Classification and diagnostic criteria in Sjögren's syndrome: a long-standing and still open controversy

Claudio Vitali,^{1,2} Nicoletta Del Papa^{2,3}

Most rheumatic diseases are multisystem disorders that are heterogeneous in their presentation, course and outcome. These conditions still lack a single clinical, laboratory, pathological or radiological feature that could serve as a 'gold standard' in support of diagnosis and/or classification. Thus, the development of criteria for use in clinical care and research studies has been an important challenge in these disorders.¹

From the theoretical and methodological point of view, classification and diagnostic criteria are quite different. Classification criteria are standardised tools that are aimed at selecting well-defined and homogenous groups of patients for research and at guaranteeing comparability across studies. They are not designed to be used for the clinical diagnosis in individual patients and may be defective in capturing some cases with a less common clinical presentation or course.²

Diagnostic criteria are generally less stringent and usually include a wider variety of disease features. Their aim is to accurately identify as many people with that condition as possible.¹ Given the complexity of systemic rheumatic disorders, the development of diagnostic criteria in these diseases is certainly difficult. Therefore, optimal diagnostic criteria have not been defined for most of the rheumatic diseases and the diagnosis, given the suspicion of one of these disorders, is commonly based on a decision-making process by physicians who have to evaluate a complex combination of symptoms, signs, diagnostic tests and rule out other confounding or similar diseases.

As a consequence of these theoretical assumptions, the diagnostic criteria are commonly characterised by high sensitivity

and negative predictive value, whereas the classification criteria classically possess high specificity and positive predictive value to minimise the risk of classifying false positive patients as having the disease. Sensitivity and specificity show an inverse relationship where to any increase of the former corresponds a decrease of the latter and vice versa.³ The receiver operating characteristic (ROC) curve is the statistical and graphical description of this process, showing the equilibrium between sensitivity and specificity.⁴

Because of the lack of diagnostic criteria for many rheumatic disorders, no studies on direct comparison between classification and diagnostic criteria for the same disease are traceable in the medical literature.⁵ Conversely, the performance of classification criteria as a diagnostic tool has been explored in a number of studies where the expert clinician's judgement was considered the gold standard for the diagnosis. As expected, specific classification criteria did not demonstrate to be a reliable instrument in making a correct diagnosis in the different disorders.⁵ In spite of their deceptive diagnostic performance in individual patients, the use of classification criteria as a diagnostic tool is commonplace in daily rheumatological practice. Classification criteria are regarded as a useful guide for diagnosis, and, in addition, they may have a role in education and training in medicine.⁵

A large number of studies comparing different classification criteria for the same disorder have been carried out, often aimed at measuring the performance of newly proposed criteria to that of the older ones. In this regard, Tsuboi *et al*⁶ report a study performed in a large cohort of Japanese (JPN) patients where the sensitivity and specificity of the new 2016 American College of Rheumatology (ACR)-European League Against Rheumatism (EULAR) classification criteria for primary Sjögren's syndrome (pSS)⁷ were compared with those of the 1999 revised JPN Ministry of Health diagnostic criteria,⁸ the 2002 American-European Consensus Group (AECG)⁹ and

the 2012 ACR¹⁰ classification criteria for this disease. On the whole, the results of this comparison indicate that the 2016 ACR-EULAR criteria have higher sensitivity and lower specificity in the classification of patients with pSS than the other three sets of criteria. Furthermore, the degree of agreement of the ACR-EULAR classification criteria with all the other three sets of criteria was low.

Looking in details at the results of this study,⁶ and namely at the subanalysis of 383 cases—that is certainly more reliable for the higher similarity of the considered diagnostic items across the different criteria sets—it is rather surprising to see that the JPN criteria are the ones with the highest specificity and the lower sensitivity. This result is rather unexpected since the JPN criteria are the only ones defined as diagnostic criteria among the criteria compared in the study.

