

# Large Epidemiologic Studies of Gout: Challenges in Diagnosis and Diagnostic Criteria

José M. A. Wijnands · Annelies Boonen ·  
Ilja C. W. Arts · Pieter C. Dagnelie ·  
Coen D. A. Stehouwer · Sjef van der Linden

Published online: 17 December 2010  
© The Author(s) 2010. This article is published with open access at Springerlink.com

**Abstract** Large epidemiologic studies of gout can improve insight into the etiology, pathology, impact, and management of the disease. Identification of monosodium urate monohydrate crystals is considered the gold standard for diagnosis, but its application is often not possible in large studies. Therefore, under such circumstances, several proxy approaches are used to classify patients as having gout, including ICD coding in several types of databases or questionnaires that are usually based on the existing classification criteria. However, agreement among these methods is disappointing. Moreover, studies use the terms *acute*, *recurrent*, and *chronic gout* in different ways and without clear definitions. Better definitions of the different manifestations and stages of gout may provide better insight into the natural course and burden of disease and can be the basis for valid approaches to correctly classifying patients within large epidemiologic studies.

**Keywords** Gout · Epidemiology · Classification

---

J. M. A. Wijnands (✉) · A. Boonen · S. van der Linden  
Department of Internal Medicine, Division of Rheumatology,  
Maastricht University Medical Center and Research School  
Caphri, Maastricht University,  
P.O. Box 5800, NL-6202 AZ, Maastricht, The Netherlands  
e-mail: j.wijnands@maastrichtuniversity.nl

I. C. W. Arts · P. C. Dagnelie  
Department of Epidemiology,  
Maastricht University,  
Maastricht, The Netherlands

C. D. A. Stehouwer  
Department of Internal Medicine,  
Maastricht University Medical Center,  
Maastricht, The Netherlands

## Introduction

Epidemiology can be defined as the study of the occurrence and distribution of disease and its determinants [1]. In a broader approach, the areas of research in epidemiology include disease definition, occurrence, causation, outcome, management, and prevention. The occurrence of a disease may be studied in relation to factors that can identify or predict the disease (diagnostic factors) or are thought to influence its occurrence (eg, prognostic or etiologic factors). Furthermore, the association between a particular intervention and a change in the occurrence of the disease is of importance [2].

In a recent review on rheumatic diseases, Gabriel and Michaud [3•] acknowledged that until recently, very few studies had been conducted on the epidemiology of gout. Notwithstanding the rising incidence [4, 5], the burden of disease associated with gout and the frequency of gout as a comorbidity in patients with multiple morbidities highlight the importance of gaining a better understanding of the etiology, pathology, and management of gout through epidemiologic studies.

This article focuses on the challenges of epidemiologic studies that aim to estimate the occurrence of gout in terms of prevalence or incidence. Such studies require large sample sizes, especially in heterogeneous populations. However, it is not only sample size that needs consideration when conducting epidemiologic studies on the occurrence of gout. This article considers some other deliberations. First, the definitions used in studies of gout are discussed, as well as the differences between diagnostic and classification criteria. Next, the criteria regularly used in epidemiologic studies are reviewed and critically appraised. Finally, several additional challenges encoun-

tered when interpreting results of epidemiologic research on gout are addressed.

### Definition of Disease

The gold standard to diagnose gout is to demonstrate the presence of monosodium urate monohydrate (MSU) crystals in synovial fluid at the time the patient experiences a gout attack [6]. However, this is easier said than done in clinical practice, as gout patients are often seen by general practitioners [7] or specialists other than rheumatologists, who rarely perform a synovial fluid analysis to demonstrate urate crystals [8]. Reasons for nonperformance include lack of expertise, limited access to polarizing microscopes, lack of time, or the concern of getting a “dry tap” [9]. If a joint aspiration is performed, several difficulties remain, as both false-positive and false-negative results may occur [10]. In some cases, cholesterol crystals may appear as needle-shaped birefringent crystals [11, 12]. Furthermore, it is debatable whether one can diagnose a patient as having gout when only one or two MSU crystals are seen, or whether one should perform a second joint aspiration if no crystals are detected the first time. On the other hand, Lumberras et al. [13] showed that when observers are trained in crystal detection and identification, their results are usually consistent.

