Published in final edited form as:

Best Pract Res Clin Rheumatol. 2014 December ; 28(6): 921–934. doi:10.1016/j.berh.2015.04.004.

The Use and Abuse of Diagnostic/Classification Criteria

Rayford R. June¹ and Rohit Aggarwal²

¹Department of Medicine, Division of Rheumatology, Penn State College of Medicine

²Department of Medicine, Division of Rheumatology and Clinical Immunology, University of Pittsburgh

Abstract

In rheumatic diseases, classification criteria have been developed to identify well-defined homogenous cohorts for clinical research. Although, they are commonly used in clinical practice, their use may not be appropriate for routine diagnostic clinical care. Classification criteria are being revised with improved methodology and further understanding of disease pathophysiology, but still may not encompass all unique clinical situations to be applied for diagnosis of heterogeneous, rare, evolving rheumatic diseases. Diagnostic criteria development is challenging primarily due to difficulty for universal application given significant differences in prevalence of rheumatic diseases based on geographical area and clinic settings. Despite these shortcomings, the clinician can still use classification criteria for understanding the disease as well as a guide for diagnosis with a few caveats. We present the limits of current classification criteria, describe their use and abuse in clinical practice, and how they should be used with caution when applied in clinics.

Keywords

Diagnostic Criteria; Classification Criteria; Rheumatology

Introduction

Rheumatology is not a field of black and white, but a specialty full of gray. Multisystem clinical syndromes and diseases in Rheumatology attract clinicians and researchers whom seek to unify different shades of “gray” into a single diagnosis or classification criteria. While understanding of the pathophysiology in each disease has advanced, single laboratory

Corresponding Author: Rohit Aggarwal, M.D. MSc, Assistant Professor of Medicine, Division of Rheumatology, University of Pittsburgh Medical Center, Address: Biomedical Science Tower S 725, 3500 Terrace Street, Pittsburgh, PA 15217, aggarwalr@upmc.edu, Telephone: 412-383-8100, Fax: 412-383-8864.

Co-Author: Rayford R. June, M.D., Assistant Professor of Medicine, Division of Rheumatology, Penn State College of Medicine, Address: 500 University Drive, P.O. Box 850, Hershey, PA 17033, rjune@hmc.psu.edu, Telephone: 717-531-5384, Fax: 717-531-8274

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of Interest Statement

RRJ and RA report no relevant conflicts of interests related to this manuscript.

tests with high sensitivity and specificity sufficient to make a diagnosis, still do not exist for most of the rheumatic diseases. As opposed to a positive blood culture in infectious disease suggestive of bacteremia or a fasting blood glucose in endocrinology suggestive of diabetes mellitus, even the most common and well-studied clinical conditions in rheumatology such as rheumatoid arthritis can have significant diagnostic uncertainty of so-called “sero-negativity” up to 30% of the time (1). Despite making significant technological advances with diagnostic tests such as anti-cyclic citrullinated peptides (CCP), diagnosis is still imperfect given lack of 100% specificity for rheumatoid arthritis, and even worse sensitivity (2). This diagnostic uncertainty has led to the development of multiple sets of disease classification criteria for use in research on disease characterization, epidemiology, prognosis, and designing clinical trials for therapeutic investigation (3). While designed for clinical research, classification criteria are used and abused for clinical practice in patient care. This article will help define both classification and diagnostic criteria, describe limitations of current classification criteria and how their use in clinical practice, while not sufficient alone for diagnosis, can be an aid or aide-mémoire in making a diagnosis.

Statistical Principles

Prior to further discussion of classification and diagnostic criteria, review of certain statistical principles is necessary to clarify differences between classification and diagnostic criteria. *Sensitivity* is the percentage of true positives with the disease. A highly sensitive test is useful for ruling out a disease with a negative test but not necessarily ruling in the disease. Whereas *specificity* is the percentage of true negatives without disease, and is useful for ruling in a positive test (if high specificity) but not necessarily ruling out a disease. In the setting of a highly sensitive and specific test, while sensitivity is easily understood (if you do not have the test positive, then the disease is not present) specificity leads to confusion because, rather than being focused on having the disease, the focus is on not having the disease (4). Highly specific tests have low false positive rates and highly sensitive tests have low false negative rates. For instance, anti-cyclic citrullinated peptides (CCP) antibodies have been shown to have a high, greater than 90%, specificity for rheumatoid arthritis in established rheumatoid arthritis cohorts, where as it has moderate sensitivity of 66% (5). For knowing the true clinical applicability of sensitivity and specificity for a given test, the population in which it is studied or developed is important. For example, *CCP is useful in ruling in rheumatoid arthritis in subjects with polyarthritis secondary to its high specificity in this particular population* (6, 7). Without knowing the population in which CCP specificity was attributed to, the meaning of the specificity is lost. For example, CCP is positive in many non-inflammatory arthritis including infections (2). Therefore, *sensitivity* and *specificity* of any diagnostic or classification criteria are dependent on the reference gold standard used for its development as well as target population it is intended for. For example, the 2010 ACR/EULAR RA classification criteria were developed for use on early RA cohorts and therefore not intended to be used on burnt out deforming nodular RA.

Sensitivity and specificity are on a continuum with an inverse relationship where perfect sensitivity (close to 100%) will lead to loss in specificity and vice-versa. This is more evidenced in rheumatology where sensitivity and specificity of any criteria depend on multiple disease variables (4). When one gold standard test is used for diagnosis, as in gout

acute or septic arthritis (8), both sensitivity and specificity can remain high. However, as the number of variables needed for a disease classification increase, i.e. elevated c-reactive protein, number of swollen joints, seropositivity, the specificity in classification criteria increases, but sensitivity decreases, and vice-versa. The receiver operator curve (ROC) is the statistical and graphical description of this process showing the equilibrium between sensitivity and specificity (9). This same continuum is found when describing sensitivity and specificity of any classification and/or diagnostic criteria (4).

