Is ultrasound changing the way we understand rheumatology? Including ultrasound examination in the classification criteria of polymyalgia rheumatica and gout.

Catalin Codreanu, Luminita Enache

Center for Rheumatic Diseases, Bucharest, Romania

Abstract

Ultrasonography (US) is widely used in the diagnosis of rheumatic conditions, and its value for the classification criteria of rheumatic diseases has been recently suggested. According to the EULAR/ACR provisional criteria for polymyalgia rheumatica, adding US to the clinical and serological features will significantly improve the sensitivity of proposed criteria. The ability of high resolution US to detect crystalline deposits of monosodium urate in joints and soft tissues is well recognized. For the first time, the new 2014 ACR/EULAR set of proposed criteria for gout includes advanced imaging techniques for the detection of disease: US and dual-energy computed tomography. Due to low costs and affordability, use of US evaluation for patients with suspected gout will increase both specificity and sensibility of classification criteria. The recent inclusion of US in the classification criteria of various rheumatic diseases, such as PMR and gout, implies that this imaging technique is not only useful as a valued diagnostic tool for individual cases, but also on a larger scale, it will improve doctors’ ability to classify diseases. Its use is thus changing our understanding of rheumatic diseases allowing further advances in research and clinical practice.

Keywords: rheumatic diseases, ultrasonography, classification criteria, polymyalgia rheumatica, gout

Diagnostic imaging in rheumatology enables the identification and assessment of pathological conditions, allowing improved accuracy of early diagnosis as well as monitoring disease activity [1]. For many years conventional radiology has been used as a gold standard of imaging in rheumatic diseases. Only recently ultrasonography (US) and magnetic resonance imaging (MRI) have been implemented more widely in the clinical practice of rheumatology [2]. Both imaging techniques perform better than conventional radiography in soft tissue assessment and earlier detection of bone erosions [3]. Significantly less expensive than MRI, more readily available, and patient-friendly, US could be considered an ideal tool for improving the diagnostic ability of medical imaging in clinical practice for rheumatic diseases [1].

There is no doubt that US is a valuable tool for diagnosing rheumatic diseases in individual cases. Nevertheless, the recent sophistication and standardization [4,5] of its use could potentially lead to its inclusion in the classification criteria for various rheumatic diseases. The value of this imaging tool could thus be significantly increased by allowing its use on a larger scale.

Classification and/or diagnostic criteria have been validated in several musculoskeletal diseases. Even though classification and diagnostic criteria may consist of similar clinical, laboratory, or imaging parameters, they differ in their aims and requirements. Classification criteria enable scientific research by grouping patients with the same disease in homogeneous groups, thus making comparable the data obtained by researchers in different populations [6]. On the other hand, diagnostic criteria are used by physicians to make diagnoses in individual patients, so they should have high positive predictive value to ensure that if a patient fulfills the criteria there is a high chance to have the disease [6]. Nevertheless, in the absence of diagnostic criteria, classification criteria are sometimes used for diagnosis [7].

Whereas the American Rheumatism Association (ARA) 1987 Revised Classification Criteria for Rheu-
matoid Arthritis (RA) [8] included the detection of bone erosions using conventional radiography, the new 2010 American College of Rheumatology (ACR) / European League Against Rheumatism (EULAR) RA classification criteria excluded imaging evaluation from the accepted criteria, due to the inability of conventional radiology to detect early structural changes [9].

It was not until the publication of these criteria that the use of US, particularly power Doppler US, has been shown to improve the sensitivity and specificity in the diagnosis of RA according to the 2010 ACR/EULAR classification criteria [10]. Additionally, in the setting of very early inflammatory arthritis, the power Doppler signal is able to predict the development of persistent inflammatory arthritis [11] and can facilitate early diagnosis of RA.

Following from the earlier points, the recently developed diagnostic criteria for two common rheumatic diseases will be discussed: polymyalgia rheumatica and gout. The inclusion of US in the accepted criteria for both conditions is allowing us to reconsider the value of this imaging technique in modern rheumatology.

**Polymyalgia Rheumatica**

Polymyalgia Rheumatica (PMR), a common inflammatory condition of elderly people [12], is characterized by pain in both shoulders and hips, elevated acute-phase reactants and a prompt response to corticosteroid therapy [13]. With these clinical features lacking specificity, it is often difficult to diagnose or classify subjects suspected of PMR [14].

