Arthritis Care & Research
Vol. 67, No. 7, July 2015, pp 891–897
DOI 10.1002/acr.22583

REVIEW ARTICLE

Distinctions Between Diagnostic and Classification Criteria?

ROHIT AGGARWAL,[1] SARAH RINGOLD,[2] DINESH KHANNA,[3] TUHINA NEOGI,[4] SINDHU R. JOHNSON,[5] AMY MILLER,[6] HERMINE I. BRUNNER,[7] RIKKE OGAWA,[8] DAVID FELSON,[4] ALEXIS OGDIE,[9] DANIEL ALETAHA,[10] AND BRIAN M. FELDMAN[11]

Introduction

Rheumatologists face unique challenges in discriminating between rheumatologic and nonrheumatologic disorders with similar manifestations and in discriminating among rheumatologic disorders with shared features. The majority of rheumatic diseases are multisystem disorders with poorly understood etiology; they tend to be heterogeneous in their presentation, course, and outcome and do not have a single clinical, laboratory, pathologic, or radiologic feature that could serve as a "gold standard" in support of diagnosis and/or classification. Thus, the development of criteria for use in routine clinical care and in clinical research has been an important focus in rheumatology. An improved understanding of disease pathogenesis and the availability of new diagnostic tools have led to a reexamination of the existing classification and diagnostic criteria; updated classification criteria for some diseases have been endorsed recently (1,2).

The American College of Rheumatology (ACR) Subcommittee on Classification and Response Criteria is responsible for guiding the development and validation of new classification and response criteria that are eventually considered for endorsement by the ACR. This responsibility includes reviewing proposals for the development of new criteria sets and providing the ACR leadership with recommendations for the development and approval of new classification and response criteria sets (1,3–5). Members of the Subcommittee previously offered recommendations for updating the standards for considering classification and response criteria (6). That prior work provided details about the rationale for the position of the ACR regarding classification criteria, but clarification regarding the issue of diagnostic criteria was lacking. Indeed, in 2010 the ACR endorsed preliminary diagnostic criteria for fibromyalgia (7), which prompted discussions about whether the Subcommittee should also support the development and ACR endorsement of diagnostic criteria, in addition to that of classification and response criteria.

The primary objectives of this current article, by former and current members of the Subcommittee on Classification and Response Criteria, are to compare diagnostic and classification criteria, using specific examples from the published literature, and to clarify the ACR's position on both types of criteria (Table 1).

Diagnostic criteria

Diagnosis may be defined as determination of the cause or nature of an illness by evaluating the signs, symptoms, and results of supportive tests in an individual patient. Diagnostic criteria are a set of signs, symptoms, and tests for use in routine clinical care to guide the care of individual patients.

Supported by the American College of Rheumatology Classification and Response Criteria Subcommittee of the Committee on Quality of Care.

[1]Rohit Aggarwal, MD, MS: University of Pittsburgh, Pittsburgh, Pennsylvania; [2]Sarah Ringold, MD, MS: Seattle Children's Hospital, Seattle, Washington; [3]Dinesh Khanna, MD, MS: University of Michigan, Ann Arbor; [4]Tuhina Neogi, MD, PhD, FRCPC, David Felson, MD: Boston University School of Medicine, Boston, Massachusetts; [5]Sindhu R. Johnson, MD, PhD: Toronto Western Hospital, Mount Sinai Hospital, and University of Toronto, Toronto, Ontario, Canada; [6]Amy Miller: American College of Rheumatology, Atlanta, Georgia; [7]Hermine I. Brunner, MD, MSc: Cincinnati Children's Hospital and Medical Center, Cincinnati, Ohio; [8]Rikke Ogawa, MLIS: University of California, Los Angeles; [9]Alexis Ogdie, MD, MSCE: University of Pennsylvania, Philadelphia; [10]Daniel Aletaha, MD, MSc: Medical University of Vienna, Vienna, Austria; [11]Brian M. Feldman, MD, MSc, FRCPC: Toronto Western and Mount Sinai Hospitals, The Hospital for Sick Children, and University of Toronto, Toronto, Ontario, Canada.

Dr. Khanna has received consultant fees, speaking fees, and/or honoraria (less than $10,000 each) from Actelion, Bayer, Biogene, Idec, BMS, Cigna, Genentech/Roche, Intermune, Merck, Sanofi-Aventis, and Genzyme. Dr. Felson has received consultant fees, speaking fees, and/or honoraria (less than $10,000) from Zimmer Knee Creations.