Furthermore, in addition to the number of variables and continuum of sensitivity and specificity involved in development of classification or diagnostic criteria, the ultimate performance of any classification or diagnostic criteria is highly dependent on the *prevalence* of the disease in the patient population being investigated (4). The principle of *positive predictive value (PPV)* illustrates this point. PPV is the proportion of true positives to the number of positive tests and is a measure of the accuracy or performance of a diagnostic test, or in the case of this discussion, diagnostic or classification criteria. Negative predictive value (NPV) is the opposite, a proportion of the number of true negatives to the number of negative tests. Both PPV and NPV are highly dependent on prevalence of disease. For instance, the prevalence of behcets disease in Turkey is almost 0.4% of the population, and in this population the international behcet's classification criteria have high sensitivity and specificity (4, 10, 11). In this population, the classification criteria can be used for diagnosis without significant numbers of false positive classifications. However, if the same criteria were used outside of Turkey where behcet's is rare, while the sensitivity would remain the same, the specificity would decrease with increase in false positivity: the positive predictive value of these classification criteria used for diagnosis would then plummet (12). With each change in individual population prevalence, the positive predictive value of the test is dependent on the frequency of disease in the population being studied. The concept of frequency of the disease in a population affecting utility of any criteria not only apply to geographical areas but also to clinical setting in which patients are being seen within the same geographical area. Whether it be a tertiary referral center for inflammatory myositis, a community rheumatology practice, or an acute care primary clinic, with each change in disease frequency, the usefulness of the test is determined by the positive predictive value in that clinical setting or population.

Classification Criteria

Classification criteria are defined as a set of disease characteristics used to group individuals into a well-defined relatively homogenous population with similar clinical disease features (3, 13). Classification criteria are essential for understanding disease pathogenesis and assessing treatment response. Classification criteria increase the specificity for underlying disease by creating a homogenous population while at times losing sensitivity on the ROC continuum.

Classification criteria are not designed to be used for clinical diagnosis or applied to individual patients but instead used to further research of the population. For instance, in 1990, the American College of Rheumatology produced many disease specific classification criteria aimed at furthering clinical research in disease specific states. The 1990 American

College of Rheumatology vasculitis classification criteria were shown by Rao et al to have particularly low positive predictive value for a *specific* vasculitis diagnosis, less than 30%, with only 38/51(75%) of patients with vasculitis fulfilling the 1990 classification criteria for *any* type of vasculitis, low sensitivity. Furthermore, the criteria also showed a low specificity with 31/147 (21%) of control patients without vasculitis actually full-filling these vasculitis classification criteria (14). Further study of the subsequent Chapel Hill Consensus Vasculitis criteria showed similar low sensitivity and specificity of classification criteria being applied as diagnostic criteria to individual patients in which the classification criteria were not designed to assess (15). Another study showed 65.8% of patients with histopathological proven vasculitis from a single center were classified according to the Chapel Hill Consensus Classification criteria (16), further showing the lack of clinical utility for the vasculitis classification criteria outside of clinical research studies. Similar to vasculitis, in knee osteoarthritis, Peat et al examined the relationships between patient diagnosis, clinician diagnosis, and classification of knee osteoarthritis by the American College of Rheumatology (ACR) criteria and found poor levels of agreement between patient or physician diagnosis with clinical classification of knee osteoarthritis (17). Classification criteria work best in the study of groups rather than the study, or care of, the individual patient (12).

In 2007, Johnson et al reviewed the methodological properties of various classification criteria for the rheumatic diseases (3) and emphasized that there are marked variation and significant deficiencies in methods used for criteria development, which affects the face validity and reliability of classification criteria. They noted the over 50% of the classification criteria used in rheumatic disease have not even been based off of patient data sets, but instead expert opinion alone (3). There were several other shortcomings noted in previous classification criteria including lack of adequate control with only 4 criteria using controls with non-rheumatic disease, small numbers of patients, poor description of content validity and description of construct of disease, inclusions of limited patient populations, including community rheumatology practices rather than only referral centers, and lack of independent validation (3). Prospective longitudinal cohorts are the optimal cohort design for criteria development, including adequate control groups, with each criterion in the classification criteria investigated for psychometric properties (9, 18, 19). The population in which classification criteria are developed needs to be considered as often classification criteria are specific to the patient population being studied (12, 14). Often times, classification criteria are developed on populations from academic medical centers without including community practices, which might limit use in community practices. Moreover, following development of classification criteria, their use is not confirmed until validated in independent cohorts. Overtime, with advancement in technology, methodological rigor and better understanding of the disease, older classification criteria need to be updated. The subcommittee of quality of care – classification and response criteria committee of the American College of Rheumatology (ACR) and the European League Against Rheumatism (EULAR) have been active in classification criteria development and further validation of use for in clinical trials (18).

With the efforts and support of ACR and EULAR in last 5 years, newer classification criteria are being developed which may perform better than some older versions in terms of sensitivity and/or specificity. Furthermore, with advancement in therapeutics (better risk-benefit profile of drugs) and need for recognizing and treating early disease in order to change the natural history of the disease, there is a current trend in newer classification criteria development, particularly for both systemic lupus and rheumatoid arthritis, of having increased sensitivity. For instance, the 1987 ACR RA criteria had low sensitivity for definitive RA diagnosis in early disease and subsequently the criteria were updated in 2010 with improved sensitivity in order to include early onset RA patients in clinical trials (1). However, with this increase sensitivity in lupus and RA classification criteria, there is a loss of specificity (20, 21), leading to further limitation of its use for clinical diagnosis. Ideal diagnostic criteria need to have very high (almost perfect, 100%) sensitivity and specificity, ensuring the clinician confidence in the diagnosis. By increasing sensitivity, early disease is included in the population for the classification criteria. However, this increased sensitivity risks including both undifferentiated arthritis that does not progress to rheumatoid arthritis and also subjects with incomplete lupus that do not develop systemic lupus. On the other hand, the inclusion of early disease gives the opportunity to develop newer therapeutics with improved outcomes and chance of cure. Moreover, the risk-benefit ratio of therapeutics needs to be considered when determining the sensitivity and specificity of criteria. For instance, in gout, the existing classification criteria have low specificity for early disease and should be used with caution in early disease when investigating agents with unclear safety profiles (8).

Furthermore, classification criteria change over time with development of new technology and longitudinal evaluation of disease. This process is most evident in the evolution of the classification criteria for systemic sclerosis (SSc). Originally, the 1980 ACR classification for systemic sclerosis were shown to have a low sensitivity with only 34 out of 54 patients with a clinical diagnosis of SSc fulfilling the classification criteria; these criteria were designed for research studies thus had high specificity but low sensitivity (22–24). High specificity of classification criteria is essential for clinical research of disease to create homogeneity amongst the subjects but the decreased sensitivity leaves the criteria without utility in patient care. Furthermore, stringent criteria are required, especially when cytotoxic treatment is necessary for severe disease, to ensure only uniform patient populations with disease are enrolled in clinical trials for therapeutic investigation. Over time and with increased understanding of disease, the SSc classification criteria have been updated to include extra-cutaneous manifestation of disease such as interstitial lung disease and pulmonary hypertension and also included advanced laboratory evidence of disease with SSc related antibodies (25). With this improvement of understanding of disease and updated criteria, the criteria have now much improved sensitivity and specificity (up to 96% sensitive and 90% specific) (25, 26). Given advances in understanding pathogenesis of disease, multiple older classification criteria, including vasculitis and myositis, are in the process of being updated along the same lines as the SSc criteria (27, 28). These updated criteria are likely to have better sensitivity as well as specificity due to improved knowledge as compared to few decades ago.

