Editorial

Are Disease Classification Criteria for Diagnosis or for Research? In Fact, for Neither

Hasan Yazici1 and Yusuf Yazici2

A popular viewpoint is that disease classification criteria, usually developed for conditions with uncertain disease mechanisms, should not be used in diagnosing patients. They are for research purposes and aim to homogenize study patient characteristics to better interpret research outcomes from diverse research centers.1 We believe what made this viewpoint so popular is that it is deceptively convincing. We agree that these criteria should not be used in diagnosing a patient in clinical care in that they should not be the only determinants of our clinical diagnoses. Instead, we say they can be used as guidelines. On the other hand, we strongly disagree with the notion that a set of similarly universal disease classification criteria for research (DCCR) in a specific disease can or should be formulated. It is the hypothesis of a particular study, rather than the DCCR, that determines who needs to be studied. We also propose that the use of the word *criteria* causes an unjustified impression of certainty about facts that are inherently much more probabilistic and therefore should be replaced by the word *guidelines*.

### Disease classification criteria for diagnosis

Although the advice is not to use these criteria for diagnoses, in fact, they are used even in the face of data showing the validity of these criteria among the already diagnosed patients is rather unsatisfactory. Notable examples are related to the 1990 American College of Rheumatology (ACR) Vasculitis Criteria,2,3 the 2010 ACR/European Alliance of Associations for Rheumatology (EULAR) rheumatoid arthritis (RA) disease classification criteria,4 a previously endorsed ACR criteria set for systemic lupus erythematosus (SLE),5,6 and Assessment of Spondyloarthritis international Society (ASAS) criteria for spondyloarthritis (SpA).9,10 There are indications8,9 that the same will be true for the 2022 ACR/EULAR criteria for eosinophilic granulomatosis with polyangiitis (EGPA)11,12 and granulomatosis with polyangiitis (GPA).13,14 The authors of these criteria sets uniformly emphasize that these criteria had not been developed for making diagnoses in clinical care. The appeal to authority follows: “Conceptually, classification criteria are the same as diagnostic criteria and, in a perfect world, might indeed be termed diagnostic criteria. That is, if sensitivity and specificity were both 100%, criteria would be diagnostic criteria, and would apply to every individual case.”15 We agree. Indeed, based on many factors, including disease prevalence, disease expression, as well as the type of medical practice,16,17 such criteria would be unsuitable for rendering diagnoses in clinical practice.

The Table shows some examples of how accurate such universal disease classification criteria turned out to be when assessed in different clinical settings. It is interesting to note that out of the 6 examples given (Table), 2 of them6,10 are assessments by the main authors of these classification schemes.

We agree that the use of disease classification criteria for diagnosis (DCCD) in patient care would have serious consequences1 of wrongly treating or not treating some patients. Such mismanagement can surely lead to lawsuits and/or problematic issues with third-party payors. However, we consider that an essential component of a diagnosis has totally been neglected in the discussion around the usefulness of disease criteria in patient care. In medical practice, our main concern is how to best assist the patient who seeks our help. The name of the disease or the pathological aberration behind the illness are only secondary. Our primary and privileged concern as physicians is the “why” of a diagnosis. This “why” overrides, or should override, all other considerations we have just discussed in rendering a diagnosis in clinical care, including the name of or the pathological aberration in an illness. Here, we also want to appeal to the same authority, Dr. Fries. At the end of the same paragraph to which our disease classification criteria colleagues frequently quote, he says that “the physician, considering features of an individual patient beyond those represented in the criteria, is the only one who can..."
establish a diagnosis for an individual patient.”¹⁵ In other words, the buck stops at the physician.

**DCCR**

Whereas the “why” of a diagnosis is what to do with it when we diagnose a patient, the “why” of research is to test a hypothesis. Thus, the main determinant of patient enrollment into a study is the very hypothesis of that specific study, and the forms to formulate universal disease classification criteria are not well suited for this purpose, either. Different research questions frequently mandate studying different patient groups, and in the case of interventional studies, such as drug trials, the nature of the planned intervention has full bearing on patient enrollment, where safety considerations become all-important.

For example, we would consider it ethically justifiable to do a controlled trial of colchicine for the control of arthritis among the relatives of patients with Behçet syndrome with only oral ulcers and symmetrical oligoarthritis of, say, a month’s duration. On the other hand, we would not at all consider it justifiable to conduct a trial of tumor necrosis factor agents in this group, who certainly would not fit any classification criteria for Behçet syndrome. The example we give underlines another unwanted aspect of DCCR—that of virtually precluding clinical research in incomplete forms of our diseases.

A rationalization for the concept of DCCR is that such criteria are to be applied to already diagnosed patients.¹⁸ This argument, however, does not in any way address the issue of the hypothesis-driven dependency of inclusion criteria for different research projects. We maintain that every research project, unless it aims at strict replication, needs a unique patient selection criteria, and not a generalized criteria set.