Address correspondence to Rohit Aggarwal, MD, MS, Assistant Professor of Medicine, University of Pittsburgh, 3601 5th Avenue, Suite 2B, Pittsburgh, PA 15217. E-mail: aggarwalr@upmc.edu.

Submitted for publication December 30, 2013; accepted in revised form March 10, 2015.

Table 1. Summary of the key points*

Background question	Should ACR support the development and endorsement of diagnostic criteria?	
Objective	1. To compare diagnostic and classification criteria 2. To clarify the ACR's position on ACR endorsement of diagnostic criteria	
Methods	Systematic review of literature, recommendation of ACR Classification and Response Criteria Subcommittee, and subsequent ACR Quality of Care Committee and ACR Board of Directors approval	
Results	**Classification Criteria**	**Diagnostic Criteria**
	Classification criteria are standardized definitions that are primarily intended to enable clinical studies to have uniform cohorts for research	Diagnostic criteria are a set of signs, symptoms, and tests developed for use in routine clinical care to guide the care of individual patients
	Need to define (relatively) homogenous group that can be compared across studies and geographic regions	Need to be broad and must reflect the all possible different features and severity of a disease (heterogeneity)
	Very high specificity is required, even if some loss in sensitivity	Both specificity and sensitivity need to be very high, approaching 100%, which is difficult to achieve
	Single universal classification criteria can be applied to different geographical regions, race and ethnicities	Single universal diagnostic criteria cannot be used for making diagnosis due to different disease prevalence in different geographic areas, race and ethnicities
	Classification criteria are possible for disease with and without true "gold standards" (e.g., MSU crystals in gout)	Diagnostic criteria are possible for disease with a true "gold standard" like MSU crystals in gout. For such diseases, the classification and diagnostic criteria can be very similar. For diseases without gold standard, development of diagnostic criteria poses significant challenges
	Differences in resources and feasibility has limited effect on classification criteria	Differences in resources and feasibility significantly effect development of diagnostic criteria
	Classification criteria are for research, and therefore, should have no or little impact on billing and reimbursement	Diagnostic criteria are for diagnosis, and therefore, have implications for billing and reimbursement
	Classification criteria have no treatment implications for patients	Diagnostic criteria have treatment implications for patients
	Health care priorities of different geographical areas do not influence classification criteria	Health care priorities of different geographic areas may influence diagnostic criteria
	Classification of a disease can be accomplished by a set of criteria with reasonable sensitivity and specificity	Diagnosis is a complex multi-step process by a physician, which is difficult to accomplish with a single set of criteria
ACR recommendation	Given the difficulty in establishing a uniform diagnostic criteria as noted above, the ACR will only provide approval for classification criteria and will no longer consider funding or endorsement of diagnostic criteria	

* ACR = American College of Rheumatology; MSU = monosodium urate.

Diagnostic criteria are generally broad and must reflect the different features of a disease (heterogeneity), with a goal of accurately identifying as many individuals with the condition as possible. Given this complexity, the development and validation of diagnostic criteria can be quite challenging. The Diagnostic and Statistical Manual of Mental Disorders (DSM) is likely the best-known example of diagnostic criteria. Its initial development was prompted by the observation of extremely poor agreement among providers regarding psychiatric diagnoses. There are only a few validated diagnostic criteria in rheumatology, and clinicians commonly establish a diagnosis based on a subjective combination of clinical signs/symptoms, results of available clinical tests, and knowledge about epidemiology in their geographic area.

Classification criteria

Classification criteria are standardized definitions that are primarily intended to create well-defined, relatively homogeneous cohorts of patients for clinical research; they are not intended to capture the entire universe of possible patients but rather to capture the majority of patients who share key features of the condition. Hence, the goal of classification differs from the intent of diagnostic criteria. Validated classification criteria are considered critical to the interpretation of study findings and comparisons of results between studies. Despite facilitating the comparison of study results, classification criteria have the potential to restrict the external validity of studies, because interventions may perform differently in study participants who fulfill classification criteria for a disease than in the broader group of patients in whom the same disease was diagnosed (i.e., those who share only some of the disease manifestations considered in the classification criteria).

Although classification criteria may provide some framework to aid in diagnosis and are frequently used this way in teaching, these criteria traditionally have high specificity (defined as the proportion of patients who are known not to have the disease who will test negative for it), which generally comes at the expense of somewhat lower sensitivity (defined as the proportion of patients who are known to have the disease and who will test positive for it). Consequently, few individuals are incorrectly labeled as having a disease (false positives), but a proportion of those with the disease diagnosis may be "missed," i.e., labeled as not having the disease based on the classification criteria (false negatives). This may make classification criteria inappropriate for use in routine clinical care (8).