Although synovial fluid examination remains best practice in clinical medicine, in the context of epidemiologic studies, this may not be feasible, especially in view of the intermittent nature of gout and the sample sizes that are needed in population studies.

In studies, the terms *gout flare*, *chronic gout*, and *acute gout* are regularly used. However, comparing their definitions, if they are provided at all, these terms seem to be used inconsistently by different authors. To facilitate the comparability of study results, clear definitions of the various manifestations and stages of gout therefore are needed.

Taylor et al. [14••] proposed key components of a standard definition of gout flares using the Delphi methodology. The final list of elements includes a swollen, tender, and warm joint; patient self-report of pain and global assessment; time to maximum pain level; time to complete resolution of pain; functional status; and an acute-phase marker. This definition specifically aims to be used in clinical trials.

Another frequently used term is chronic gout. Interestingly, although a core set of outcome domains to be used in clinical trials was proposed by OMERACT (Outcome Measures in Rheumatology Clinical Trials) [15], the group did actually not provide a clear definition

of chronic gout but agreed on serum urate, gout flare recurrence, tophus regression, joint damage imaging, health-related quality of life, musculoskeletal function, patient global assessment, participation, safety, and tolerability as core outcome domains to be assessed in trials on gout.

For Choi et al. [16], acute gout is typically intermittent, whereas chronic tophaceous gout develops after years of acute intermittent gout. However, they point out that tophi can also be part of the initial presentation. This is in line with the American College of Rheumatology (ACR) criteria for acute gout, which incorporate the item “more than one attack of acute arthritis” as well as suspicion of a tophus.

Others state that after a certain undefined number of attacks, a patient has reached a stage called *recurrent gout*. Then the attacks come more frequently and last longer. If a patient cannot recover from the flares, it becomes chronic gout. In that case, there is an almost permanent state of inflammation and pain. Some authors have a more inclusive definition of chronic gout, incorporating all patients who have had more than one attack of acute gout. Acute gout, in that case, is synonymous with gout flare. The underlying reasoning is that once a patient has had a flare, a persistent metabolic disorder exists.

In addition to these deliberations, one might think of gout as a continuum of increasing severity. Currently, a clear definition of severity is lacking. A first step would be to define the domains that are of importance when deciding on the severity of gout. Of interest is a recent study that explored which variables are associated with patients’ as opposed to physicians’ assessment of gout severity [17]. It was found that physicians base their judgment of severity on the presence of tophi, frequency of gout attacks during the past year, recent serum uric acid levels, and rheumatologist utilization, whereas patients’ judgment of severity is associated with concerns regarding gout during an attack and the time since the last gout attack. It will be a challenge to try to measure each of these domains and to define thresholds to distinguish levels of severity based on the selected domains. This may also include addressing issues such as total load of uric acid and total load of tophi.

In summary, the question remains whether intermittent or recurrent gout must be distinguished from chronic gout and, if so, at which point acute or recurrent gout develops into chronic gout. Can we speak of chronic gout after a number of attacks or after a specific period of recurrent attacks, or only when a patient has bone destruction and chronic synovitis, even during remission of acute flares? Should we make a distinction between chronic gout and tophaceous gout, and should we distinguish different levels of severity of gout?

Currently, no clear answers to these questions exist, and the lack of insight into the natural course of gout among all patients complicates this issue. Some patients may not recall attacks of symptomatic gout, and patients with tophaceous gout may no longer experience acute attacks. This underlines the problem of assessing the true prevalence of gout. Moreover, the gold standard for diagnosing gout does not discriminate between different “stages” of gout. Clearly, a need exists for consensus on definitions that help distinguish the different manifestations of the disease.