Although, traditionally classification criteria have high specificity, 1975 Bohan and Peter classification criteria for polymyositis (29) is an example where the classification criteria were too non-specific, leading to the wrong diagnosis of PM in patients with metabolic myopathy, muscle dystrophies and inclusion body myositis (30). These criteria only require a minimum of symmetrical proximal muscle weakness and abnormal serum skeletal muscle enzymes, myopathic EMG or biopsy for probable classification of PM. Given that PM often requires intensive and potentially toxic treatment with long courses of corticosteroids and additional immunosuppressive medications, clinicians should do due diligence of ruling out PM mimics before making clinical diagnosis, rather than blindly applying the classification criteria for diagnosis (19, 30).

Diagnostic Criteria

Diagnostic criteria are a conglomeration of signs, symptoms, or supportive tests used in routine clinical care to aid in a clinical diagnosis in an individual patient. Clinical diagnoses are used for two major purposes during the care of the individual patient: first to guide medical care and second to help the patient gain an understanding of the disease process with prognosis. The most well-known diagnostic criteria in medicine are in psychiatry for the Diagnostic and Statistical Manual of Mental Disorders (31). As there was poor agreement among providers regarding patient's psychiatric diagnosis, specific diagnostic criteria were developed to assist in patient care and diagnosis. The goal of diagnostic criteria is to have a high positive predictive value of the diagnosis and high likelihood ratio whereas the goal of classification criteria is to create a well-defined group of patients with the same underlying diagnosis.

Regarding the evolution of diagnostic criteria development, diagnostic criteria were initially developed to purely help guide the care of patients but later were translated for use in clinical research. Following the initial Jones diagnostic criteria development for rheumatic fever, clinical research realized more specific criteria were needed to prevent misclassification of non-disease patients and exposing them to therapeutic agents with unclear evidence in a clinical trials (4). As formal clinical research studies were developed, classification criteria soon followed to increase specificity for disease and minimize variation across study populations. Professional societies have focused on classification criteria development whereas diagnostic criteria have not been updated.

Given lack of optimal diagnostic criteria, the majority of rheumatic disease diagnoses are now made based on complex decision-making process by physicians-using a combination of symptoms, signs, available diagnostic tests, and ruling out other competing diagnosis while considering knowledge about geographic disease prevalence affecting pretest probability to make a diagnosis. One particular difficulty in the development of diagnostic criteria is that the disease prevalence in the population (different geographical areas or clinic settings) significantly affects the specificity and PPV of any criteria. For instance, in morning report at an academic medical center, a patient presenting with significant nausea, hepatomegaly and transaminitis, has a broad differential diagnosis requiring multiple diagnostic tests to evaluate the differential diagnosis (imaging, recurrent laboratory work, liver biopsy). However, the same patient presenting in endemic area for hepatitis A will be diagnosed and

treated as such given high prevalence of the disease in that area. These variations in disease frequency by geographic areas are especially true for some rheumatic diseases. For example, the prevalence of Takayasu vasculitis in Asian countries is very different than in USA, leading to vast difference in performance characteristics especially PPV of any criteria for diagnosis. There are also wide variations in rheumatic disease frequency in different clinical setting within same geographic areas such as academic referral center as compared to community rheumatology practice or primary care physician. The same concept applies to differences in disease prevalence in individual races or ethnicities within a geographic area.

Moreover, the rarity and heterogeneity of many rheumatic disease, given complex and variable clinic presentations, lack of gold standard tests, and reliance of multiple clinical factors rather than few makes the development of diagnostic criteria extremely difficult in rheumatology. Due to these limitations, only few diagnostic criteria have been developed in rheumatology with the most recent diagnostic criteria being the recently validated 2010 American College of Rheumatology preliminary Fibromyalgia Syndrome (FMS) diagnostic criteria (32, 33). However, these criteria have been criticized secondary to their lack of specificity (34). For instance, when the modified 2010 FMS criteria were applied to patients with cirrhosis, 53 of 193 patients (27%) met diagnostic criteria for FMS syndrome (35). Additionally, when the 2010 FMS diagnostic criteria were applied to 100 pregnant patients, 27 patients (27%) fulfilled the modified 2010 FMS criteria whereas the previous 1990 classification criteria, which were more specific, only 1/100 of the patient fulfilled (36). This example illustrates the difficulty in putting a complex decision made by a trained rheumatologist, who considers several factors simultaneously to make a clinical diagnosis, into a simple algorithm of diagnostic criteria.

Differences in Classification and Diagnostic Criteria

At first glance, diagnostic and classification criteria may seem similar but the two are distinct yet still complement each other. As discussed earlier, diagnostic and classification criteria are on a continuum for an individual disease with sensitivity and specificity inversely related along the receiver operating curve (4). Whereas diagnostic criteria are specifically designed for the individual, classification criteria create well defined groups of subjects for clinical research. Diagnostic criteria are meant to be broad with high sensitivity and negative predictive value as to not exclude individuals with possible disease. Additionally, diagnostic criteria ideally also have very high specificity to prevent wrong diagnosis, whereas classification criteria are mean to have very high specificity and positive predictive value as to include patients only with definite disease. In a disease with very high sensitivity and specificity, diagnostic and classification criteria may be similar or the same. For instance, in acute gout, where intracellular monosodium urate crystals by fluorescent microscopy in synovial fluid in the setting of inflammatory arthritis clinically is diagnostic of acute gout with high sensitivity and specificity, diagnostic and classification criteria could potentially be the same when a gold standard test such as synovial fluid analysis is required for classification; new classification criteria of gout are in development for gout as previous classification criteria were not validated (37). However, despite ideal criteria, it does not mean that physicians cannot make a clinical diagnosis of gout without arthrocentesis in a patient with risk factors of gout presenting with classic podagra. It may not be confirmed or

definite gout diagnosis, but is enough to treat the patient with NSAIDs. Whenever a physician decides to treat, he/she has made a diagnosis (knowingly or unknowingly) even if it is not definite. In contrast to the example of acute gout, in a disease such as chronic gouty arthropathy, considerable disagreement will occur between diagnostic and classification criteria given the ambiguity of diagnosis. Classification criteria would need to be much more specific for clinical research studies but diagnostic criteria could be broad to aid the clinician in diagnosis and prescribing non-toxic treatment. The clinical example of acute gout versus chronic gouty arthropathy highlights continuum of diagnostic and classification criteria.