Continuum of diagnosis and classification

Although *diagnostic criteria* may be different from *classification criteria*, at least with regard to their intended purpose, in reality they represent 2 ends of a continuum (9). The "distance" between diagnostic and classification criteria on this continuum depends on various factors, including disease prevalence, geographic area, and prevalence of "mimickers," among others. When the etiology of a disease is well defined (as in gout and Lyme disease), diagnostic and classification criteria may be very similar and can be used interchangeably. If sufficient internal and external validity for diagnosis is demonstrated in a given population, classification criteria can be diagnostic. In theory, a diagnosis applies classification criteria to an individual patient (9). Therefore, when classification criteria have perfect (100%) sensitivity and specificity, classification and diagnostic criteria are synonymous and would correctly identify every single individual case (10). However, because disease features typically are not identical among patients with a given disease, classification criteria are not 100% accurate, thus leaving a certain proportion of patients misclassified. Because of this possibility of misclassification, meeting classification criteria is not equivalent to carrying a given diagnosis. Only physicians considering features of an individual patient, beyond those represented in the classification criteria, in addition to extraneous factors (such as the local prevalence of conditions that are included in the differential diagnosis) can establish a diagnosis for an individual patient.

Due to the lack of gold standards in the field of rheumatology, any criteria (classification or diagnostic) are difficult to establish. As detailed below, compared with classification criteria, an array of factors pose greater challenges and clinical implications for the development and validation of diagnostic criteria. Even in those situations in which diagnostic criteria can be established, the question remains as to whether the ACR and/or other international organizations should endorse a single set of diagnostic criteria.

Literature review

A systematic search of articles addressing classification and/or diagnostic criteria for the rheumatic diseases was performed by an experienced librarian (RO), considering the PubMed database (1940–2011) and the Cochrane Central Register of Controlled Trials database (from 1996 to 2011). In combining the search terms for "Diagnosis AND Classification AND Rheumatic Disease AND Methodology," 3,825 citations from PubMed and 88 from the Cochrane database were identified. Two reviewers (RA and SR) independently screened the titles and abstracts of articles for relevance to classification and diagnostic criteria in rheumatic diseases. Abstracts were screened to identify articles that defined, updated, addressed, reviewed, or commented on methodologic aspects of classification or diagnostic criteria for the rheumatic diseases. Screening of titles and abstracts excluded 3,681 articles, leaving 232 articles for detailed review. This led to the identification of 97 articles that were deemed relevant for the evaluation of classification and diagnostic criteria that are considered in this review (Figure 1). Relevant articles were defined as those that either illustrate differences between classification criteria and diagnostic criteria, identify key advantages and disadvantages of classification or diagnostic criteria, or evaluate performance characteristics of classification or diagnostic criteria. Eighteen additional relevant articles were identified through handsearching the bibliography of the initially identified 97 articles (Figure 1). Articles fit into one of the following 6 categories: 1) study of or commentary on differences and/or similarities between classification criteria and physician assessment, 2) description of the performance of classification or diagnostic criteria in various populations, geographic regions, or different practice settings, 3) proposal of original or revised classification or diagnostic criteria, 4) comparison of the performance of established classification criteria, 5) description of various cohorts using established classification criteria, and 6) discussion of either classification or diagnostic criteria not otherwise related to one of the above categories.

Summary of the published literature

1. Examples of the differential performance of classification criteria in relation to physician assessment. No examples of direct comparison between diagnostic criteria and classification criteria were identified in the literature review. The performance of classification criteria as diagnostic tools has been assessed in a handful of studies. Because of the lack of gold standards for diagnosis and