In response to the lack of standardized classification criteria for PMR, in 2005 an international working group [15,16] convened to develop new EULAR/ACR provisional classification criteria for PMR, being published in 2012 [17].

The target population for the new PMR criteria is made up of patients who have met three required criteria: 1) aged 50 or older; 2) presentation with new-onset (<12 weeks) bilateral shoulder pain; 3) abnormal acute phase response, either CRP or ESR (limiting the applicability of the new criteria to patients with biologic signs of inflammation, in part overcome by the use of the more sensitive CRP levels to measure inflammation).

The scoring algorithm includes several criteria: four clinical/serological and two US criteria (table I).

The clinical or serological features are sufficient for recognizing PMR according to these classification criteria; however, the addition of US examination focused on shoulders and hips significantly increased their specificity [17].

<table>
<thead>
<tr>
<th>Optional classification criteria</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morning stiffness &gt;45 minutes</td>
<td>2</td>
</tr>
<tr>
<td>Hip pain, limited range of motion</td>
<td>1</td>
</tr>
<tr>
<td>Normal Rheumatoid factor (RF) or Anti–citrullinated protein antibodies (ACPA)</td>
<td>2</td>
</tr>
<tr>
<td>Absence of other joint pain</td>
<td>1</td>
</tr>
</tbody>
</table>

**Ultrasound criteria**

At least 1 shoulder with:

- Subdeltoid bursitis and/or biceps tenosynovitis and/or glenohumeral synovitis
- At least 1 hip with: synovitis and/or trochanteric bursitis
- Both shoulders with: subdeltoid bursitis, biceps tenosynovitis or glenohumeral synovitis

The optimal cut-off point for classifying a case as PMR is 4 (without US) and 5 (including US).

For the purposes of establishing the usefulness of the addition of US to the new criteria, systematic US evaluations of both shoulders and hips were made in order to assess patients for features previously reported to be associated with PMR: bicipital tenosynovitis, subacromial and subdeltoid bursitis, trochanteric bursitis, glenohumeral and hip effusion [18-20] (fig 1).

According to the criteria proposed by Dasgupta et al, the diagnostic value of US has not been taken into account on its own, but only in addition to the clinical and serological features included in the final set. Even though in the context of the new criteria, the use of US was optional, it has shown to improve the specificity of criteria from 68% to 89% for discriminating shoulder conditions from PMR [17].

Recently, Macchioni al [21] published a study that compares the performance of the new ACR/EULAR provisional criteria for PMR with the previously published criteria. They included 136 patients with PMR and 149 controls (mostly with RA) and tested the performance of the new 2012 EULAR/ACR provisional classification criteria...
for PMR compared with the former PMR diagnostic/classification criteria. US of shoulders and hips was performed at the first and subsequent visits according to a standardized protocol by the same rheumatologist in order to reduce operator-dependent variability and to improve the homogeneity of US testing [22]. According to this study, the most sensitive criteria were the new 2012 EULAR/ACR classification criteria (92.6%). The results replicated the findings of Dasgupta et al [17], but with the addition of US causing an even greater increase in specificity, in this case from 81.5% to 91.3%. Additionally, the study revealed that shoulder and hip US examination increased the specificity of the new EULAR/ACR criteria not only in distinguishing between PMR and total controls, but also in differentiating between RA and PMR patients [21].

Similar findings resulted from a recent meta-analysis [23]: US, especially when used to evaluate shoulders, is confirmed to be a potentially useful tool to improve the classification and management of patients with PMR.

Gout

A review published in the last number of this journal by Fodor et al discussed in detail the ability of high resolution US to detect crystalline deposits of monosodium urate (MSU) in joints and soft tissues, analyzing the value of the method in diagnosis and treatment of gout [24].

According to the 2006 EULAR recommendations for gout management, a definitive diagnosis of gout is granted only by the demonstration of MSU crystals in synovial fluid or tophus aspirates [25].