### Diagnostic Criteria Versus Classification Criteria

The title of this article may seem contradictory because in epidemiologic studies, no diagnostic criteria other than classification criteria are applied. Classification criteria aim to define homogeneous groups of patients with a particular disease. These criteria can be used to select patients for clinical (interventional) studies, to compare the results of clinical trials, or to assess the occurrence of a disease in epidemiologic studies [18]. In contrast to diagnostic criteria, classification criteria do not have the purpose of early detection of a disease in an individual patient [19]. Instead, classification criteria are used to detect established cases.

As for diagnostic tests, calculating the sensitivity and specificity assesses the usefulness of criteria. Sensitivity is the percentage of individuals with a certain disease correctly classified as “ill” (true positives). The percentage of individuals without the disease correctly labeled as “not ill” is the specificity (100% - percentage of false positives). If the sensitivity and specificity of criteria are both 100%, diagnostic and classification criteria are the same [19]. Note that diagnostic criteria would require sufficient sensitivity in early stages of the disease to enable early diagnosis. However, the nature of medicine makes it unlikely that there will ever be tests that offer 100% sensitivity and specificity. Therefore, misclassification poses a challenge, and the type of misclassification that is least desirable will depend on the setting in which the test is applied.

In health care, physicians must identify which disease a patient has rather than whether a disease exists at all [19]. They do not want to misdiagnose a patient and may prefer high sensitivity against acceptable specificity. In contrast, in epidemiologic studies in large populations, to study homogeneous groups that are likely to have the diagnosis of interest but do not include many false positives, the researchers must balance specificity and sensitivity and will often sacrifice part of sensitivity against better specificity.

Of course, any misclassification, which is common in classification criteria, is undesirable [19]. An approach to minimize misclassification is the use of cut-off points. In deciding on a cut-off point, one has to choose between a sensitive approach—involving false positives—and a more specific approach that results in a more homogeneous group and more false negatives [19].

Classification criteria often are developed by comparing groups with the disease of interest with control patients having other (usually related or resembling) diseases that should be taken into account in the differential diagnosis. However, one must keep in mind that if these criteria are applied in population studies, the positive predictive value (PPV) may decrease, especially when the prevalence of the disease of interest is low. The PPV is defined as the number of individuals with a true positive test result divided by all individuals with a positive test result (true positives + false positives). In other words, it indicates the probability that in case of a positive test, the patient truly has the specified disease. The value of PPV depends on the prevalence of the disease of interest in the particular setting and will decrease when the prevalence goes down, due to the increasing number of false positives.

Thus, when applying the criteria for gout, a disease with a relatively low prevalence at the general population level, it is important to keep in mind that the estimated prevalence may be overestimated due to the unintended inclusion of false positives.

### Overview of Criteria

In this section, criteria to assess the prevalence of gout in epidemiologic studies are described. For this purpose, PubMed was searched using the search terms “gout,” “incidence,” “prevalence,” and “epidemiology.” Only original articles describing the prevalence and incidence of gout were considered. The EULAR (European League Against Rheumatism) criteria, which are purely diagnostic criteria and intended for use in individual patients with arthritis and not for use in groups, are excluded.

In 1963, the Rome criteria for gout were proposed during a symposium on population studies (Table 1). These and the 1966 New York criteria, which are a modification of the Rome criteria, are based on expert opinion and aimed for application in epidemiologic studies (Table 1) [18]. They rely heavily on the presence of tophi and the observation of MSU crystals in synovial fluid, which causes some feasibility issues. This is probably why both criteria sets are rarely used in large epidemiologic studies. The Rome criteria have only been used in two population studies assessing the prevalence or incidence of gout, once as interview [20] and once as questionnaire [21], and both