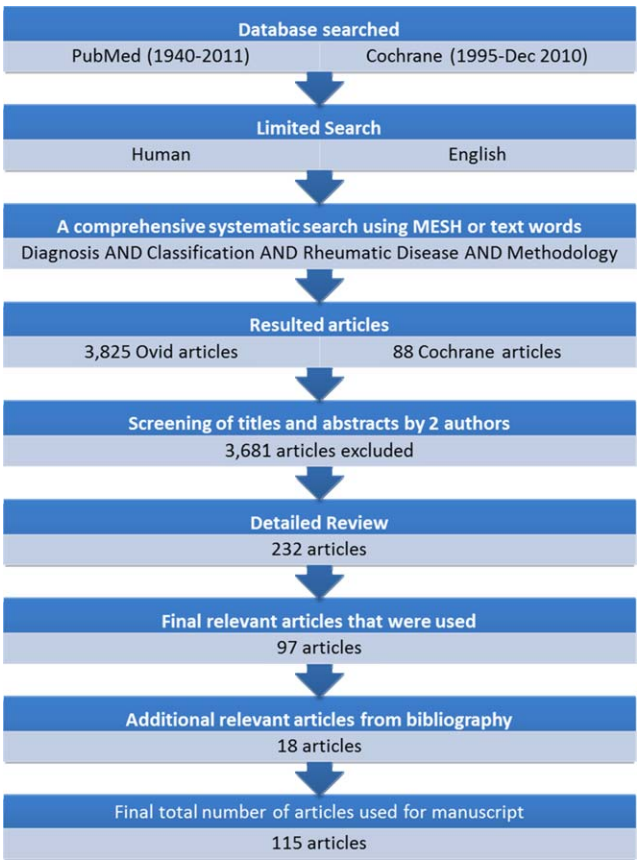

Figure 1. Methods used for comprehensive literature search for articles relevant to classification and/or diagnostic criteria.

classification, the performance of criteria sets was compared with physician judgment in these studies. For example, Rao et al assessed the measurement properties of the 1990 ACR classification criteria for vasculitis when used as diagnostic criteria, relying on the treating rheumatologist's final diagnosis as the gold standard (11). In that study, only 38 (75%) of 51 patients with vasculitis fulfilled the ACR classification criteria for one or more types of vasculitis, and 31 (21%) of 147 patients without vasculitis also fulfilled these criteria (79% specificity). This illustrates that compared with a physician's diagnosis, the 1990 ACR classification criteria had relatively low sensitivity and specificity for predicting the presence of a specific type of vasculitis in an individual patient seen in routine clinical practice. A separate study demonstrated that the Chapel Hill Consensus Conference (CHCC) classification criteria for vasculitis correctly identified only 30% of patients with the disease, when compared with a physician's assessment (gold standard) (12). Likewise, Patarroyo et al reported that 65.8% of patients with histopathologically proven vasculitis from a single center could not be classified as having a discrete type of vasculitis as defined by the CHCC criteria (13).

Clinicians have expressed concern about the high number of patients in whom systemic lupus erythematosus (SLE) is diagnosed in clinical practice who fail to meet the SLE classification criteria (14,15). Similarly, the diagnosis of knee osteoarthritis (OA) made by community physicians is in only fair agreement with the ACR criteria for the classification of knee OA ($\kappa = 0.28$) (16). Other examples of diseases in which the classification criteria do not perform well when compared with clinical diagnoses by treating rheumatologists or experts include juvenile idiopathic arthritis and systemic sclerosis (17–19). These examples may reflect not only differences between physician decision-making versus classification criteria but also that older classification criteria may require revision; several criteria sets have now been updated or are in the process of being updated (e.g., for rheumatoid arthritis [RA], systemic sclerosis [1,2]) or are being developed (e.g., for vasculitis [20] and gout [21]).

Nonetheless, these examples suggest that classification differs from diagnosis, and that use of classification criteria may result in underreporting the presence of a disease. This is because classification criteria capture a narrower range of disease severity than that treated in routine clinical practice, because classification criteria tend to identify a uniform population for participation in clinical trials at the expense of excluding some patients with less common phenotypes, as suggested by the above-described examples (22). The rationale for perhaps favoring specificity over sensitivity for classification criteria in the setting of clinical trials is to avoid exposing patients who may not have the disease to the possible risks associated with experimental interventions.

2. Need for revision of classification criteria. Newer revised classification criteria may perform better than some older classification criteria in terms of sensitivity and/or specificity, with use of better data sets and methodology (1,5,23). The more favorable risk–benefit profiles of therapeutics and recognition that early therapy may affect long-term prognosis have prompted trials in patients with rheumatic diseases who do not yet meet the thresholds for traditional classification. Consequently, some recent classification criteria have focused more on improving the sensitivity of criteria (1,23). For example, the notion that the ACR 1987 revised criteria for the classification of RA (24) missed early disease (i.e., lacked sensitivity) led to development of the ACR/European League Against Rheumatism 2010 criteria for RA (5) (better sensitivity) so that patients with early disease could be identified for intervention studies. It should be considered, however, that this approach may increase the chance of false-positive results in the absence of gold standard tests and likely has implications for prevalence estimates of these diseases as well as for clinical practice. On the other hand, a lack of sufficient specificity of criteria (i.e., false positives) also has bearing on the enrollment of patients into trials of agents with unclear safety/efficacy profiles; as a result, some recent classification criteria have aimed to improve specificity (21).

