Classification Criteria and Diagnostic Tests for Vasculitides

Most rheumatic diseases, in adults or children, do not have objective clinical or laboratory measures that allow physicians to diagnose a specific disease with certainty. In order to overcome the clinical dilemma of a proper diagnosis, classification criteria have been proposed for several rheumatic diseases. Classification criteria are therefore an essential tool for clinical research since they allow comparison of patients with similar clinical and laboratory characteristics across studies.

METHODOLOGICAL APPROACHES TO CLASSIFICATION CRITERIA

In the vasculitis research field several classification criteria have been proposed for the adult population, and more recently for children. These criteria have been derived through 2 major methodological approaches: consensus and consensus/data-driven.

In brief, the consensus-based criteria are specifically derived through the combination of judgments from a group of experts in a particular field after thorough literature review and sometimes developed using consensus techniques such as the Delphi technique and nominal group technique[1,2]. To the consensus-based criteria we can ascribe the Chapel Hill International Consensus Conference[3] that essentially provided proper nomenclature for systemic vasculitides, and the Childhood Vasculitis Working Group of the Paediatric Rheumatology European Society (PRES) for preliminary classification criteria of childhood vasculitis[4]. These criteria have their main strength in the panel of contributors who are usually very well recognized experts/experienced clinicians in that particular research/clinical field.

The second approach of consensus/data-driven usually starts from consensus-based criteria and adds a formal statistical validation based on collected data in order to provide accuracy measures (classic methods), or other alternative methods such as the classification tree[5]. Only the classical approach is discussed here.

With the classic approach, the accuracy measures that are considered include sensitivity (ability of a classification criterion to identify a patient as having the disease based on a gold standard; also called true-positive rate); specificity (ability of classification criteria to exclude that a patient has the disease; also called true-negative rate); positive predictive value (number of patients correctly classified with the disease by classification criteria, divided by all positive patients); negative predictive value (number of subjects correctly classified without the disease by classification criteria, divided by all negative subjects); positive likelihood ratio (ratio between sensitivity and 1 minus specificity); negative likelihood ratio (ratio between $1 -$ specificity divided by sensitivity); and diagnostic odds ratio (ratio between positive and negative likelihood ratio)[6]. In addition the kappa statistic[7] can be used to measure strength of agreement between proposed classification criteria and a related gold or reference standard (see below), usually with the following cutoffs proposed by Landis and Koch[8]: 0.01–0.2 = slight; 0.21–0.4 = fair; 0.41–0.6 = moderate; 0.61–0.8 = substantial; 0.81–1.0 = almost perfect agreement. Where more than 1 set of criteria is tested, only the criteria with kappa statistic > 0.7 (substantial agreement), sensitivity and specificity > 80%, and false-positive and false-negative < 20% are usually retained for face or content validity, i.e., the evaluation of which a set of criteria (if more than one is tested) is easiest to use and most credible.

Ideally the best set of criteria should have the highest sensitivity and specificity; however, it has been suggested that for epidemiologic studies one may wish to maximize sensitivity so as to have all cases in the study, while for drug trials specificity is probably more important to properly exclude patients without the disease[5]. Regarding consensus/data-driven criteria, we include the American College of Rheumatology (ACR) endeavor for vasculitides in the 1990s (both classic and alternative methods were used)[9,10,11,12], the classification algorithm endorsed by the

European Medicines Agency (EMA, formerly EMEA)[13,14,15], which consider in a 4-step approach the Churg-Strauss criteria of the ACR and Lahman, et al[16], the ACR criteria for granulomatosis with polyangiitis (GPA, formerly Wegener's granulomatosis), the Chapel Hill Consensus nomenclature for GPA, the Chapel Hill histological definition for microscopic polyangiitis, and lastly, the Chapel Hill panarteritis definition. The EMA algorithm essentially proposes a sequential method in order to identify diseases according to published criteria. More recently, criteria for childhood vasculitides were proposed and validated by the European League Against Rheumatism (EULAR)/ Paediatric Rheumatology International Trials Organisation (PRINTO)/PRES[17,18,19].

GOLD STANDARD FOR CLASSIFICATION CRITERIA VERSUS PHYSICIAN DIAGNOSIS IN CLINICAL PRACTICE

Vasculitides are complicated conditions that are difficult to properly diagnose because they lack a pathognomonic test(s) to objectively identify them (diagnostic criteria). Several classification criteria have been proposed for research purposes to provide a common terminology allowing comparison of different studies. Because of the dichotomy between diagnostic and classification criteria, classification criteria are often used in clinical practice in place of diagnostic criteria. Whenever a pathognomonic test is available (e.g., enzyme determination for metabolic diseases), there is by definition no need to establish classification criteria.