Taking in mind the theoretical considerations discussed above on the critical differences between classification and diagnostic criteria, to compare JPN criteria, which were defined as diagnostic, to other classification criteria for pSS could be 'per se' an invalidating procedural defect. However, considering the general policy and procedures adopted in the development of the revised JPN diagnostic criteria, in which it was outlined that one of the goals should be to make only definite diagnoses and to exclude probable cases, in other words, to have a high specificity,⁸ one can conclude that the JPN criteria should have been more correctly defined as classification rather than diagnostic criteria.

Other factors may have conditioned the results of this study. It is well known that both the classification and diagnostic criteria performance may vary in different clinical and geographical settings.^{11 12} This may greatly depend on the prevalence and clinical pattern of presentation that a disease may have in different geographical regions and in different clinical backgrounds.^{12 13} Thus, it is likely that the best performance of any criteria set can be reached in the clinical setting and geographical area where the criteria set has been developed. This performance variability is expected to be wider for diagnostic criteria that include more disease descriptors, but may also be observed, to a lesser extent, in applying classification criteria.

The low level of agreement between the ACR/EULAR and AECG criteria observed in the study by Tsuboi *et al* is in contrast with what was reported before.⁷ This discrepancy can be largely reduced, and the agreement between the ACR/EULAR

¹Rheumatology section, Istituto Santo Stefano, Villa San Giuseppe, Como, Italy

²Study Group on Sjögren's Syndrome, Gaetano Pini Hospital, Milan, Italy

³Day Hospital of Rheumatology, ASST Gaetano Pini-CTO, Milan, Italy

Correspondence to Dr Nicoletta Del Papa, Day Hospital of Rheumatology, ASST Gaetano Pini-CTO, Piazza Cardinal Ferrari 1, 20122 Milan, Italy; nicoletta.delpapa@asst-pini-cto.it

and AECG criteria consistently improved, reconsidering the 19 patients of this cohort who were classified as having pSS by only the ACR/EULAR criteria.⁶ They have positive lip biopsy (11 patients) or positive anti-SSA/Ro antibodies (8 patients), plus reduced salivary (18 patients) or lachrymal flow (1 patient). Most of these patients could also have met the AECG criteria if the presence of dry eye and dry mouth symptoms had been investigated by the AECG-validated questionnaires for sicca symptoms. The authors did not specify the way they explored sicca complaints in their retrospective study.

The fact that, in the study of Tsouboi *et al*, the ACR/EULAR classification criteria for pSS have demonstrated higher sensitivity and, consequently, lower specificity than all of the other criteria sets is not completely unexpected. The appearance of new therapeutic agents with a favourable risk-benefit profile and the potential to change the long-term prognosis of rheumatic disorders has outlined the need to define new classification criteria with a higher sensitivity and therefore able to recognise patients with early disease. With the support of and following the methodological procedures approved by both the ACR and EULAR 'ad hoc' committees,^{4,14} newer classification criteria for different disorders have been proposed and validated in multicentre multinational frameworks.^{7,15–18} Of course, a loss of specificity may be the counterpart to the increased sensitivity of the new criteria. Consequently, more 'liberal' criteria should be used with caution when a therapeutic agent with an unclear safety profile is under investigation in a trial. By moving along the ROC curve designed for the new classification criteria, and then applying the criteria in a flexible way, one can find a different sensitivity/specificity ratio capable of greatly reducing the risk of selecting and treating false positive cases. A cut-off point of 5 instead of 4, for instance, raises the specificity of the ACR-EULAR classification criteria for pSS from 89% to 98%.⁷

The new ACR-EULAR classification criteria for pSS are the final result of an international cross-cultural collaboration and are derived by a well-established and validated methodology. At the best of present knowledge, these criteria describe the key shared features defining this condition and may represent the common language to be used in the next future to make the scientific communication easier, favour the exchange of information and stimulate the development of collaborative studies.

Contributors CV is the main author. NDP contributed to writing the paper and revising the cited literature.

Competing interests None declared.

Provenance and peer review Commissioned; externally peer reviewed.

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

CrossMark

To cite Vitali C, Del Papa N. *Ann Rheum Dis* 2017;**76**:1953–1954.