3. Effects of geographic area, practice setting, and race on criteria performance. The performance of any criteria (classification or diagnostic) is dependent on the prevalence of the disease in a given geographic area or clinical setting (e.g., community clinic versus tertiary care facility). While sensitivity and specificity are functions of the screening test or criteria set and are not influenced by disease incidence or prevalence, the predictive validity

changes with the prevalence of the disease. The performance of criteria depends on both the pretest probability of the disease, (which reflects the prevalence of the disease as well as potential "mimickers") and the sensitivity and specificity of the criteria themselves. Given the low prevalence of certain rheumatologic diseases, the positive predictive value (defined as the proportion of positive test results that are true positives) of any criteria set will generally be low. This is probably the reason why there are few diagnostic criteria sets for rheumatology. For example, in areas where Behçet's disease is endemic (high pretest probability), patients with recurrent oral ulcers may be accurately diagnosed and treated based on physician's judgment in the absence of supporting criteria, whereas in the US, where the disease prevalence is low, any set of diagnostic criteria will have a low positive predictive value. Similarly, the performance of the European Spondylarthropathy Study Group (ESSG) preliminary criteria for the classification of spondylarthropathy varies across patients seen in private practices, those seen at outpatient departments, and those admitted to the hospital, depending on the prevalence of spondyloarthritis (SpA) in different regions (25). For example, in Spanish rheumatology services where the prevalence of SpA is 49%, the likelihood of a patient younger than age 35 years fulfilling the ESSG criteria for SpA is 87%. Conversely, in clinics in France where the age-adjusted prevalence of SpA is only 22.5%, the likelihood of fulfilling the ESSG criteria is estimated to be 70% (25). While classification and diagnostic criteria perform differently in different clinical and geographic settings, this difference is less pronounced for classification criteria, because their intended purpose is to identify patients with similar disease features for studies within different populations rather than to guide clinicians in establishing diagnoses and making treatment decisions.

The same concept applies to patients of different races or ethnicities within a geographic area. Indeed, the measurement properties of classification criteria can differ markedly when used in populations other than those used for development of the classification criteria. As one example, a study conducted in Asia showed that only 12 (17%) of 71 patients with IgA vasculitis (Henoch-Schönlein) fulfilled both the ACR vasculitis classification criteria and those of the CHCC (26). The variable performance of both classification and diagnostic criteria in different settings highlights the differences between these types of criteria and further illustrates the difficulty of developing diagnostic criteria for which performance is consistent across populations. Diagnostic criteria will typically need to be based on the local prevalence of the disease of interest as well as that of other diseases included in the differential diagnosis, which is not practical given the vast differences in epidemiology of most rheumatic diseases in different clinical settings and geographic areas. The performance of classification criteria is also affected by their application to patients other than the intended target population (e.g., if the 2010 criteria for the classification of RA was applied to patients with "burned out" deforming nodular RA, when it was intended for use in patients with early active RA).

4. Well-defined disease phenotypes. One of the main differences between classification and diagnostic criteria is that classification criteria are aimed at assembling a study sample that is well defined and representative of the vast majority of patients with the disease. In contrast, diagnostic criteria aim to identify all patients with the disease, including those with unusual features or presentations. Achieving a relatively homogeneous disease population is important for any classification criteria so that multiple studies and populations can be compared or combined. On the other hand, to be highly sensitive while preserving acceptable specificity, diagnostic criteria have to allow for all of the heterogeneous manifestations of the disease (which may be difficult to achieve in rheumatic diseases). SLE is a prototypical example of a disease with heterogeneous presentations. Although the SLE classification criteria can support a diagnosis of SLE, clinicians still must diagnose SLE based upon the totality of a patient's disease manifestations (27,28). Classification criteria for SLE perform reasonably well for making a diagnosis in academic medical centers that attract patients with more severe or advanced disease, who typically have a higher pretest probability of having the diagnosis (27). However, the SLE classification criteria may fail to recognize patients with milder phenotypes or uncommon presentations of the disease. Classification criteria tend to include phenotypic features that have sufficiently high prevalence, whereas low-prevalence features that may be very specific and helpful in diagnosis are typically not included in criteria sets due to the expected low yield from including such a feature.