Unlike malignancy and infection, rheumatic diseases evolve over time and cross-sectional application of any criteria as either “disease present or absent” is too simplistic. Rheumatic diseases are dynamic and at a single point in time, a patient may not fit the classification criteria for a given disease but still have the clinical diagnosis. Moreover, rheumatic diseases may present at undifferentiated stage which may or may not later evolve into more established syndrome. Undifferentiated arthritis is an example of this phenomenon of differentiating into separate diagnosis such as psoriatic arthritis or rheumatoid arthritis. Following initially presenting with undifferentiated arthritis, 32–53% of patients remained unclassified after one year of observation and in another follow up study of 270 patients with early undifferentiated arthritis, 23% remained unclassified at 3 years and the diagnosis changed 46% of the time between the first and last examination (38, 39).

Heterogeneity of rheumatic disease also highlights differences between making diagnostic and classification criteria. Systemic lupus erythematosus (SLE), perhaps the most heterogeneous rheumatologic disease, requires multiple clinical characteristics to classify patients with systemic lupus (4/11 criteria with the revised 1982 criteria and 4/11 clinical criteria and in addition 1 immunological criteria for the 2012 SLICC criteria) (40–42). Many patients with SLE may not have had the entire spectrum of clinical features at the time of presentation or may present with uncommon features not considered in classification criteria, but still need the bedside clinical diagnosis as well as treatment by a physician (43). Diagnosis, even if presumptive, is a prerequisite for treatment. Does a patient with a high titer + ANA, anti-DS DNA, malar rash and Raynaud’s phenomenon not have SLE given no other clinical features? This patient would not meet current SLE classification criteria but most clinicians would diagnosis the patient with SLE and treat accordingly.

Use

Limitations of classification criteria in clinical practice have thus been described, but their use in the practice of rheumatology is commonplace and should not be negated. The art of medicine with differentiation of the gray from the black and white is essential in clinical diagnosis and an integral part of the rheumatology subspecialty. While classification criteria remove part of the art required in diagnosis and can often be misleading if used inappropriately for diagnosis in each individual case, they can be helpful to guide clinicians in diagnosis and are complementary in many ways in following areas of rheumatology:

Education and training

In addition to the obvious use of classification criteria in clinical research, published classification criteria are useful in medical education for disseminating knowledge of disease process not only in common rheumatic disease processes such as rheumatoid arthritis, osteoarthritis, and systemic lupus but also in rare diseases such as Behcet's disease and Takayasu's Arteritis that are not seen in daily medical practice. Attending physicians commonly teach medical students and residents key clinical features of a rheumatic disease from classification criteria. It often serves as the first step for a new learner to understand the most common presentation, clinical features, which diagnostic tests are required, etc. Moreover, classification criteria are a segway for discussion during rounds to highlight differences in clinical diagnosis and decision making as compared to applying a criteria for research as well as to explain concepts of sensitivity, specificity, PPV, NPV, etc. In addition to medical education and understanding, classification criteria are important for *patient* understanding of disease process and prognosis. Often derived from longitudinal cohorts that have defined the course of clinical disease features and change in disease outcomes over time (19, 44), classification criteria have improved understanding of natural history and pathogenesis of rheumatic disease. Therefore, developed classification criteria, which have been validated across geographically and broad patient populations, are useful to gain understanding of disease for clinicians, patients, and researchers.

Research

Classification criteria are essential for clinical research of rheumatic diseases and have an impact in 3 areas of rheumatology research.

a) Defining the Disease—For initial study of disease, classification criteria enable clinician researchers in various parts of the world to investigate disease in a uniform manner, which facilitates understanding pathophysiological basis for disease, natural history of disease, and epidemiology (3, 13, 18). Use of classification criteria creates a common language and shared interest among scientists in multiple geographic areas about a well-defined homogenous population of a particular disease. Moreover, classification criteria facilitate knowledge transfer between investigators, multi-center collaborative studies and advancement of the field. Rather than multiple research programs running in parallel, each working on a different ill-defined group of patients yet all calling it the same disease, classification criteria allow basic and clinical research to progress in a sequential multicenter, and even multi-specialty, manner. For example, new Sjogren's Syndrome classification criteria brought 3 different specialist including rheumatology, oral medicine and ophthalmology, on a common platform regarding the disease (45).

b) Developing Therapeutics—Classification criteria aid in therapeutics research. Classification criteria are a prerequisite for defining inclusion criteria for clinical trials. Furthermore, achieving a well-defined homogenous disease population, as classification criteria do, is necessary in order for multiple therapeutic studies and populations to be compared or combined. Multi-center clinical trials, which in the last decade have provided various therapeutic options for rheumatic patients, would not have been possible without the use of classification criteria in patient selection. For example, rheumatoid arthritis today has

9 approved biologic therapies in addition to multiple oral DMARDs, due to large international multicenter clinical trials facilitated by classification criteria (46). Moreover, therapeutic investigation findings in one part of the world can be used by clinicians to treat patients in other parts of the world if their patients have clinical characteristics meeting classification criteria used in the original studies. Following development of classification criteria, response criteria development is next logical step and is essential for comparative effectiveness research and defining outcomes of disease (47). As RA is the leader in development of classification criteria and response criteria, response criteria development will be the future for all diseases after development of classification criteria.

c) Innovative Areas of research—Evolution of classification criteria overtime can lead to new areas of research. Initially, when a disease is rare and less defined, classification criteria are useful for creating homogenous populations for comparison and study of natural history of disease. Subsequently, as disease understanding advances, classification criteria, as in SSc, can be updated to include scleroderma specific antibodies and other specific diagnostic tools (25, 48). Additionally, classification criteria can change over time, enabling a shift focus to new areas of research as occurred in RA. In RA, as the disease has uniformly been defined and improvement in outcome of disease with multiple disease modifying anti-rheumatic drugs, classification criteria have recently been targeted for early disease in aims of improving early diagnosis and treatment. This change in classification criteria focus from investigation of established disease to early disease enables innovative areas of research (1, 49, 50).