A similar, but somehow more nuanced approach is used in recently published evidence-based multinational recommendations for the diagnosis and management of gout, endorsed by the 3e (Evidence, Expertise, Exchange) Initiative, a unique multinational collaboration aimed at promoting evidence-based practice in rheumatology [26]. In which circumstances can a diagnosis of gout be made on clinical grounds with or without laboratory tests or imaging, and when is the identification of crystals really necessary [27]? To find an answer to this extremely relevant question, a panel of 78 international rheumatologists stated that the identification of MSU crystals should be performed for a definite diagnosis of gout. However, when the examination of synovial fluid or tissue samples is not feasible, a diagnosis of gout can be supported by a classical clinical presentation with podagra, tophi, or a prompt response to initiation of colchicine treatment, as well as typical findings on advanced imaging techniques, such as US [28-30] and dual-energy computed tomography (DECT) [15,31].

Important advances in gout management have recently been made: the development not only of modern imaging techniques [32-35] that have the ability to recognize gout better and earlier, but also of potent new drugs, including biologic agents. These advances emphasize the need for robust gout classification criteria that need to have acceptably high specificity to ensure that enrolment in trials testing new therapies is targeting patients with definite gout [36]. At the same time, with the rise in the incidence/prevalence of gout worldwide, uniform criteria with appropriate sensitivity and specificity are needed for epidemiological studies [37].

At least three sets of classification criteria for gout have been published: Rome (1963) [38], New York (1966) [39] and the 1977 ARA preliminary criteria [40] and at least four sets of diagnostic criteria or recommendations: EULAR (2006) [25], Mexico (2010) [41], Netherlands (2010) [42] and the 3e gout recommendations (2014) [26]. However, none of the currently available criteria has been adequately validated and none included advanced imaging techniques for the detection of gout [37].

A collaborative EULAR and ACR international project -Gout Classification Criteria Project- has been initiated in order to develop new gout classification criteria, aiming to improve the case definition for gout among both primary and secondary care populations [37]. The intended use of classification criteria in this setting includes case ascertainment for recruitment into clinical studies, as well as epidemiologic studies and therapeutic randomized controlled trials.

The first phase of the project [43] was a two step Delphi questionnaire applied to rheumatologists and patients with gout, in order to identify a comprehensive list of features that might discriminate between gout and other rheumatic musculoskeletal conditions, to be used subsequently for a case-control study to develop and test new classification criteria for gout. Physicians rated a list of potentially discriminating features identified by literature review and expert opinion; a self-generated list of items was used for the patient survey.

Following item generation from the Delphi exercise, two parallel approaches have been used to determine the optimal combination of features to define gout [37]: a prospective multicentre international study including patients with suspected gout [44] and a paper patient exercise where patient profiles representing a spectrum of gout probability have been ranked by an expert panel. Additional data were derived from a systematic review of the diagnostic utility of advanced imaging for gout [45]. A structured consensus process was used to integrate these sources of data into classification criteria that have been validated using a test sample from the multicentre international study. The proposed new preliminary gout classification criteria were presented at the last ACR meeting in Boston [46].
The Study for Updated Gout Classification Criteria (SUGAR) was performed as part of the Gout Classification Criteria Project and was conceived to perform a direct comparison of existing criteria in early disease versus established disease [44]. This was a cross-sectional study including 983 consecutive patients with joint swelling; all patients underwent arthrocentesis or tissue aspiration for polarizing microscopy to identify MSU crystals. Gout was defined by presence of MSU crystals identified by a certified examiner and confirmed in 509 cases (52%). Non-gout was defined as absence of MSU crystals, irrespective of the clinical diagnosis. Early disease was defined as patient-reported onset of symptoms of 2 years or less. Sensitivity across criteria was better in established disease (95.3% vs 84.1%, p<0.001) and specificity was better in early disease (79.9% vs 52.5%, p<0.001). Existing classification criteria for gout have a sensitivity of over 80% in early and established disease but currently available criteria that do not require synovial fluid analysis have inadequate specificity, especially later in the disease. The SUGAR study [44] confirmed the stringent need for gout classification criteria with a better specificity.

The third important component in the Gout Classification Criteria Project was a systematic review and meta-analysis [45], including 11 studies examining the accuracy of imaging features for the classification of gout. The objective of this study was to analyze the usefulness of imaging modalities in the classification of symptomatic gout when compared to MSU crystal confirmation as the gold standard.