5. Rheumatic diseases where both diagnostic and classification criteria are feasible. Single sets of criteria that serve for both classification and diagnosis appear feasible for diseases for which there is a diagnostic "gold standard." The presence of monosodium urate crystals (MSU) in synovial fluid during an episode of acute arthritis is widely considered diagnostic for gout. In fact, compared with the presence of MSU crystals, the 1977 ACR preliminary criteria for the classification of acute gout (29) have shown limited diagnostic accuracy (30–32). In one study, diagnosis by primary care physicians correctly identified 93.5% of patients with MSU crystal positivity, indicating that a clinician's diagnosis of gout can be at least as sensitive as the classification criteria for acute gout (30). Similar support for the use of diagnostic criteria can be applied to any rheumatic disease in which the pathology is well understood and/or the etiology is well defined. Infectious arthritides, such as septic arthritis, can be diagnosed based on gold standard tests, and a diagnostic criteria set can be devised. For such diseases, diagnostic criteria are also suited to guiding subject identification for research studies, because diagnostic criteria perform as well as classification criteria in terms of sensitivity and specificity.

6. Resources and feasibility. Feasibility, acceptability, and availability of resources are other potentially limiting factors in establishing universally accepted diagnostic criteria. Clinicians may be faced with limited access to or lack of affordability of testing in certain geographic regions, patients' own financial and/or insurance limitations, patient preferences, and overall health condition, among others, when deciding on strategies to establish a diagnosis. This could necessitate making a diagnosis and subsequently initiating treatment based solely on a clinical basis. More stringent

diagnostic criteria that require a particular laboratory or imaging test or surgical procedure could constitute a hurdle for patients and clinicians, and use of such criteria has the potential to postpone the initiation of effective therapy.

7. Health priorities of a country or geographical area. In a malaria-endemic area, a physician can empirically diagnose malaria in a patient with high-grade fever and chills and start empiric treatment (9). In most cases, the initiation of malaria treatment, without waiting for the results of a confirmatory test, prevents serious complications and even death. The above clinical approach in these malaria-endemic regions outweighs the harm of over-diagnosis and overtreatment with a relatively nontoxic medication in patients without malaria. Conversely, use of a similar approach in a Nordic country would likely be unacceptable and irrational. Therefore, the health priorities and conditions in different countries/geographic areas often dictate the diagnostic approaches to be used, which suggests that a single universal diagnostic criteria set cannot always be applied equally in different regions of the world. However, classification criteria for the purpose of enrollment into clinical trials and epidemiologic studies may be used across the globe, with high specificity even if a few cases are missed, without affecting the internal validity of the study.

8. Legal, financial, and treatment implications. Unlike classification of a disease for research purposes, the accurate diagnosis of a particular disease has important implications for a patient's treatment as well as for billing and reimbursement. Highly specific diagnostic criteria will leave some patients undiagnosed (because no criteria will ever be 100% sensitive). This means that such patients may be denied treatment coverage if insurance companies and government agencies use the diagnostic criteria as a standard for reimbursement. Similarly, patients in whom an illness was incorrectly diagnosed (because 100% specificity is difficult) can encounter difficulty in obtaining health insurance or life insurance and may be exposed unnecessarily to incorrect, potentially harmful, therapies.

9. Undifferentiated rheumatic diseases. Many patients with rheumatic disease present to their physicians when their disease is at an undifferentiated stage, which may later evolve into more established disease. Although classification criteria are typically applied at a given time point, they can be reevaluated, because individuals may fulfill the criteria as the disease manifestations change over time. Unlike acute infection, many rheumatic diseases evolve over time, and cross-sectional application of any criteria as either "disease present or absent" is too simplistic. Moreover, some cases may never evolve into well-established disease, and others may transform from one presumed condition to another. Strict universal diagnostic criteria may limit the ability to make a clinical diagnosis and treat undifferentiated diseases based on symptoms. This notion was emphasized in recent studies on the outcome of early arthritis, in which 32–53% of patients remained unclassified after 1 year of observation (33,34). Similarly, in a 3-year followup study of 270 patients with early arthritis, the diagnosis remained unclear in 61 (23%) of the patients and changed between the first and last examination in 96 (46%) of the other 209 cases (35).

10. Complex decision making for diagnosis. Finally, clinicians perform a complex multistep process in order to make a diagnosis of rheumatic disease. This process includes balancing the post-test probability of the disease with thresholds for further action based on factors such as disease severity, risks of further testing, side effects of treatment, and ruling out other conditions in the differential diagnosis (e.g., infection or malignancy). It is difficult to establish diagnostic criteria that may satisfactorily perform this complex multistep process.