A major challenge in establishing a new "test" (in this case a set of criteria) is to have a "gold standard" to evaluate the accuracy of measures. The gold standard for a new laboratory test could be the old test(s) previously used for that particular determination, for example, comparison of new versus older equipment used in laboratory evaluation. In rheumatology a lack of pathognomonic test(s) has led to the standard practice of using as the "gold standard" (or better, as a reference standard) the consensus panels composed of people with different backgrounds, including researchers and experienced clinicians. The consensus panel has the goal to evaluate, usually with the help of consensus formation techniques, if a group of patients has clinical signs/symptoms, laboratory or imaging data compatible with the disease under evaluation. This is normally done by presenting a comprehensive summary of data from a group of real patients suspected to have the index disease (e.g., GPA) and one or more groups of patients with confounding diseases. To ensure unbiased disease attribution based on available data, the consensus panel is normally blinded to the original disease attribution by the attending physician[18,19]. The disease attribution by the consensus panel is then used as the reference standard against which the accuracy measures of proposed ad hoc-derived classification criteria are evaluated. The accuracy measures should help to identify the classification criteria that best reflect the consensus panel evaluation (high sensitivity and specificity of the proposed classification criteria vs the consensus panel). When several classification criteria have similar accuracy measures, the consensus panel is also required to evaluate face validity. This approach, only briefly described here, has the advantage of combining evaluation of real data with the consensus of a panel of experts.

Both the ACR[9,10,11,12] and EULAR/PRINTO/ PRES[17,18,19] classification criteria for adult and childhood vasculitides are based on a consensus panel of experts as reference standard and the most common forms of vasculitis as confounding conditions in order to propose classification criteria.

CURRENT EVIDENCE

In this issue of The Journal Uribe, et al evaluate the accuracy of measures for classification of GPA (formerly Wegener's granulomatosis)[21] of the ACR criteria, EULAR/ PRINTO/PRES criteria, and EMA algorithm in a large series of pediatric patients with antineutrophil cytoplasmic antibody (ANCA)-associated vasculitides of the ARChiVe registry (A Registry for Childhood Vasculitis) set up by the Childhood Arthritis and Rheumatology Research Alliance. The authors report that the EMA algorithm has higher sensitivity (but lower specificity) than the EULAR/PRINTO/ PRES and ACR criteria.

The authors should be lauded for their initiative and for such a large data collection of pediatric cases of ANCA-associated vasculitis; however, it is still difficult to discern the correct place of each set of criteria because of several differences in the methodological approach used by the authors with respect to the published criteria. For example, the kappa level of agreement between EMA algorithm, ACR criteria, and EULAR/PRINTO/PRES classification criteria and physician diagnosis was fair to moderate (kappa 0.34–0.49), while the agreement between the criteria and the consensus panel evaluation for GPA was almost perfect in the original work from EULAR/PRINTO/PRES (kappa = 0.9)[18,19] and in the EMA algorithm (kappa = 0.886)[13].

While physician diagnosis is often a subjective intuitive process based on available clinical and laboratory data and individual expertise, classification criteria have to rely on clinical/laboratory/imaging features that could be applied to all patients in order to discriminate one disease from the other. So the choice to use physician diagnosis as a reference standard is probably the main reason for the low level of agreement reported by the authors. Indeed, this choice is a reflection of the well known difficulty in daily clinical practice of properly diagnosing different vasculitides. If research could be based on diagnosis of the treating physician when no pathognomonic criteria are available, there would be no need to work on classification criteria.

The other problem is that while the EMA algorithm pro-

poses a method to separate MPA, GPA, and polyarteritis nodosa, the other 2 criteria sets do not provide this distinction. Indeed, the ACR criteria did not cover MPA, which were later proposed by the Chapel Hill nomenclature consensus conference[3], while the childhood effort explicitly excluded MPA because just a few cases were available in the dataset.

With these limitations, the fact remains that at the moment none of the 3 proposed criteria sets is able to properly differentiate GPA from MPA, at least when the treating physician diagnosis is used as reference.

In conclusion, classification criteria have been identified as the best scientific method for providing reproducible definitions of disease for which diagnostic tests are not yet available. The current available classification criteria provide accurate definitions for identification of GPA in both adults and children, while more work is still needed for the proper identification of patients with MPA.

