A Critical Look at Diagnostic Criteria
Time for a Change?

Hasan Yazici, M.D.

Abstract
There are certain thought barriers involved in making diagnostic-classification criteria in diseases of unknown origin. Among these are a lack of appreciation of the issue of circular logic, the basic oneness of diagnostic and classification criteria, the lack of appreciation as to why we make such criteria in the first place, and the lack of importance informing our patients that we do as well as should treat them without a firm diagnosis in many instances. The relevance of these thought barriers to the new American College of Rheumatology/European Union League Against Rheumatism (ACR/EULAR) Rheumatoid Arthritis (RA) classification criteria are also discussed.

In a series of talks and papers about methodology, I had already talked and written about one popular pastime of our profession, that of making diagnostic or classification criteria for diseases of unknown etiology and pathogenesis. This is obviously very pertinent when we have yet to discern distinct causes for a majority of major illnesses that we currently manage. What brings me back to the same issue is 1. on reflecting back, perhaps I was not clear enough about why I considered diagnosis and classification as basically the same, and 2. since my last paper, the new ACR/EULAR classification criteria for RA has appeared, a critical look at which will perhaps better explain my objectives.

I had underlined in a previous paper that diagnosis and classification are on a continuum. To me, there is actually no difference in the mental or mathematic activity in the formulation of criteria for diagnosis, classification or both. Diagnosis is classification in the individual patient. The comparison groups, the odds ratios, and the pre-test probabilities are the essential components of the mental process. Again, as I had emphasized the pre-test (or diagnostic criteria) probabilities are the two most essential components of whether we would name a set of criteria as diagnostic or classification. Perhaps a further look into the origins of the two words will better clarify the issue.

The verb “to diagnose” is originally Greek, and it means “to discern.” Thus, it also has the connotation of “understanding the nature of.” The verb “to classify,” on the other hand, is more contemporary. It simply means to allocate into classes. As such, it implies less precision and perhaps less of an attempt “to better comprehend the nature of.” When confronted with a patient, we prefer to diagnose and when we do research, we like to classify. However, I respectfully suggest that this is a bit of double talk. I suggest that there is no sound reason to say that we have to be more precise in making a diagnosis in every day practice as compared to when we enroll a patient into a drug trial.

As the traditional healer, we are of the opinion that the patient expects a diagnosis rather than a classification from us. This is perhaps, we still maintain, as our predecessors of the previous centuries did, ours is a profession of “omniscience.” We all like to give a patient a diagnosis; in other words, we are not content with a classification only—we also want to attach perhaps an “omniscience” to what we classify. On the other hand, we as rheumatologists painfully know that while we do not know the exact nature of many of the illnesses we recognize and manage, more recently and happily, we have been rather successful in their management. This brings me directly to the issue of the new ACR/EULAR criteria.

I am concerned there is much that is critical that needs to be said about the new RA criteria. The whole exercise starts with a group of experts from each side of the Atlantic. The
experts are of even number and similar gender distribution between these different sides of the Atlantic. This emphasis by the ACR/EULAR on equal numbers across continents and gender suggests that the data is therefore more accurate. This message might indeed impress the third-party payers or the policy makers in Washington, D.C. or Brussels but surely is out of context in a rheumatology journal. Are there data that indicate that one nationality or gender is different from their counterparts in managing RA?

Also the related manuscripts fail to address the nagging question of how many patients in daily practice with psoriasis and early onset arthritis are diagnosed as psoriatic arthritis when they present, while they eventually turn out to have RA in the months following. Similarly, how many young females with synovitis of the small joints of the hand, Reynaud’s phenomenon, and a positive ANA are initially diagnosed with systemic lupus erythematosus (SLE) and, in time, turn out to have erosive RA?

There are still other important issues related to the way a consensus was reached in assessing the disease severity, actually showing pronounced variance among the paper patients. Furthermore, when validity estimations were made, the reader was not given formal statistics that compared disease severity between, this time actual, patients from different cohorts. However, I will digress here from the disease severity between, this time actual, patients from different cohorts. However, I will digress here from the ACR/EULAR RA criteria and try to make a list of thought barriers in criteria making for diseases of unknown origin. They surely apply to the ACR/EULAR example at hand.

As depicted in Table 1, the first issue is trying to avoid circularity in thinking. This is unrealistic. Fries indicated such many years ago:5

Presence of disease “criteria” affirms or ignorance of the essence of disease. If we understand a disease, we can ascribe the elements that are necessary and sufficient for its diagnosis. One can so define gouty arthritis, in which joint fluid crystals serve as a “gold standard” against which to measure the usefulness of other observations. No other major rheumatic disease, including SLE, has such a standard. Thus, criteria must be constructed in a circular manner, by testing variables against a diagnosis based on intuition [emphasis added]. The “best” criteria therefore only describe the current conventional wisdom in an efficient manner.