**Table 1** Classification criteria for gout

Rome criteria	New York criteria	American College of Rheumatology criteria
Two of the following 4 criteria must be present to make a diagnosis of gout:	Urate crystals in synovial fluid or tissue or presence of at least 2 of the following:	Preliminary criteria for the classification of the acute arthritis of primary gout
1. Serum uric acid level $\geq 7.0$ mg/dL in men, or $\geq 6.0$ mg/dL in women	1. History or observation of at least 2 attacks of painful limb swelling with remission within 1–2 week	A. Monosodium urate monohydrate crystals in synovial fluid, <i>or</i>
2. Tophus	2. History or observation of podagra	B. Tophus, <i>or</i>
3. Urate crystals in synovial fluid or tissues	3. Presence of tophus	C. Presence of at least 6 of the following:
4. History of attacks of painful joint swelling of abrupt onset with remission within 1–2 week	4. History or observation of a good response to colchicines (major reduction in objective signs of inflammation within 24 h of onset of therapy)	1. More than 1 attack of acute arthritis
		2. Maximal inflammation developed within 24 h
		3. Monoarthritis attack
		4. Redness observed over joints
		5. First metatarsophalangeal joint painful or swollen
		6. Unilateral first metatarsophalangeal joint attack
		7. Unilateral tarsal joint attack
		8. Tophus (suspected)
		9. Hyperuricemia
		10. Asymmetric swelling within a joint on radiograph
		11. Joint fluid culture negative for organisms during attacks

times in combination with a physical examination. The New York criteria have been used in several population studies in the same way as the Rome criteria [21–24].

It should be noted that for use in population surveys, the items that make up criteria likely need to be rephrased into questions that are answerable by patients using questionnaires or participating in interviews.

Currently, the most frequently used methods to identify people with gout in epidemiologic studies are the ACR criteria—former American Rheumatism Association criteria—using an interview approach and the *ICD-9*.

The ACR criteria for gout have been developed to achieve a uniform system for reporting and comparing data from studies (Table 1) [6]. They have been developed by comparing different sets of criteria among gout patients and patients with classic rheumatoid arthritis of 2 years' or less duration, definite or classic rheumatoid arthritis of more than 2 years' duration, pseudogout, or acute septic arthritis. All have been diagnosed by rheumatologists. As such, the ACR criteria for gout focus on acute arthritis of primary gout and can be used in single patients as well as in population surveys [6].

In large studies on the occurrence of gout, the ACR criteria are often applied by interviewing patients with or without a standardized questionnaire, or by chart review of medical records. Although the ACR criteria were developed for the diagnosis of acute gout, they also have been used to identify so-called chronic gout, when patients fulfill the item tophi or radiographic abnormalities. Compared with

the Rome and New York criteria, the ACR criteria rely less on the presence of tophi or identification of MSU crystals and even allow classification based on clinical criteria alone. Malik et al. [25] applied the ACR, New York, and Rome criteria in patients who had joint effusions in the setting of a rheumatology clinic. They asked patients whether they had experienced any of the clinical features of these three sets of criteria. The researchers found the highest specificity (89%) and PPV (77%) for the Rome criteria. However, the criteria were slightly less sensitive (67%). The New York criteria showed sensitivity and PPV of 70% and specificity of 83%. The ACR criteria (6 of 12 clinical items) had 70% sensitivity and 79% specificity and a PPV of only 66%. Clearly, one should not extrapolate such findings to an epidemiologic population study, because the PPV varies with the pretest probability, which, as mentioned previously, is highly dependent on the prevalence of the disease. Janssens et al. [26] compared the ACR criteria with synovial fluid analysis as a gold standard in monoarthritic patients presenting to primary care. Only patients who were suspected of having gout were included in the study. They found a PPV and a sensitivity of 80%, while specificity was 64%. According to Janssens et al. [26], these findings stress the importance of interpreting with caution the results of gout studies that made use of the ACR criteria.