Clinical Care

The use of classification criteria in clinical care for diagnosis is limited and should be cautioned as explained above. Classification criteria provide a framework of common clinical features and tests of a disease to the clinician which he/she can use to consider a disease diagnosis in question along with other clinical information about pretest probability, differential diagnosis, available tests, clinical features, etc. Also, classification criteria can be a useful tool clinically for ruling in common rheumatic disease. In a common disease (i.e. high prevalence in the geographic area and clinical setting on which the classification criteria is to be applied) given high specificity of the classification criteria, the PPV of classification criteria for diagnosis is usually high. Thus, clinicians can use the classification criteria to rule in a diagnosis of common disease like RA, sjogren syndrome, SLE, etc, but not rule out a diagnosis if classification criteria are not met. We still advise careful consideration given to rule out mimickers of the disease, even if classification criteria are met. For example, RA classification criteria can be safely used to make a clinical diagnosis of RA in a patient presenting with polyarthritis, after ruling out other possibilities such as hepatitis C, psoriatic arthritis, etc. Moreover, classification criteria are appropriately used to decide on future investigation or tests to rule in or rule out various differentials. For example, a primary care physician may order anti-CCP in a patient who present with polyarthritis, after reading classification criteria for RA.

Abuse

The use of classification criteria has been described but the primary abuse of classification criteria is its use, or misuse, in making a diagnosis. This abuse of using classification criteria in making a diagnosis is explained in various clinical scenarios below.

Ruling out a common disease based on the classification criteria

Given the traditional lower sensitivity and homogenous disease representation by classification criteria, a clinical diagnosis should not be ruled out based on not meeting the classification criteria. For instance, in rheumatoid arthritis, the 2010 ACR/EULAR classification criteria showed a relatively high specificity and PPV in early arthritis cohorts, but the NPV varied from 36–78% (50). Classification criteria do not have 100% sensitivity; hence simply put, cannot be used in individual patients to rule out a diagnosis if it is a relatively common disorder.

Ruling in or out a rare disease based on the classification criteria

Use of classification criteria differs between rare and common disease as disease prevalence affects the positive predictive value when classification criteria are applied to individual patients (4, 12). When diseases have a very low frequency in a population, clinicians often cannot depend on classification criteria because even excellent classification criteria (95%) may lead to misclassification in too many patients. For example estimated frequency of Takayasu vasculitis in US is 2.6 per million (51–53) and even if we have a classification criteria which is 95% sensitive and specific, the positive predictive value of that criteria in the US will be extremely low. On the other hand negative predictive value will be excellent regardless of the sensitivity or specificity of the classification criteria in such a population. Therefore, clinicians should use their sound knowledge, clinical judgment and experience to make a diagnosis in such rare conditions.

Insensitive for mild or uncommon presentation of disease

Classification Criteria are not traditionally designed for uncommon presentations or milder phenotypes of disease and rather are designed to create homogenous populations of common and severe phenotypes. Therefore, in rare or early presentation of disease, the use of classification criteria would be an abuse for the patient. For instance, although the 2013 SSc classification now have greater than 90% sensitivity and specificity, the criteria would still fail to classify a patient with systemic sclerosis presenting with new onset raynaud's phenomenon, abnormal nail fold capillaries, tendon friction rubs, and a scleroderma specific antibody although most clinicians would diagnose this patient with early SSc (25). Furthermore, for undifferentiated rheumatic disease, such as in undifferentiated arthritis or incomplete lupus, classification criteria do not have the sensitivity to accurately capture the disease process (54, 55).

Resources and feasibility of diagnosis

Rheumatologist in several parts of the world, especially in under developed or developing countries or even in underserved areas of developed countries, may be faced with limited access to or affordability of testing (due to patient's own financial and/or insurance

limitations) and may not be able to do every test to fulfill classification criteria for diagnosis. For example, many clinicians would make the diagnosis of rheumatoid arthritis in a poor patient without requiring rheumatoid factor or anti-CCP in a developing country and prescribe appropriate treatment. Moreover, patient preferences and overall health conditions may prevent patients to fulfill every criterion as per classification criteria if used for diagnosis. For example, patients with suspected sjogren syndrome may refuse minor salivary gland biopsy, but the physician still often makes a clinical diagnosis and prescribes appropriate treatment. Requiring fulfillment of classification criteria to make a clinical diagnosis would be wrong in these situations. More stringent classification criteria for diagnosis that require particular laboratory tests, imaging, or surgical procedures could constitute a hurdle for patients and clinicians and has the potential to postpone the initiation of effective therapy.

Legal, financial and treatment implications of diagnosis

Using classification criteria strictly for diagnosis may have significant negative legal, financial and treatment implication on our patients and should be avoided by any doctor, health care system or governing bodies. Even the highly sensitive classification criteria in rheumatology can never be 100% sensitive, and subsequently if used as diagnostic criteria will leave some patients as undiagnosed. This means that doctors or hospitals may deny the needed treatment to these patients, or insurance coverage may be refused if insurance companies and government agencies use the classification criteria as a standard for reimbursement. On the other hand, no single criterion is 100% specific, so some patients may be incorrectly diagnosed as with an illness leading to problems with health or life insurances as well as exposing them to unnecessary or incorrect, potentially harmful therapies.

Complex decision making for diagnosis

Often clinicians will use complex multi-step decision making in a clinical diagnosis. Just as the RA patient presenting without active synovial swelling after aggressive corticosteroid use or the ICU patient with pulmonary-renal syndrome too sick to undergo biopsies to confirm a diagnosis of vasculitis, clinical diagnosis includes balancing pre and post-test probability of the disease with multiple other factors including severity of disease, risk of further testing or treatment and ruling out appropriate mimickers on the differential diagnosis. Finding the white and black out of the gray clinical situation, or the art of clinical diagnosis, is not possible with highly specific classification criteria, but is required daily in clinical medicine. The knowledge from classification criteria can guide the clinician in diagnosis but only as a guide, not a rule.

Classification criteria are often applied to disease in which they are not intended

In patients with peripheral seronegative spondyloarthropathies, a RA diagnosis is often used by the clinician secondary to a low titer rheumatoid factor in order to justify medication use despite asymmetrical inflammatory arthropathy without erosive features and clinical features of a spondyloarthropathy rather than rheumatoid arthritis. While the means to the end, being medication approval, is used as justification of the clinician for this classification criteria misuse, there are many detrimental effects from this misuse of classification criteria.

Practically, this false use of the RA diagnosis hurts both patients' understanding of disease process and also provider-provider communication given EMR documentation of rheumatoid arthritis rather than spondyloarthropathy. Ethically, rather than using a false RA diagnosis, clinicians should argue for medication approval for the treatment of spondyloarthropathy as such and not misuse the RA classification criteria.

