas J, et al. The diagnostic value of the proposal for clinical gout diagnosis (CGD). Clin Rheumatol. 2012;31:429-34. [PMID: 21979446] doi:10.1007/s10067-011-1873-4
11. Kienhorst LB, Janssens HJ, Fransen J, Janssen M. The validation of a diagnostic rule for gout without joint fluid analysis: a prospective study. Rheumatology (Oxford). 2015;54:609-14. [PMID: 25231179] doi:10.1093/rheumatology/keu378
12. Kienhorst LB, Janssens HJ, Fransen J, van de Lisdonk EH, Janssen M. Arthritis of the first metatarsophalangeal joint is not always gout: a prospective cohort study in primary care patients. Joint Bone Spine. 2014;81:342-6. [PMID: 24468668] doi:10.1016/j.jbspin.2013.12.001
13. Richette P, Clerson P, Boué S, Chalès G, Doherty M, Filpo RM, et al. Identification of patients with gout: elaboration of a questionnaire for epidemiological studies. Ann Rheum Dis. 2015;74:1684-90. [PMID: 24796335] doi:10.1136/annrheumdis-2014-204976
14. Taylor WJ, Fransen J, Dalbeth N, Neogi T, Schumacher HR, Brown M, et al. Performance of classification criteria for gout in early and established disease. Ann Rheum Dis. 2016;75:178-82. [PMID: 25351521] doi:10.1136/annrheumdis-2014-206364
15. Taylor WJ, Fransen J, Janssen TL, Dalbeth N, Schumacher HR, Brown M, et al. Study for updated gout classification criteria: identi-
fication of features to classify gout. Arthritis Care Res (Hoboken). 2015;67:1304-15. [PMID: 25777045] doi:10.1002/acr.22585

29. Neogi T, Jansen TL, Dalbeth N, Fransen J, Schumacher HR, Berendsen D, et al. 2015 Gout classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis. 2015;74:1789-98. [PMID: 26359487] doi:10.1136/annrheumdis-2015-208237

30. Jatuworapruk K, Lhakum P, Pattamapaspong N, Kasitanon N, Wangkaew S, Louthrenoo W. Performance of the existing classification criteria for gout in Thai patients presenting with acute arthritis. Medicine (Baltimore). 2016;95:e2730. [PMID: 26844519] doi:10.1097/MD.0000000000002730

31. Zhang W, Doherty M, Pascual E, Bardin T, Barskova V, Conaghan P, et al; EULAR Standing Committee for International Clinical Studies Including Therapeutics. EULAR evidence based recommendations for gout. Part I: diagnosis. Report of a task force of the Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). Ann Rheum Dis. 2006;65:1301-11. [PMID: 16707533]

32. Sivera F, Méndez M, Carmona L, Kydd AS, Moi J, Seth R, 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. [PMID: 23868909] doi:10.1136/annrheumdis-2013-203325

33. Huppertz A, Hermann KG, Diekhoff T, Wagner M, Hamm B, Schmidt WA. Systemic staging for urate crystal deposits with dual-energy CT and ultrasound in patients with suspected gout. Rheumatol Int. 2014;34:763-71. [PMID: 24619560] doi:10.1007/s00296-014-2979-1

34. Bongartz T, Glazebrook KN, Kavros SJ, Murthy NS, Merry SP, Franz WB 3rd, et al. Dual-energy CT for the diagnosis of gout: an accuracy and diagnostic yield study. Ann Rheum Dis. 2015;74:1072-7. [PMID: 24671771] doi:10.1136/annrheumdis-2013-205095

35. Glazebrook KN, Guimarães LS, Murthy NS, Black DF, Bongartz T, Manek NJ, et al. Identification of intraarticular and periarticular uric acid crystals with dual-energy CT: initial evaluation. Radiology. 2011;261:516-24. [PMID: 21926378] doi:10.1148/radiol.11102485

36. 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. Clin Rheumatol. 2014;33:555-9. [PMID: 24510062] doi:10.1007/s10067-014-2505-6

37. Lai KL, Chiu YM. Role of ultrasonography in diagnosing gouty arthritis. J Med Ultrasound. 2011;19:7-13. doi:10.1016/j.jmu.2011.01.003.

38. Rettenbacher T, Ennenmoser S, Weirich H, Ulmer H, Hartig F, Klotz W, et al. Diagnostic imaging of gout: comparison of high-resolution US versus conventional X-ray. Eur Radiol. 2008;18:621-30. [PMID: 17994238]

39. Löffler C, Sattler H, Peters L, Löffler U, Uppenkamp M, Bergner R. Distinguishing gouty arthritis from calcium pyrophosphate disease and other arthritides. J Rheumatol. 2015;42:513-20. [PMID: 25399385] doi:10.3899/jrheum.140634

40. Zufferey P, Pascal Z, Valcov R, Fabreguet I, Dumusc A, Omoumi P, et al. A prospective evaluation of ultrasound as a diagnostic tool in acute monocrystalline arthritis. Arthritis Res Ther. 2015;17:188. [PMID: 26198435] doi:10.1186/s13075-015-0701-7

41. Taylor WJ, Fransen J, Dalbeth N, Neogi T, Ralph Schumacher H, Brown M, et al. Diagnostic arthrocentesis for suspicion of gout is safe and well tolerated. J Rheumatol. 2016;43:150-3. [PMID: 26628602] doi:10.3899/jrheum.140684

42. Park JW, Ko DJ, Yoo JJ, Chang SH, Cho HJ, Kang EH, et al. Clinical factors and treatment outcomes associated with failure in the detection of urate crystal in patients with acute gouty arthritis. Korean J Intern Med. 2014;29:361-9. [PMID: 24851071] doi:10.3904/kjim.2014.29.3.361

43. Hunder GG. The use and misuse of classification and diagnostic criteria for complex diseases [Editorial]. Ann Intern Med. 1999;129:417-8. [PMID: 9735071]

44. Aggarwal R, Ringold S, Khanna D, Neogi T, Johnson SR, Miller A, et al. Distinctions between diagnostic and classification criteria? Arthritis Care Res (Hoboken). 2015;67:891-7. [PMID: 25776731] doi:10.1002/acr.22583
Current Author Addresses: Drs. Newberry, FitzGerald, and Shekelle; Ms. Motala; Ms. Booth; Ms. Maglione; Mr. Han; Mr. Tariq; Ms. O’Hanlon; Ms. Shanman; and Ms. Dudley: RAND Corporation, 1776 Main Street, Santa Monica, CA 90401.

Author Contributions: Conception and design: A. Motala, M.A. Maglione.
Analysis and interpretation of the data: J.D. FitzGerald, M. Booth, C.E. O’Hanlon, W. Dudley.
Drafting of the article: M. Booth.
Critical revision for important intellectual content: J.D. FitzGerald, A. Motala, M.A. Maglione, P.G. Shekelle.
Final approval of the article: S.J. Newberry, J.D. FitzGerald, A. Motala, M. Booth, M.A. Maglione, D. Han, A. Tariq, C.E. O’Hanlon, R. Shanman, W. Dudley, P.G. Shekelle.
Statistical expertise: M. Booth.
Obtaining of funding: P.G. Shekelle.
Collection and assembly of data: A. Motala, M. Booth, M.A. Maglione, C.E. O’Hanlon, R. Shanman, W. Dudley.