A common method used to estimate the prevalence of gout is the use of large medical databases that have registered diseases by *ICD-9* coding. Examples of data-

bases used in gout research include medical patients' record systems, administrative claims, and insurance programs. Advantages of such databases are the large numbers available at low expenses and the efficient time investment. A disadvantage is that it remains unclear how the diagnosis was made by the variety of health professionals and how to generalize the results because the denominator is often unclear. Malik et al. [27] evaluated the possibility of documenting the accuracy of the *ICD-9* code for gout in three databases (National Patient Care Database, Pharmacy Benefits Management Database, and the Clinical and Administrative Database) by identifying patients with two *ICD-9*-coded encounters for gout during a 6-year time period. They found that identifying the items of the ACR, New York, or Rome criteria in medical records could not validate the majority of gout diagnoses recorded by *ICD-9*. This discrepancy may be caused by inadequate documentation in medical records, inaccurate diagnostic coding, or the inappropriateness of current criteria. According to Malik et al. [27], it is the poor documentation in medical records rather than inaccurate diagnostic coding. Harrold et al. [28] analyzed a random sample of medical records of patients with two or more coded diagnoses of gout from four managed care plans. The PPV of two or more ambulatory claims (during a time period of 5 years) for a diagnosis of gout was assessed using the investigators' rating of the presence or absence of definite or probable gout as the gold standard. The PPV turned out to be 61%. Substantial improvement in the PPV was not achieved by increasing the number of visits to three or four. Explanations for the disappointing PPV include the ambiguity of a diagnosis of gout compared with, for example, a more firm diagnosis of myocardial infarction; the assignment of an *ICD* code before the diagnosis was firmly established; the underutilization of synovial fluid analysis; and inadequate documentation in medical records [28].

Self-reported disease, sometimes completed with information from other medical sources, is often applied in epidemiologic studies of gout. However, it is difficult to distinguish between questionnaires that inquire about physician-diagnosed gout and questionnaires that inquire about manifestations typical of gout based on the existing criteria described above. Furthermore, such questionnaires vary in the time frame in which gout occurred, which can be one or more attacks at some point in the past or several attacks in a specific (limited) period of time preceding the survey. Miller et al. [29] analyzed the agreement between self-reported diseases and *ICD-9* coding. These authors indicated that self-report is fairly reliable. However, only 50% of self-reported cases of arthritis could be confirmed by *ICD* coding [29]. Reasons for this lack of agreement may be that responders have interpreted a question

incorrectly or have recalled a diagnosis that was not actually established or was recalled inaccurately. However, if patients are not receiving medication or other treatment, the diagnostic code for a certain condition may not be written down in the record [29]. Miller et al. [29] pointed out that acute events that occurred in the past and conditions that are episodic in nature are not always captured if the reviewing period is too short.

It should be noted that questionnaires are often operationalized through an interview approach. Although self-completed questionnaires may cover a large population in a relatively short time period at low cost, the downside is a possible low response rate. Using an interview approach, it is possible to ensure that all questions are answered in the correct manner. However, this method is more prone to interviewer bias and interviewer variability [1]. Bergmann et al. [30] reported that the agreement between a face-to-face interview and a self-administered questionnaire was moderate ( $\kappa$ , 0.61). Less serious, less defined, or less persistent diseases such as gout may be perceived by patients as not being important enough to report in questionnaires [30].

Of interest might be the diagnostic rule for acute gouty arthritis recently developed by Janssens et al. [31]. It is intended to be applied in primary care and obviate joint aspiration. Based on validated clinical variables using synovial fluid analysis as a reference test, a multivariate logistic regression model was developed. Hereafter, they developed two models based on external knowledge and availability of the tool in clinical practice. Their final model includes seven variables: male sex, previous patient-reported arthritis attack, onset within 1 day, joint redness, first metatarsophalangeal joint involvement, hypertension, or one or more cardiovascular diseases and serum uric acid level exceeding 5.88 mg/dL (0.35 mmol/L). Although developed for use in primary care, the diagnostic rule may be useful in a research setting. However, it is not known how well this model performs in a population study (work in progress).