Ogdie et al [45] systematically reviewed the published literature concerning diagnostic performance of plain radiography, MRI, US, conventional CT, and DECT. Eleven studies examining the accuracy of imaging features for the classification of gout were included in the analysis, but only three imaging features were examined in more than one study and judged to have significance for inclusion in the proposed criteria: 1) the double contour sign (DCS) on US (fig 2), with a pooled (95% CI) sensitivity of 0.83 and specificity of 0.76; 2) presence of tophus on US, with a pooled (95% CI) sensitivity of 0.65 and specificity 0.80; 3) MSU crystal deposition on DECT, with a pooled (95% CI) sensitivity of 0.87 and specificity of 0.84.

Recognizing the limitations of small patient samples in most studies and the inclusion of long-standing, established cases with gout, available data [45] support the promise that both imaging techniques, US and DECT, showed for the classification of symptomatic gout.

Based on data provided by the Delphi exercise [43], the SUGAR study [44] and the imaging systematic review [45], as well as the Consensus meeting held during the 2014 EULAR Congress in Paris, a set of new preliminary classification criteria for gout have been proposed [46].

In order to apply new criteria, a main entry criterion needs to be present: patients must have at least one episode of peripheral joint or bursal swelling, pain, or tenderness. If the entry criterion is present, a second criterion, the so called sufficient criterion is evaluated: MSU crystals in synovial fluid extracted from a symptomatic joint or bursa or in a tophus, assessed by a competent examiner. If the sufficient criterion is present, the case is classified as gout, without the need for any further assessment or criteria. No exclusion criteria have been set, as gout can coexist with many other conditions.

However, if the sufficient criterion is absent, a number of other criteria have to be evaluated in order to classify a case as gout; these criteria are distributed in three domains: clinical, laboratory and imaging [46].

**The clinical domain** includes four aspects:

- **The pattern of joint or bursal involvement**: symptomatic episode at 1st MTF joint, ankle, midfoot (in a mono-, oligo- or polyarticular presentation)

  Characteristics of symptomatic episodes:
  - erythema overlying affected joint,
  - can’t bear touch or pressure to affected joint,
  - great difficulty to walk or use joint.

- **Time-course of episodes**: to be judged typical, an episode should present at least two of the following:
  - time to maximal level of pain < 24 hours,
  - resolution of symptoms ≤ 14 days and
  - complete resolution between symptomatic episodes.

- **Clinical evidence of tophus**: draining or chalk-like subcutaneous nodule under transparent skin with overlying vascularity.
The laboratory domain relates to serum urate measurement, recommended by the uricase method, ideally during the intercritical period (at least 4 weeks after an acute episode), using for scoring the highest value recorded.

The imaging domain is fulfilled when able to demonstrate presence in a joint or bursa, ever symptomatic, of MSU crystal deposition, by:

- **US evidence of DCS**
- **DECT evidence of urate deposition**
- **and/or imaging evidence of gout-related joint damage**: at least one erosion (defined as a cortical break with sclerotic margin and overhanging edge, excluding osteoarthritis related findings as gull wing appearance or involvement of DIP joints) on conventional radiography of hands and feet.

The threshold value for classifying a case as gout is 8 points. **The performance of the criteria is significantly improved by imaging**, with a sensitivity of 0.92 and a specificity of 0.89, compared with 0.85 and 0.78 respectively when applying criteria without imaging.

For the first time the new set of proposed criteria for gout includes advanced imaging techniques for the detection of gout [37]. Recent reports suggest that US [47] and DECT [48] may allow accurate identification of some patients with gout. Due to low costs and affordability, as well as easy access to US machines, inclusion of US in these criteria will be an essential step forward in advancing the research agenda for gout management [49-51].

### Conclusions

Without any doubt, the recent inclusion of US in the classification criteria of various rheumatic diseases, such as PMR and gout, implies that this imaging technique is not only useful as a valued diagnostic tool for individual cases, but also on a larger scale it is improving doctors' ability to classify diseases. Its use is thus changing our understanding of rheumatic diseases allowing further advances in research and clinical practice [52-54].

Is this the end of the story? Most probably not: shortly after the publication of the ACR Classification Criteria for Sjogren’s Syndrome [55], a number of recent studies [56-59] confirmed that US examination is able to significantly improve the diagnostic performance of the ACR classification criteria for SS and recommended its inclusion in the criteria.

### Conflict of interest

None.

### References