Controversial Issues

The controversy between diagnostic and classification criteria is longstanding and likely will continue. Only in certain rheumatic disease such as acute gout, where there is a potential of high sensitivity and specificity, is a convergence of the two types of criteria possible to occur. Clinicians currently use the guidance of classification criteria for clinical diagnosis but understanding all the specific psychometric properties on the underlying cohort in which the criteria were developed and subsequent validation studies, will lead to better use of these criteria by practicing clinicians. Increased transparency in the methods of cohort design, classification criteria development, laboratory or other test development as well as dissemination of knowledge about using or not using classification criteria in rheumatology office is necessary to understand the true clinical and diagnostic applicability of classification criteria. For example, understanding difference between new and old RA classification criteria is critical for clinicians as both are currently valid but have different purposes. The 2010 ACR RA classification criteria are best applied early polyarthritis patients rather than to long-term RA patients (1). These patients with end stage RA may not have active swollen or tender joints but rather ulnar deviation, swan neck deformities and rheumatoid nodules, and the 1987 ACR RA criteria would be better served in their diagnosis rather than the 2010 criteria (49).

In addition to transparency of classification criteria, Rheumatology has been plagued by lack of transparency and standardization in the laboratory testing of rheumatologic labs for example ANA. The method of ANA testing varies considerably amongst laboratories and subsequently, whether indirect immunofluorescence versus multiplex assay method, can vastly affect the sensitivity in SLE (56), leading to difficulty when assessing presence or absence of disease and SLE diagnosis. CCP is highly specific for rheumatoid arthritis in patients with chronic inflammatory arthritis, but its utility and clinical significance outside of inflammatory arthritis is not necessarily known or understood (2). Any classification or diagnostic criteria are only as good as its individual criterion such as CCP in RA or ANA in SLE, and lack of standardization or transparency in rheumatic diagnostic testing seriously hampers appropriate clinical use of these criteria.

Summary

Rheumatic diseases lack a gold standard test with 100% sensitivity and specificity, making it necessary to develop classification and diagnostic criteria to guide clinicians and researchers. Classification criteria are used as standardized means for defining homogenous cohorts in research studies, whereas diagnosis is determination of illness in an individual after examining all clinical features and the differential diagnosis, keeping in mind the epidemiology of the specific clinical setting. The prevalence of disease in the area of

practice directly affects the PPV and utility of any criteria when used for diagnosis. Improved understanding of disease pathogenesis and development of new diagnostic tools has led to better classification criteria with increased sensitivity and/or specificity. In general, despite a recent trend of better psychometric properties of newer classification criteria, it is still difficult to capture the full range of disease manifestations by any criteria due to the trade-off between sensitivity and specificity. Any diagnostic criteria, given the heterogeneous nature of rheumatic disease and lack of 100% sensitivity and specificity, may lead to both underrepresentation of disease and incomplete description of the breadth of disease treated by rheumatology (12).

As few accepted diagnostic criteria exist for rheumatic disease, clinicians often depend on classification criteria for guidance in making a clinical diagnosis, which may not be the best strategy in all cases. Rather than inappropriately generalizing classification criteria for diagnosis in order to use, rather than abuse, classification criteria for individual patient diagnostic decisions, clinicians need to understand the underlying psychometric properties in the classification criteria and cohort characteristics in which the criteria were developed. This controversy and balance between diagnostic and classification criteria will remain as diagnosis of a complicated multisystem rheumatic disease inherently involves a highly complex cognitive process requiring synthesis of many clinical factors, typically beyond a simple algorithm-based set of criteria. Nonetheless, classification and diagnostic criteria both play central roles in rheumatology and with improved understanding of their use and abuse they can help us better serve our patients.

Acknowledgements

Dr. June's work on this publication was supported, in part, by Grant UL1 TR000127 and KL2 TR000126 from the National Center for Advancing Translational Sciences (NCATS).

References

1. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO 3rd, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis and rheumatism*. 2010 Sep; 62(9):2569–2581. PubMed PMID: 20872595. Epub 2010/09/28. eng. [PubMed: 20872595]
2. Aggarwal R, Liao K, Nair R, Ringold S, Costenbader KH. Anti-citrullinated peptide antibody assays and their role in the diagnosis of rheumatoid arthritis. *Arthritis Rheum*. 2009 Nov; 61(11):1472–1483. PubMed PMID: 19877103. Pubmed Central PMCID: PMC2859449. eng. [PubMed: 19877103]
3. 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 and rheumatism*. 2007 Oct 15; 57(7):1119–1133. PubMed PMID: 17907227. [PubMed: 17907227]
4. Yazici H. Diagnostic versus classification criteria - a continuum. *Bull NYU Hosp Jt Dis*. 2009; 67(2):206–208. PubMed PMID: 19583555. eng. [PubMed: 19583555]
5. Lee DM, Schur PH. Clinical utility of the anti-CCP assay in patients with rheumatic diseases. *Ann Rheum Dis*. 2003 Sep; 62(9):870–874. PubMed PMID: 12922961. Pubmed Central PMCID: PMC1754666. eng. [PubMed: 12922961]
6. van Gaalen FA, Linn-Rasker SP, van Venrooij WJ, de Jong BA, Breedveld FC, Verweij CL, et al. Autoantibodies to cyclic citrullinated peptides predict progression to rheumatoid arthritis in patients with undifferentiated arthritis: a prospective cohort study. *Arthritis Rheum*. 2004 Mar; 50(3):709–715. PubMed PMID: 15022309. eng. [PubMed: 15022309]