## Interpretation of Results

In addition to the above considerations that are critical in the appraisal of data on the occurrence of disease, several other issues merit consideration when appraising the results of such studies [32]. First is the question of which type of epidemiologic measure of occurrence was applied. The nature of disease will influence the relevant study design and measure of occurrence that is most informative [1]. As acute flares of arthritis mainly characterize gout, the point prevalence estimate (Table 2) is not likely of primary interest. In fact, in a cross-sectional population study, the

**Table 2** Conceptual framework of incidence and prevalence in studies of gout

Incidence	
Cumulative incidence	To be assessed in a cohort: number of new cases of gout per year divided by the population at risk (ie, all cohort members who at the initiation of the cohort or at the start of the year of incidence assessment had never experienced any manifestation of gout)
Incidence density	To be assessed in a dynamic population (eg, the inhabitants of New York City): number of new cases of gout per year divided by the number of person-years of individuals at risk. One person-year is defined as 1 person who is at risk for a 1-year period (eg, if an individual gets gout after 3 months, he or she is counted as 0.25-person years in the denominator). Thus, the denominator of incidence density (ie, the number of person-years in a dynamic population) is not only determined by the changing size of the total population (eg, accounting for individuals entering and leaving the municipality of New York, as well as newborns and deaths) but also by the number of hitherto-healthy individuals who for the first time get gout during the observation period and are therefore (from that moment onward) no longer “at risk” of newly getting gout
Prevalence	
Point prevalence	Number of cases of gout in the study population at a given point in time divided by the total study population. This comprises the (usually few) individuals who suffer from an acute attack of gout at the concerned point in time together with all those who then have chronic (tophaceous or nontophaceous) gout
Period prevalence	Number of cases of gout in the study population during a specified period of time divided by the mean size of the total study population over the concerned period. This comprises all individuals who have experienced an acute attack of gout during that period together with all cases of chronic (tophaceous or nontophaceous) gout

chance that someone will suffer from a gout attack at exactly the time of the survey is low. In this case, the period prevalence, which represents individuals who experienced one or more episodes over a specified period preceding the survey (Table 2), will be more informative. Only if “chronic gout” would be described in terms of persistent joint inflammation, presence of irreversible joint damage, or presence of tophi would the point prevalence be interesting.

Another important concept in epidemiologic studies is the incidence or incidence rate, which refers to the number of new cases of gout in a population (Table 2). Cumulative incidence refers to new cases of gout per year divided by all members of a cohort (ie, a closed population) who are at risk (ie, who never experienced any signs of gout before the observation period). In contrast, incidence density refers to new cases of gout per person-year in a dynamic population, such as the inhabitants of a region or municipality (Table 2).

It is also important to carefully consider the population that has been studied [2]. As for any epidemiologic study, the sample should be a correct representation of the population of interest; this requires insight into the sampling frame and participation rate. The participation rate should be at least as high as 80%; however, a rate between 60% and 80% with a description of the non-responders is often considered acceptable. Furthermore, to guarantee the representativeness of the samples in studies on gout, it is important to take into account sources of (selection) bias, such as the age, sex, and race of the study population. Determining the prevalence in a preponderant older male population will overestimate the occurrence of gout in the general population. One also should be aware of confounding factors such as alcohol consumption, body mass index, and comorbidities. Obesity, diabetes, and

hypertension are common among patients with gout [33], and the prevalence of the metabolic syndrome is higher than in patients without gout [34, 35].

## Conclusions

Although the pathophysiology of gout is relatively well-understood, it is surprisingly difficult to define good classification criteria for use in large population studies to validly assess the prevalence and burden of gout. This is partially due to the nature of the disease, which is typically intermittent, which limits the ability to use MSU crystals as the gold standard in large epidemiologic studies.

Also challenging is the absence of clear insight into the natural course of the disease, which would require better definitions of the manifestations and stages of gout. Although there is general agreement that gout is likely the result of a longstanding metabolic disorder that eventually leads to clinically manifest gout, it is less clear how many patients will progress to tophaceous gout and develop joint damage.