NICOLINO RUPERTO, MD, MPH,
for the Paediatric Rheumatology International Trials Organisation (PRINTO),
Istituto Giannina Gaslini,
Pediatria II — Reumatologia,
PRINTO, EULAR Centre of Excellence in Rheumatology 2008-2013,
Via Gaslini 5,
16147 Genova, Italy

Address correspondence to Dr. Ruperto;
E-mail: nicolaruperto@ospedale-gaslini.ge.it

REFERENCES

1. Delbecq AL, Van de Ven AH, Gustafson DH. Group techniques for program planning. A guide to nominal group and Delphi processes. Glenview, IL: Scott, Foresman and Company; 1975.
2. Ruperto N, Meiorin S, Iusan SM, Ravelli A, Pistorio A, for the Paediatric Rheumatology International Trials Organisation (PRINTO). Consensus procedures and their role in pediatric rheumatology. Curr Rheumatol Rep 2008;10:142-6.
3. Jennette JC, Falk RJ, Andrassy K, Bacon PA, Churg J, Gross WL, et al. Nomenclature of systemic vasculitides: Proposal of an international consensus conference. Arthritis Rheum 1994; 37:187-92.
4. Ozen S, Ruperto N, Dillon MJ, Bagga A, Barron K, Davin JC, et al. EULAR/PReS endorsed consensus criteria for the classification of childhood vasculitides. Ann Rheum Dis 2006;65:936-41.
5. Bloch DA, Michel BA, Hunder GG, McShane DJ, Arend WP, Calabrese LH, et al. The American College of Rheumatology 1990 criteria for the classification of vasculitis. Patients and methods. Arthritis Rheum 1990;33:1068-73.
6. Sox HC Jr. Probability theory in the use of diagnostic tests. An introduction to critical study of the literature. Ann Intern Med 1986;104:60-6.
7. Cohen J. Statistical power analysis for the behavioral sciences. New York: Academic Press; 1977.
8. Landis JR, Koch GC. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
9. Mills JA, Michel BA, Bloch DA, Calabrese LH, Hunder GG, Arend WP, et al. The American College of Rheumatology 1990 criteria for the classification of Henoch-Schonlein purpura. Arthritis Rheum 1990;33:1114-21.
10. Arend WP, Michel BA, Bloch DA, Hunder GG, Calabrese LH, Edworthy SM, et al. The American College of Rheumatology 1990 criteria for the classification of Takayasu arteritis. Arthritis Rheum 1990;33:1129-34.
11. Lightfoot RW Jr, Michel BA, Bloch DA, Hunder GG, Zvaifler NJ, McShane DJ, et al. The American College of Rheumatology 1990 criteria for the classification of polyarteritis nodosa. Arthritis Rheum 1990;33:1088-93.
12. Leavitt RY, Fauci AS, Bloch DA, Michel BA, Hunder GG, Arend WP, et al. The American College of Rheumatology 1990 criteria for the classification of Wegener's granulomatosis. Arthritis Rheum 1990;33:1101-7.
13. Watts R, Lane S, Hanslik T, Hauser T, Hellmich B, Koldingsnes W, et al. Development and validation of a consensus methodology for the classification of the ANCA-associated vasculitides and polyarteritis nodosa for epidemiological studies. Ann Rheum Dis 2007;66:222-7.
14. Mohammad AJ, Jacobsson LT, Mahr AD, Sturfelt G, Segelmark M. Prevalence of Wegener's granulomatosis, microscopic polyangiitis, polyarteritis nodosa and Churg-Strauss syndrome within a defined population in southern Sweden. Rheumatology 2007;46:1329-37.
15. Liu LJ, Chen M, Yu F, Zhao MH, Wang HY. Evaluation of a new algorithm in classification of systemic vasculitis. Rheumatology 2008;47:708-12.
16. Lanham JG, Elkon KB, Pusey CD, Hughes GR. Systemic vasculitis with asthma and eosinophilia: A clinical approach to the Churg-Strauss syndrome. Medicine 1984;63:65-81.
17. Ruperto N, Martini A. International research networks in pediatric rheumatology: The PRINTO perspective. Curr Opin Rheumatol 2004;16:566-70.
18. Ruperto N, Ozen S, Pistorio A, Dolezalova P, Brogan P, Cabral DA, et al. EULAR/PRINTO/PRES criteria for Henoch-Schonlein purpura, childhood polyarteritis nodosa, childhood Wegener granulomatosis and childhood Takayasu arteritis: Ankara 2008. Part I: Overall methodology and clinical characterisation. Ann Rheum Dis 2010;69:790-7.
19. Ozen S, Pistorio A, Iusan SM, Bakkaloglu A, Herlin T, Brik R, et al. EULAR/PRINTO/PRES criteria for Henoch-Schonlein purpura, childhood polyarteritis nodosa, childhood Wegener granulomatosis and childhood Takayasu arteritis: Ankara 2008. Part II: Final classification criteria. Ann Rheum Dis 2010;69:798-806.
20. Falk RJ, Gross WL, Guillevin L, Hoffman GS, Jayne DR, Jennette JC, et al. Granulomatosis with polyangiitis (Wegener's): An alternative name for Wegener's granulomatosis. Arthritis Rheum 2011;63:863-4.
21. Uribe AG, Huber AM, Kim S, O'Neil KM, Wahezi DM, Abramson L. Increased sensitivity of the European Medicines Agency algorithm for classification of childhood granulomatosis with polyangiitis. J Rheumatol 2012;39:1687-97.