7. Whiting PF, Smidt N, Sterne JA, Harbord R, Burton A, Burke M, et al. Systematic review: accuracy of anti-citrullinated Peptide antibodies for diagnosing rheumatoid arthritis. *Ann Intern Med.* 2010 Apr; 152(7):456–464. PubMed PMID: 20368651. eng. [PubMed: 20368651]
8. 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. *Annals of the rheumatic diseases.* 2014 Oct.;28. PubMed PMID: 25351521.
9. Dougados M, Gossec L. Classification criteria for rheumatic diseases: why and how? *Arthritis and rheumatism.* 2007 Oct 15; 57(7):1112–1115. PubMed PMID: 17907226. Epub 2007/10/02. eng. [PubMed: 17907226]
10. Yurdakul S, Günaydin I, Tüzün Y, Tankurt N, Pazarli H, Ozyazgan Y, et al. The prevalence of Behçet's syndrome in a rural area in northern Turkey. *J Rheumatol.* 1988; 15(5):820–822. PubMed PMID: 3172095. eng. [PubMed: 3172095]
11. Criteria for diagnosis of Behçet's disease. International Study Group for Behçet's Disease. *Lancet.* 1990 May; 335(8697):1078–1080. PubMed PMID: 1970380. eng. [PubMed: 1970380]
12. Hunder GG. The use and misuse of classification and diagnostic criteria for complex diseases. *Ann Intern Med.* 1998 Sep; 129(5):417–418. PubMed PMID: 9735071. eng. [PubMed: 9735071]
13. Felson DT, Anderson JJ. Methodological and statistical approaches to criteria development in rheumatic diseases. *Baillieres Clin Rheumatol.* 1995 May; 9(2):253–266. PubMed PMID: 7656339. Epub 1995/05/01. eng. [PubMed: 7656339]
14. 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 Sep; 129(5):345–352. PubMed PMID: 9735061. eng. [PubMed: 9735061]
15. Sørensen 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 Jun; 59(6):478–482. PubMed PMID: 10834866. Pubmed Central PMCID: PMC1753166. eng. [PubMed: 10834866]
16. 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. PubMed PMID: 19250526. Pubmed Central PMCID: 2654892. Epub 2009/03/03. eng. [PubMed: 19250526]
17. 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 Feb; 22(1):96–102. PubMed PMID: 15640301. Epub 2005/01/11. eng. [PubMed: 15640301]
18. Singh JA, Solomon DH, Dougados M, Felson D, Hawker G, Katz P, et al. Development of classification and response criteria for rheumatic diseases. *Arthritis Rheum.* 2006 Jun; 55(3):348–352. PubMed PMID: 16739201. eng. [PubMed: 16739201]
19. Medsger TA, Oddis CV. Classification and diagnostic criteria for polymyositis and dermatomyositis. *J Rheumatol.* 1995 Apr; 22(4):581–585. PubMed PMID: 7791146. eng. [PubMed: 7791146]
20. Mjaavatten MD, Bykerk VP. Early rheumatoid arthritis: the performance of the 2010 ACR/ EULAR criteria for diagnosing RA. *Best Pract Res Clin Rheumatol.* 2013 Aug; 27(4):451–466. PubMed PMID: 24315048. eng. [PubMed: 24315048]
21. Inês L, Silva C, Galindo M, López-Longo FJ, Terroso G, Romão VC, et al. Classification of Systemic lupus erythematosus: Systemic Lupus International Collaborating Clinics versus American College of Rheumatology criteria. *Arthritis Care Res (Hoboken).* 2015 Jan. PubMed PMID: 25581417. ENG.
22. 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 Oct 20; 137(41–42):586–590. PubMed PMID: 17990152. Epub 2007/11/09. eng. [PubMed: 17990152]
23. LeRoy EC, Maricq HR, Kahaleh MB. Undifferentiated connective tissue syndromes. *Arthritis Rheum.* 1980 Mar; 23(3):341–343. PubMed PMID: 7362686. eng. [PubMed: 7362686]

24. Maricq HR, LeRoy EC, D'Angelo WA, Medsger TA, Rodnan GP, Sharp GC, et al. Diagnostic potential of in vivo capillary microscopy in scleroderma and related disorders. *Arthritis Rheum.* 1980 Feb; 23(2):183–189. PubMed PMID: 7362667. eng. [PubMed: 7362667]
25. 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 Nov; 65(11):2737–2747. PubMed PMID: 24122180. Pubmed Central PMCID: PMC3930146. eng. [PubMed: 24122180]
26. Hoffmann-Vold AM, Gunnarsson R, Garen T, Midtvedt Ø, Molberg Ø. Performance of the 2013 American College of Rheumatology/European League Against Rheumatism Classification Criteria for Systemic Sclerosis (SSc) in Large, Well-defined Cohorts of SSc and Mixed Connective Tissue Disease. *J Rheumatol.* 2015 Jan; 42(1):60–63. PubMed PMID: 25274890. eng. [PubMed: 25274890]
27. Zong M, Lundberg IE. Pathogenesis, classification and treatment of inflammatory myopathies. *Nat Rev Rheumatol.* 2011 May; 7(5):297–306. PubMed PMID: 21468145. eng. [PubMed: 21468145]
28. 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). Clinical and experimental nephrology. 2013 Oct; 17(5):619–621. PubMed PMID: 23996327. [PubMed: 23996327]
29. Bernstein RM, Morgan SH, Chapman J, Bunn CC, Mathews MB, Turner-Warwick M, et al. Anti-Jo-1 antibody: a marker for myositis with interstitial lung disease. *Br Med J (Clin Res Ed).* 1984 Jul 21; 289(6438):151–152. PubMed PMID: 6430389. Pubmed Central PMCID: 1442006.
30. Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). *N Engl J Med.* 1975 Feb; 292(7):344–347. PubMed PMID: 1090839. eng. [PubMed: 1090839]
31. American Psychiatric Association. American Psychiatric Association. DSM-5 Task Force. Diagnostic and statistical manual of mental disorders : DSM-5. 5th ed. Vol. xlv. Washington, D.C: American Psychiatric Association; 2013. p. 947
32. Bennett RM, Friend R, Marcus D, Bernstein C, Han BK, Yachoui R, et al. Criteria for the diagnosis of fibromyalgia: validation of the modified 2010 preliminary American College of Rheumatology criteria and the development of alternative criteria. *Arthritis Care Res (Hoboken).* 2014 Sep; 66(9):1364–1373. PubMed PMID: 24497443. eng. [PubMed: 24497443]
33. Wolfe F, Clauw DJ, Fitzcharles MA, 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 May; 62(5):600–610. PubMed PMID: 20461783. Epub 2010/05/13. eng. [PubMed: 20461783]
34. McBeth J, Mulvey MR. Fibromyalgia: mechanisms and potential impact of the ACR 2010 classification criteria. *Nat Rev Rheumatol.* 2012 Feb; 8(2):108–116. PubMed PMID: 22270077. eng. [PubMed: 22270077]
35. Rogal SS, Bielefeldt K, Wasan AD, Szigethy E, Lotrich F, DiMartini AF. Fibromyalgia Symptoms and Cirrhosis. *Dig Dis Sci.* 2014 Nov. PubMed PMID: 25433921. ENG.
36. Saa'd S, Many A, Jacob G, Ablin JN. High prevalence of fibromyalgia symptoms among healthy full-term pregnant women. *Rheumatol Int.* 2013 Jun; 33(6):1555–1560. PubMed PMID: 23263499. eng. [PubMed: 23263499]
37. Dalbeth N, Fransen J, Jansen TL, Neogi T, Schumacher HR, Taylor WJ. New classification criteria for gout: a framework for progress. *Rheumatology.* 2013 Oct; 52(10):1748–1753. PubMed PMID: 23611919. Pubmed Central PMCID: 4047282. [PubMed: 23611919]
38. 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 Oct; 37(10):1084–1088. PubMed PMID: 9825747. eng. [PubMed: 9825747]
39. 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 May; 28(5):975–981. PubMed PMID: 11361225. eng. [PubMed: 11361225]

40. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* 1997 Sep;40(9):1725. PubMed PMID: 9324032. eng. [PubMed: 9324032]
41. Petri M, Orbai AM, Alarcón GS, Gordon C, Merrill JT, Fortin PR, et al. Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. *Arthritis Rheum.* 2012 Aug; 64(8):2677–2686. PubMed PMID: 22553077. Pubmed Central PMCID: PMC3409311. eng. [PubMed: 22553077]
42. Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* 1982 Nov; 25(11): 1271–1277. PubMed PMID: 7138600. eng. [PubMed: 7138600]
43. Panush RS, Schur PH. It is lupus? *Bull Rheum Dis.* 1997 Oct; 46(6):3–8. PubMed PMID: 9336999. eng. [PubMed: 9336999]
44. Steen VD, Medsger TA. Changes in causes of death in systemic sclerosis, 1972–2002. *Ann Rheum Dis.* 2007 Jul; 66(7):940–944. PubMed PMID: 17329309. Pubmed Central PMCID: PMC1955114. eng. [PubMed: 17329309]
45. Shiboski SC, Shiboski CH, Criswell L, Baer A, Challacombe S, Lanfranchi H, et al. American College of Rheumatology classification criteria for Sjogren's syndrome: a data-driven, expert consensus approach in the Sjogren's International Collaborative Clinical Alliance cohort. *Arthritis Care Res (Hoboken).* 2012 Apr; 64(4):475–487. PubMed PMID: 22563590. Pubmed Central PMCID: 3349440. Epub 2012/05/09. eng. [PubMed: 22563590]
46. Singh JA, Furst DE, Bharat A, Curtis JR, Kavanaugh AF, Kremer JM, et al. 2012 update of the 2008 American College of Rheumatology recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis. *Arthritis Care Res (Hoboken).* 2012 May; 64(5):625–639. PubMed PMID: 22473917. Epub 2012/04/05. eng. [PubMed: 22473917]
47. Anderson J, Caplan L, Yazdany J, Robbins ML, Neogi T, Michaud K, et al. Rheumatoid arthritis disease activity measures: American College of Rheumatology recommendations for use in clinical practice. *Arthritis Care Res (Hoboken).* 2012 May; 64(5):640–647. PubMed PMID: 22473918. Epub 2012/04/05. eng. [PubMed: 22473918]
48. Walker JG, Pope J, Baron M, Leclercq S, Hudson M, Taillefer S, et al. The development of systemic sclerosis classification criteria. *Clin Rheumatol.* 2007 Sep; 26(9):1401–1409. PubMed PMID: 17285223. eng. [PubMed: 17285223]
49. 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 Mar; 31(3):315–324. PubMed PMID: 3358796. eng. [PubMed: 3358796]
50. van der Helm-van Mil AH, Huizinga TW. The 2010 ACR/EULAR criteria for rheumatoid arthritis: do they affect the classification or diagnosis of rheumatoid arthritis? *Ann Rheum Dis.* 2012 Oct; 71(10):1596–1598. PubMed PMID: 22580580. eng. [PubMed: 22580580]
51. Richards BL, March L, Gabriel SE. Epidemiology of large-vessel vasculidities. *Best Pract Res Clin Rheumatol.* 2010 Dec; 24(6):871–883. PubMed PMID: 21665132. eng. [PubMed: 21665132]
52. Hall S, Barr W, Lie JT, Stanson AW, Kazmier FJ, Hunder GG. Takayasu arteritis. A study of 32 North American patients. *Medicine (Baltimore).* 1985 Mar; 64(2):89–99. PubMed PMID: 2858047. eng. [PubMed: 2858047]
53. de Souza AW, de Carvalho JF. Diagnostic and classification criteria of Takayasu arteritis. *J Autoimmun.* 2014 Feb-Mar;48–49:79–83. PubMed PMID: 24461381. eng.
54. Krabben A, Huizinga TW, van der Helm-van Mil AH. Undifferentiated arthritis characteristics and outcomes when applying the 2010 and 1987 criteria for rheumatoid arthritis. *Ann Rheum Dis.* 2012 Feb; 71(2):238–241. PubMed PMID: 21953339. eng. [PubMed: 21953339]
55. Costenbader KH, Schur PH. We need better classification and terminology for “people at high-risk of or in the process of developing lupus”. *Arthritis Care Res (Hoboken).* 2014 Oct. PubMed PMID: 25302656. ENG.
56. Bruner BF, Guthridge JM, Lu R, Vidal G, Kelly JA, Robertson JM, et al. Comparison of autoantibody specificities between traditional and bead-based assays in a large, diverse collection of patients with systemic lupus erythematosus and family members. *Arthritis Rheum.* 2012 Nov;

64(11):3677–3686. PubMed PMID: 23112091. Pubmed Central PMCID: 3490432. Epub 2012/11/01. eng. [PubMed: 23112091]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Practice points

- Meeting disease classification criteria is not necessary for clinical diagnosis or treatment
- Classification criteria are designed for clinical research to create well defined relatively homogenous groups for disease investigation
- Classification criteria should be used as a guide to clinicians in clinical diagnosis and evaluation of common diseases, but clinicians should use their clinical knowledge for making final diagnosis
- Physicians should understand the sensitivity, specificity, positive predictive value, negative predictive values and influence of epidemiology on any criteria before applying them
- Clinicians and researchers should be aware of the trend of decreasing specificity with increased sensitivity of some of the newer classification criteria—understanding that more early disease may be classified with risk of misclassification

Research agenda

- Further development and validation of classification criteria for rare diseases is needed for clinical trials of therapeutic drug development
- Transparency of laboratory testing for sensitivity and specificity in specific patient populations is necessary to aid clinical diagnosis
- Diagnostic criteria development needs to be sensitive to encompass all individuals with disease and validated across diverse patient populations