In view of the aforementioned considerations, literature data on the prevalence of gout are surprisingly consistent. In developed countries, estimates vary between 1% and 2% [36]. Nevertheless, as discussed previously, different levels of misclassification must have occurred in these studies. This will hamper interpretation of the results, especially in light of risk factors and comorbidities associated with gout. More precise estimates of the prevalence and burden of gout require addressing the validity of classification criteria and proper definitions of the various manifestations and stages of severity of gout.

McAdams et al. [37] reported recently that self-report of physician-diagnosed gout has good reliability and sensitivity and that this method may seem appropriate for epidemiologic studies.

**Disclosure** No potential conflicts of interest relevant to this article were reported.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

Papers of particular interest, published recently, have been highlighted as:

- Of importance
- Of major importance

1. Silman AJ: Epidemiological studies: a practical guide. Great Britain: Cambridge University Press; 1995.
2. Bouter LM, van Dongen MCJM: Epidemiologisch onderzoek: opzet en interpretatie. Houten: Bohn Stafleu Van Loghum; 1995.
3. • Gabriel SH, Michaud K: Epidemiological studies in incidence, prevalence, mortality, and comorbidity of the rheumatic diseases. *Arthritis Res Ther* 2009, 11:229. *This review provides a clear overview of the epidemiology (incidence, prevalence, and mortality) of the rheumatic diseases. Furthermore, the article describes the role of comorbidity in determining outcome in rheumatoid arthritis.*
4. Arromdee E, Michet CJ, Crowson CS, et al: Epidemiology of gout: is the incidence rising? *J Rheumatol* 2002, 29:2403–2406.
5. Wallace KL, Riedel AA, Joseph-Ridge N, Wortmann R: Increasing prevalence of gout and hyperuricemia over 10 years among older adults in a managed care population. *J Rheumatol* 2004, 31:1582–1587.
6. Wallace SL, Robinson H, Masi AT, et al: Preliminary criteria for the classification of the acute arthritis of primary gout. *Arthritis Rheum* 1977, 20:895–900.
7. Pal B, Foxall M, Dysart T, et al: How is gout managed in primary care? A review of current practice and proposed guidelines. *Clin Rheumatol* 2000, 19:21–25.
8. Janssens HJEM, Franssen J, van de Lisdonk EH, et al: A diagnostic rule for acute gouty arthritis in primary care without joint fluid analysis. *Arch Intern Med* 2010, 170:1120–1126.
9. Dore RK: The gout diagnosis. *Clev Clin J Med* 2008, 75:S17–S21.
10. Von Essen R, Hölttä AMH, Pikkarainen R: Quality control of synovial fluid crystal identification. *Ann Rheum Dis* 1998, 57:107–109.
11. Selvi E: Needle-shaped crystals are not always urate crystals: Comment on the clinical image by Slobodin et al. *Arthritis Rheum* 2009, 60:3858.
12. Von Essen R, Hölttä AMH: Quality control of the laboratory diagnosis of gout by synovial fluid microscopy. *Scand J Rheumatol* 1990, 19:232–234.
13. Lumberras B, Pascual E, Frasquet J, et al: Analysis for crystals in synovial fluid: training of the analysts results in high consistency. *Ann Rheum Dis* 2005, 64:612–615.
14. •• Taylor WJ, Shewchuk R, Saag KG, Schumacher HR, et al: Toward a valid definition of gout flare: Results of consensus exercises using Delphi methodology cognitive mapping. *Arthritis Rheum* 2009, 61:535–543. *This study used the Delphi method to identify a short list of potential features of gout flare.*
15. Schumacher HR, Edwards LN, Perez-Ruiz F, et al: Outcome measures for acute and chronic gout. *J Rheumatol* 2005, 32:2452–2455.