Role of the American College of Rheumatology

Classification criteria have demonstrated utility for identifying well-defined, relatively homogeneous groups of patients for clinical research purposes across different regions and have some utility as teaching tools in the clinical setting; however, they may not capture all physician-assigned diagnoses. Conversely, compared with classification criteria, diagnostic criteria appear to be more impacted by practice setting, and the performance characteristics of diagnostic criteria may vary significantly due to differences in disease prevalence and the severity and manifestations of disease in different settings. Given these differences, concerns regarding the challenges in generating diagnostic criteria with consistent performance properties, and the legal and financial implications associated with diagnostic criteria, the ACR will provide approval only for classification criteria and will no longer consider funding or endorsing diagnostic criteria. However, the ACR recognizes the importance of diagnostic tools to aid rheumatologists in their clinical practice and encourages their development. The ACR anticipates that both types of criteria will continue to evolve as the pathogenesis of rheumatic diseases becomes better understood and as the emphasis on studies of comparative effectiveness increases.

In conclusion, diagnostic and classification criteria play central roles in clinical rheumatology practice. Unfortunately, the existing criteria for rheumatic diseases are not always properly applied, most often due to confusion regarding the differences between the 2 types of criteria. Classification criteria are used as a standardized means of including a well-defined set of patients in research studies to ensure comparability across studies. Given the heterogeneous nature of rheumatic diseases, it is difficult to capture the full range of disease presentations using any single set of criteria. Therefore, any criteria would be expected to fail to identify some cases of a disease due to the criteria capturing a more homogeneous population and a narrower range of disease severity than that seen in routine clinical practice. Nonetheless, classification criteria are critically important for advancing research in the field of rheumatology, enabling the conduct of clinical trials and epidemiologic studies with well-defined patient populations. The process of diagnosis, particularly for complicated multisystem involvement typical of rheumatic diseases, is a highly complex cognitive process that requires synthesis of many data points typically beyond a simple algorithm-based set of criteria.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Aggarwal had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Aggarwal, Ringold, Khanna, Brunner, Felson, Johnson, Miller, Neogi, Feldman.

Acquisition of data. Aggarwal, Ringold, Khanna, Ogawa, Neogi.

Analysis and interpretation of data. Aggarwal, Ringold, Khanna, Brunner, Ogdie, Aletaha, Neogi, Feldman.