There is a further, and mainly an ethical, issue. It has been suggested that incorrect classification in enrolling patients into a study using DCCR does not have the unwanted consequences of an incorrect diagnosis in clinical care.¹⁷ We disagree. As we already discussed, some disease classification criteria have low specificities for targeting the disease we plan to study, such as the 2010 RA criteria. Applying these criteria might cause many patients to be erroneously enrolled into studies, leading to misleading biological data, and in the case of interventional studies, administering wrong treatments to, or withholding the necessary treatments from, some patients. The litmus test to underline the importance of the ethical aspect of using universal classification criteria in interventional studies lies in our honest answer to the question of whether we are ready to acknowledge the specificity of these criteria in the informed consent forms of these studies. It is important to remember that a patient is primarily enrolled in a trial not for their own benefit but for societal good. For this reason, we must be even more respectful to the “first do no harm” principle in enrolling a patient into a trial than in routine care. In fact, it might be said that a 100% sensitive and 100% specific disease criteria set, if ever reached, would be more fitting for research than for a diagnosis in medical care.

An additional and practical consideration is that DCCR are usually prepared in academic centers, where more severe forms of a disease are taken care of. It follows that, similar to the problems with preparing DCCD, as we discussed, preparing research criteria based on this skewed population might result in many patients who have less severe forms of the disease ending up not having easy access to new and more effective remedies.¹⁹

We are of the opinion that, historically, at least in rheumatology, where diseases of unknown etiology are many, the word criteria, an unrealistic and unnecessarily strong word, first began to be used increasingly to describe our mere classification schemes for various diseases. This increase was augmented by our increasing abilities of data processing. Moreover, there has been our ever-strengthening perceived conception that contemporary managed care and healthcare reimbursement schemes expected us to be more and more specific—hence, less probabilistic—in the practice of our trade. Instead of realistically announcing and advising, starting first with our patients and then the public, of the still many imprecise diagnoses we have to make in clinical practice, we decided to comply with what we thought was expected of us. In addition, when we realized that the classification criteria sets we prepared were not as helpful as we expected them to be in clinical care, in order not to throw out the baby with

---

**Table. Diagnostic performance of the universal classification criteria among patients with RA, SLE, SpA, EGPA, and GPA, from 6 selected publications.**

<table>
<thead>
<tr>
<th>Disease Criteria</th>
<th>Highlights</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA⁴</td>
<td>Viewpoint article² analyzing the application of the 2010 RA criteria on 7 datasets of 3862 already diagnosed patients with RA. A systematic review⁵ reporting 17 full papers (6816 patients) and 17 meeting abstracts (4004 patients) analyzes the application of the 2010 RA criteria to already diagnosed patients. It is estimated that about 18% of the patients fulfilling the assessed criteria would, in time, have other conditions.⁶</td>
<td>74</td>
<td>71</td>
</tr>
<tr>
<td>SLE⁷</td>
<td>A cohort of 628 patients with SLE all followed at a single dedicated SLE center with 51% with pure and 49% with variant forms of SLE. Assessment was made using ACR-endorsed criteria available in 2017.⁵</td>
<td>83 (pure SLE); 33 (SLE variants)</td>
<td>NR</td>
</tr>
<tr>
<td>SpA⁷</td>
<td>A systematic review and metaanalysis of 9 publications, with a total of 5739 patients.¹⁰</td>
<td>73</td>
<td>88</td>
</tr>
<tr>
<td>EGPA¹¹</td>
<td>A cohort of 51 patients with EGPA in Korea.¹²</td>
<td>84</td>
<td>NR</td>
</tr>
<tr>
<td>GPA¹³</td>
<td>A cohort of 65 patients with GPA in Korea.¹⁴</td>
<td>74</td>
<td>NR</td>
</tr>
</tbody>
</table>

EGPA: eosinophilic granulomatosis with polyangiitis; GPA: granulomatosis with polyangiitis; NR: not reported; RA: rheumatoid arthritis; SLE: systemic lupus erythematosus; SpA: spondyloarthritis.
the bathwater,\textsuperscript{19} we decided to call them DCCR, almost adding insult to injury. On the other hand, many concepts that would be very useful both in clinical care and research have been brought up while developing these criteria sets. We can simply read the word criteria as guidelines in all those published disease criteria sets and use them—informally, diligently, and cautiously—for either diagnosing patients we take care of or planning research. The only 2 places where the use of the word criteria would be fully justified are (1) when we describe, rather than dictate, the patient enrollment in research reports, and (2) for planning or reimbursing health care, where the doctors’ diagnoses would be used as frequency criteria and not as a basis for individual reimbursement. Another point to remember is that new disease classification guidelines should surely continue to be data driven, and care should be taken to choose our validation groups independent of derivation groups, a rule unfortunately still not much respected.\textsuperscript{20}

Finally, we envisage that there will be 2 additional and important benefits of our proposed approach to assigning names to diseases and/or conditions of unknown disease mechanisms in patient care or in research. First, it will remind our students that a unique and privileged aspect of their chosen profession is to reconcile general probabilities with individual situations both in clinical medicine and in research. Second, it will hopefully remind or teach all the stakeholders of both medical care and scientific research that respect for the professional autonomy of the physician or the scientist is too precious to give up for the sake of the currently desired smooth running of managed healthcare and managed science.

ACKNOWLEDGMENT

We are grateful to Prof. M. Boers for providing a critical review of a draft of our manuscript.

REFERENCES