16. Choi HK, Mount DB, Reginato AM: Pathogenesis of gout. *Ann Intern Med* 2005, 143:499–516.
17. Sarkin AJ, Levack AE, Shieh MM et al: Predictors of doctor-rated and patient-rated gout severity: gout impact scales improve assessment. *J Eval Clin Pract* 2010 Aug 15 (Epub ahead of print).
18. Johnson SR, Goek ON, Singh-Grewal D, et al: Classification criteria in rheumatic diseases: a review of methodologic properties. *Arthritis Rheum* 2007, 57:1119–1133.
19. Fries JF, Hochberg MC, Medsger TA, et al: Criteria for rheumatic disease. *Arthritis Rheum* 1994, 37:454–462.
20. Mikkelsen WM, Dodge HJ, Duff IF, Kato H: Estimates of the prevalence of rheumatic diseases in the population of Tecumseh, Michigan, 1959–60. *J Chronic Dis* 1967, 20:351–369.
21. O'Sullivan JB: Gout in a New England town. *Ann Rheum Dis* 1972, 31:166–169.
22. Chen S, Du H, Wang Y, Xu L: The epidemiology study of hyperuricemia and gout in a community population of Huangpu District in Shanghai. *Chin Med J* 1998, 111:228–230.
23. Darmawan J, Valkenburg HA, Muirden KD, Wigley RD: The epidemiology of gout and hyperuricemia in a rural population of Java. *J Rheumatol* 1992, 19:1595–1599.
24. Wigley RD, Prior IA, Salmond C, et al: Rheumatic complaints in Tokelau. II. A comparison of migrants in New Zealand and non-migrants. The Tokelau Island migrant study. *Rheumatol Int* 1987, 7:61–5.
25. Malik A, Schumacher R, Dinnella JE, et al: Clinical diagnostic criteria for gout. *J Clin Rheum* 2009, 15:22–24.
26. Janssens HJEM, Janssen M, van de Lisdonk EH, et al: Limited validity of the American College of Rheumatology criteria for classifying patients with gout in primary care. *Ann Rheum Dis* 2010, 69:1255–1256.
27. Malik A, Dinnella JE, Kwok K, Schumacher HR: Poor validation of medical record ICD-9 diagnoses of gout in a Veterans Affairs database. *J Rheumatol* 2009, 36:1–4.
28. Harrold LR, Saag KS, Yood RA, et al: Validity of gout diagnoses in administrative data. *Arthritis Rheum* 2007, 57:103–108.
29. Miller DR, Rogers WH, Kazis LE, et al: Patients' self-report of diseases in the Medicare Health Outcomes Survey based on comparisons with linked survey and medical data from the Veterans Health Administration. *J Ambul Care Manag* 2008, 31:161–177.
30. Bergmann MM, Jacobs EJ, Hoffmann K, Boeing H: Agreement of self-reported medical history: comparison of an in-person interview with a self-administered questionnaire. *Eur J Epidemiol* 2004, 19:411–416.
31. • Janssens HJ, Franssen J, van de Lisdonk EH, et al: A diagnostic rule for acute gouty arthritis in primary care without joint fluid analysis. *Arch Intern Med* 2010, 170:1120–1126. *This study developed the first diagnostic rule for acute gouty arthritis in primary care.*
32. Sanderson S, Tatt ID, Higgins JPT: Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol* 2007, 36:666–676.

33. Annemans L, Spaepen E, Gaskin M, et al: Gout in the UK and Germany: prevalence, comorbidities and management in general practice 2000–2005. *Ann Rheum Dis* 2008, 67:960–966.
34. Choi HK, Atkinson K, Karlson EW, Curhan G: Obesity, weight change, hypertension, diuretic use, and risk of gout in men: the health professionals follow-up study. *Arch Intern Med* 2005, 165:742–748.
35. Inokuchi T, Tsutsumi Z, Takahashi S, et al: Increased frequency of metabolic syndrome and its individual metabolic abnormalities in Japanese patients with primary gout. *J Clin Rheum* 2010, 16:109–112.
36. Richette P, Bardin T: Gout. *Lancet* 2010, 375:318–328.
37. McAdams MA, Maynard JW, Baer AN, et al: Reliability and sensitivity of the self-report of physician-diagnosed gout in the campaign against cancer and heart disease and the atherosclerosis risk in the community cohorts. *J Rheumatol* 2011, In press.