REFERENCES

1. Neogi T, Aletaha D, Silman AJ, Naden RL, Felson DT, Aggarwal R, et al. The 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis: phase 2 methodological report. Arthritis Rheum 2010;62:2582–91.
2. Van den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2013;65:2737–47.
3. Dasgupta B, Cimmino MA, Kremers HM, Schmidt WA, Schirmer M, Salvarani C, et al. 2012 provisional classification criteria for polymyalgia rheumatica: a European League Against Rheumatism/American College of Rheumatology collaborative initiative. Arthritis Rheum 2012;64:943–54.
4. Johnson SR, Goek ON, Singh-Grewal D, Vlad SC, Feldman BM, Felson DT, et al. Classification criteria in rheumatic diseases: a review of methodologic properties. Arthritis Rheum 2007;57:1119–33.
5. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO III, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2010;62:2569–81.
6. Classification and Response Criteria Subcommittee of the American College of Rheumatology Committee on Quality Measures. Development of classification and response criteria for rheumatic diseases. Arthritis Rheum 2006;55:348–52.
7. Wolfe F, Clauw DJ, Fitzcharles M, Goldenberg DL, Katz RS, Mease P, et al. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res (Hoboken) 2010;62:600–10.
8. Dougados M, van der Linden S, Juhlin R, Huitfeldt B, Amor B, Calin A, et al. The European Spondylarthropathy Study Group preliminary criteria for the classification of spondylarthropathy. Arthritis Rheum 1991;34:1218–27.
9. Yazici H. Diagnostic versus classification criteria: a continuum. Bull NYU Hosp Jt Dis 2009;67:206–8.
10. Fries JF, Hochberg MC, Medsger TA Jr, Hunder GG, Bombardier C, and the American College of Rheumatology Diagnostic and Therapeutic Criteria Committee. Criteria for rheumatic disease: different types and different functions. Arthritis Rheum 1994;37:454–62.
11. Rao JK, Allen NB, Pincus T. Limitations of the 1990. American College of Rheumatology classification criteria in the diagnosis of vasculitis. Ann Intern Med 1998;129:345–52.
12. Sorensen SF, Slot O, Tvede N, Petersen J. A prospective study of vasculitis patients collected in a five year period: evaluation of the Chapel Hill nomenclature. Ann Rheum Dis 2000;59:478–82.
13. Patarroyo PA, Restrepo JF, Rojas SA, Rondon F, Matteson EL, Iglesias-Gamarra A. Are classification criteria for vasculitis useful in clinical practice? Observations and lessons from Colombia. J Autoimmune Dis 2009;6:1.
14. Fries JF, Siegel RC. Testing the "preliminary criteria for classification of SLE." Ann Rheum Dis 1973;32:171–7.
15. Michet CJ Jr, McKenna CH, Elveback LR, Kaslow RA, Kurland LT. Epidemiology of systemic lupus erythematosus and other connective tissue diseases in Rochester, Minnesota, 1950 through 1979. Mayo Clin Proc 1985;60:105–13.
16. Peat G, Greig J, Wood L, Wilkie R, Thomas E, Croft P. Diagnostic discordance: we cannot agree when to call knee pain "osteoarthritis." Fam Pract 2005;22:96–102.
17. Ziswiler HR, Urech R, Balmer J, Ostensen M, Mierau R, Villiger PM. Clinical diagnosis compared to classification criteria in a cohort of 54 patients with systemic sclerosis and associated disorders. Swiss Med Wkly 2007;137:586–90.
18. Maricq HR, LeRoy EC, D'Angelo WA, Medsger TA Jr, Rodnan GP, Sharp GC, et al. Diagnostic potential of in vivo capillary microscopy in scleroderma and related disorders. Arthritis Rheum 1980;23:183–9.
19. Behrens EM, Beukelman T, Gallo L, Spangler J, Rosenkranz M, Arkachaisri T, et al. Evaluation of the presentation of systemic onset juvenile rheumatoid arthritis: data from the Pennsylvania Systemic Onset Juvenile Arthritis Registry (PASOJAR). J Rheumatol 2008;35:343–8.
20. Craven A, Robson J, Ponte C, Grayson PC, Suppiah R, Judge A, et al. ACR/EULAR-endorsed study to develop Diagnostic and Classification Criteria for Vasculitis (DCVAS). Clin Exp Nephrol 2013;17:619–21.
21. Dalbeth N, Fransen J, Jansen TL, Neogi T, Schumacher HR, Taylor WJ. New classification criteria for gout: a framework for progress. Rheumatology (Oxford) 2013;52:1748–53.
22. Hunder GG. The use and misuse of classification and diagnostic criteria for complex diseases [editorial]. Ann Intern Med 1998;129:417–8.
23. Johnson SR, Fransen J, Khanna D, Baron M, van den Hoogen F, Medsger TA Jr, et al. Validation of potential classification criteria for systemic sclerosis. Arthritis Care Res (Hoboken) 2012;64:358–67.
24. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.
25. Gomariz EM, del M, Guijo VP, Contreras AE, Villanueva M, Estevez EC. The potential of ESSG spondyloarthropathy classification criteria as a diagnostic aid in rheumatological practice. J Rheumatol 2002;29:326–30.
26. Murali NS, George R, John GT, Chandi SM, Jacob M, Jeyaseelan L, et al. Problems of classification of Henoch Schonlein purpura: an Indian perspective. Clin Exp Dermatol 2002;27:260–3.
27. Perez-Gutthann S, Petri M, Hochberg MC. Comparison of different methods of classifying patients with systemic lupus erythematosus. J Rheumatol 1991;18:1176–9.
28. Panush RS, Schur PH. It is lupus? Bull Rheum Dis 1997;46:3–8.
29. Wallace ST, Robinson H, Masi AT, Decker JL, McCarty DJ, Yu TF. Preliminary criteria for the classification of the acute arthritis of primary gout. Arthritis Rheum 1977;20:895–900.
30. Janssens HJ, 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. Ann Rheum Dis 2010;69:1255–6.
31. Malik A, Schumacher HR, Dinnella JE, Clayburne GM. Clinical diagnostic criteria for gout: comparison with the gold standard of synovial fluid crystal analysis. J Clin Rheumatol 2009;15:22–4.
32. 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. In press.
33. Van der Horst-Bruinsma IE, Speyer I, Visser H, Breedveld FC, Hazes JM. Diagnosis and course of early-onset arthritis: results of a special early arthritis clinic compared to routine patient care. Br J Rheumatol 1998;37:1084–8.
34. Harrison BJ, Silman AJ, Barrett EM, Scott DG, Symmons DP. Presence of psoriasis does not influence the presentation or short-term outcome of patients with early inflammatory polyarthritis. J Rheumatol 1997;24:1744–9.
35. Berthelot JM, Klarlund M, McGonagle D, Bernelot-Moens HJ, Calin A, Harrison B, et al. Lessons from an international survey of paper cases of 10 real patients from an early arthritis clinic. CRI (Club Rhumatismes et Inflammation) Group. J Rheumatol 2001;28:975–81.