I will go even further and suggest that all diagnoses and classifications, even when specific histology and microbiology are involved, are based on definitions and, thus, have some “circularity.” However, circular reasoning is only present when the conclusion is nothing more than a reiteration of the premise(s) of the reasoning. The following set of statements can be considered in this regard:

1. Our rheumatology unit does not make a diagnosis of RA unless symptoms continue for 3 months after their onset.
2. Among 180 patients with new inflammatory arthritis followed up to 3 months, the arthritis went away after symptomatic therapy in 120 patients.
3. It was interesting to note that there were no patients with RA among these 120 patients responding to symptomatic therapy.

This is clearly circular logic. It was no surprise that the rheumatologists in this setting did not see any patients with RA. The rule in their clinic was that 3 months had to pass after the initial symptoms for a patient to have RA. However, none of the 120 patients was followed for longer than 3 months.

Now let us consider the following three statements:

1. Our unit defines RA as symmetrical polyarthritis of unknown cause that lasts at least 3 months and involves at least two-thirds of joint groups made of metacarpophalangeal (MCPs) joints, wrists, or metatarsophalangeal (MTPs) joints.
2. Among 1,000 new patients seen in our clinic within a year, 490 satisfied this definition.
3. Almost half of the new patients we see in a year have RA.

There is nothing circular in the above statements. First, a rheumatology unit defined a condition. Then, they observed a number of patients fulfilling that condition within a certain time interval. They finally concluded that they saw the X-number of patients who fulfilled their definition within the time period specified.

In brief:
- Coming to a conclusion unaware that the conclusion reached was inescapable is “circular logic.”
- To search for and find what has been defined is NOT circular logic.

The second thought barrier as listed in Table 1 is the “promise of a diagnostic criteria as distinct from classification criteria.” I had addressed this issue in some detail in my previous paper. Since then and predictably, the new ACR/EULAR criteria also make the point that separate criteria are needed for diagnosis. Again, appealing to Dr. Fries:5

Conceptually, the classification criteria are the same as diagnostic criteria, and in a perfect world, might indeed be named diagnostic criteria. That is, if sensitivity and specificity were both 100%, classification criteria would be diagnostic criteria and would apply to every individual case.

I would suggest that not conceptually but in reality diagnosis and classification are the same; however, it should be added that neither can be universal, in other words “good for all purposes.” This, in turn, takes us to the third and fourth points in Table 1.

In my previous paper on classification,1 I attempted to explain that what really transforms a set of classification criteria to that of diagnostic criteria is mainly two-fold: 1. Why do we want to diagnose or classify? and 2. What is the
pre-test probability of the conditions we wish to classify or diagnose? These are also exactly the third and fourth issues in the table. As explained in the previous paper, and as I will further discuss when I go back to the ACR/EULAR RA criteria shortly, there can be many reasons for making criteria for classification or diagnosis. We can make such criteria for disease prevention, management, and allocation of resources, as well as for clinical or laboratory research. While we do this, the pretest odds will differ according to geography, subspecialty, disease course, and the curiously uniformly forgotten issue of yet undefined diseases. Any criteria making, by definition, needs suitable comparison or control groups. How does one include a yet undefined disease in a control group? This is one important reason why a specificity of 100% is most rare in any biological research.

The last point on Table 1 will, I think is better appreciated in the context of how the remaining points are relevant to the new ACR/EULAR criteria for RA. The issues of the ungrounded fear of circularity and the promise of diagnostic criteria, as distinct from classification criteria, yet to come, are explicitly present in the related manuscripts. However, my most important criticism of the said criteria is the fact that they are not criteria for diagnosing or RA. They are, in fact, criteria for the question “When and if to start methotrexate in a patient with early onset inflammatory arthritis?” The whole design of this criteria exercise was to this end. This being so, I venture to suggest that it was the barrier no. 4 in Table I that was the culprit.

Disclosure Statement
This work has been partially sponsored by the Turkish Academy of Sciences. The author has no further financial or proprietary interest in the subject matter or materials discussed, including, but not limited to employment, consultancies, stock ownership, honoraria and paid expert testimony.

References

Table 1 Thought Barriers in Criteria Making

| 1. The misconception of trying to avoid circularity in criteria making |
| 2. Promise of diagnostic criteria, somehow different from classification criteria, to come |
| 3. A lack of consideration of why and how should we diagnose |
| 4. A lack of appreciation of pre- and posttest (criteria) odds |
| 5. A lack of consideration that we should perhaps involve our patients in naming and what we plan to do with their diagnosis(es) |