► <http://dx.doi.org/10.1136/annrheumdis-2016-210758>

Ann Rheum Dis 2017;**76**:1953–1954.
doi:10.1136/annrheumdis-2017-211378

REFERENCES

- 1 Fries JF, Hochberg MC, Medsger TA, *et al*. Criteria for rheumatic disease. Different types and different functions. The American College of Rheumatology Diagnostic and Therapeutic Criteria Committee. *Arthritis Rheum* 1994;**37**:454–62.
- 2 Johnson SR, Goek ON, Singh-Grewal D, *et al*. Classification criteria in rheumatic diseases: a review of methodologic properties. *Arthritis Rheum* 2007;**57**:1119–33.
- 3 Yazici H. Diagnostic versus classification criteria - a continuum. *Bull NYU Hosp Jt Dis* 2009;**67**:206–8.
- 4 Dougados M, Gossec L. Classification criteria for rheumatic diseases: why and how? *Arthritis Rheum* 2007;**57**:1112–5.

- 5 June RR, Aggarwal R. The use and abuse of diagnostic classification criteria. *Best Pract Res Clin Rheumatol* 2014;**28**:921–34.
- 6 Tsuboi H, Hagiwara S, Asashima H, *et al*. Comparison of performance of the 2016 ACR-EULAR classification criteria for primary Sjögren's syndrome with other sets of criteria in Japanese patients. *Ann Rheum Dis* 2017;**76**:1980–5.
- 7 Shiboski CH, Shiboski SC, Seror R, *et al*. American College of Rheumatology/European League Against Rheumatism classification criteria for primary Sjögren's syndrome: A consensus and data-driven methodology involving three international patient cohorts. *Ann Rheum Dis* 2017;**76**:9–16.
- 8 Fujibayashi T, Sugai S, Miyasaka N, *et al*. Revised Japanese criteria for Sjögren's syndrome (1999): availability and validity. *Mod Rheumatol* 2004;**14**:425–34.
- 9 Vitali C, Bombardieri S, Jonsson R, *et al*. Classification criteria for Sjögren's syndrome: a revised version of the European criteria proposed by the American-European Consensus Group. *Ann Rheum Dis* 2002;**61**:554–8.
- 10 Shiboski SC, Shiboski CH, Criswell L, *et al*. American College of Rheumatology classification criteria for Sjögren's syndrome: a data-driven, expert consensus approach in the Sjögren's International Collaborative Clinical Alliance cohort. *Arthritis Care Res* 2012;**64**:475–87.
- 11 Johnson SR, Franssen J, Khanna D, *et al*. Validation of potential classification criteria for systemic sclerosis. *Arthritis Care Res* 2012;**64**:358–67.
- 12 Gomariz EM, del M, Guijo VP, *et al*. The potential of ESSG spondyloarthritis classification criteria as a diagnostic aid in rheumatological practice. *J Rheumatol* 2002;**29**:326–30.
- 13 Murali NS, George R, John GT, *et al*. Problems of classification of Henoch Schonlein purpura: an Indian perspective. *Clin Exp Dermatol* 2002;**27**:260–3.
- 14 Singh JA, Solomon DH, Dougados M, *et al*. Classification and Response Criteria Subcommittee of the American College of Rheumatology Committee on Quality Measures. Development of classification and response criteria for rheumatic diseases. *Arthritis Rheum* 2006;**55**:348–52.
- 15 Aletaha D, Neogi T, Silman AJ, *et al*. Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010;**69**:1580–8.
- 16 Neogi T, Jansen T, Dalbeth N, *et al*. Gout classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2015;**74**:1789–98.
- 17 Dasgupta B, Cimmino MA, Maradit-Kremers H, *et al*. Provisional classification criteria for polymyalgia rheumatica: a European League Against Rheumatism/American College of Rheumatology collaborative initiative. *Ann Rheum Dis* 2012;**71**:484–92.
- 18 van den Hoogen F, Khanna D, Franssen J, *et al*. Classification criteria for systemic sclerosis: an American college of Rheumatology/European league against rheumatism collaborative initiative. *Ann Rheum Dis* 2013;**72**:1747–55.