Should We Keep Changing the Diagnostic Criteria for Behçet’s Disease?

Caroline Kronborg a  Patrick D. Mahar b  Robert Kelly c

a Department of Medicine, The Alfred Hospital, Alfred Health, and Departments of b Medicine and c Dermatology, St Vincent’s Hospital, University of Melbourne, Melbourne, Vic., Australia

Key Words
Behçet’s disease · Diagnostic criteria · International Criteria for Behçet’s Disease · International Study Group

Abstract
In medicine, clinical acumen is used to achieve diagnosis, guide management and prevent disease. While for some diseases, diagnosis is reached with the assistance of objective tests, many conditions rely upon the use of clinical diagnostic criteria; Behçet’s disease is one such case. In order to remain clinically relevant, as knowledge of a condition changes over time so too must its diagnostic criteria. Preferably, when new criteria for a disease are conceptualised it is through sound methodology, followed by a confirmation of accuracy by way of systematic validation and response to treatment. The most recently proposed revised International Criteria for Behçet’s Disease for the diagnosis of Behçet’s disease have been systematically validated and should replace the use of the clinically inferior International Study Group criteria, while not displacing the role of clinical judgement. Effort should now be invested in acquiring better understanding of the pathogenesis of the disease in the hope of developing a more objective test.

© 2013 S. Karger AG, Basel

Background

Behçet’s disease (BD) is a chronic, relapsing and remitting multisystemic disease of unknown aetiology [1]. With a relative paucity of cases and a variable natural course, research and our understanding of BD remains limited. There are no universally recognised pathognomonic tests or biomarkers reflective of BD activity, so diagnosis is clinical and based on features believed to be due to vasculitis of all types of blood vessels, most characteristically painful aphthous ulcers of the mouth and genitals, as well as ocular, cutaneous and neurologic manifestations. The pathergy phenomenon (formation of a sterile pustule at the site of a needle stick) is considered the most clinically relevant examination finding, however rates of reaction vary across different populations, and its sensitivity has declined over the years [2].

Histopathologic features of BD are said to be leukocytoclastic vasculitis, with neutrophilic infiltration and obliteration of the endothelial lumen, but these features are not always present [3]. Similarly, there is a higher rate of human leukocyte antigen HLA-B*5 or Bw51 amongst people with ancestry from the Mediterranean, the Middle East and the Far East and in BD patients, but again not universally [4].
Over the years, BD has given rise to a multitude of diagnostic criteria, perhaps more than any other eponymous disease, with more than 16 since first named. Yet despite an increasing understanding of BD overall, significant disparities between clinicians in terms of important concepts regarding clinical features, diagnosis and treatment still remain [5].

The International Study Group Criteria for Behçet’s Disease

In 1990, the International Study Group (ISG) criteria for BD were developed to operationalise the definition of the disease and facilitate research [6]. It distilled four clinical elements (recurrent oral aphthosis, genital aphthosis, ocular lesions and typical cutaneous lesions), considered oral aphthosis an obligatory manifestation and valued a positive pathergy test supportive of diagnosis, nominating it the 5th criterion (table 1). The ISG criteria remain the most widely used and well-accepted criteria among experts. While the study that led to their development should be commended for its size, collating data on 914 patients from 12 centres in 7 countries, the methodology applied for its development is not without flaws. Criticism cannot be levelled heavily against the authors however, as these biases were largely inherent to the research conducted and difficult to avoid.

The authors’ intention to develop internationally consistent and valid diagnostic criteria led to contamination and reduced objectivity of their data with the pre-existing clinical opinions of experts. Their primary cohort of patients labelled with ‘Behçet’s disease’ were diagnosed based on the decision of an experienced clinician according to the diagnostic criteria with which they were most familiar. Further bias was employed when the authors excluded 28 patients (3%) without recurrent oral ulcers because they believed that the presence of such ulcers was an imperative criterion for diagnosis, despite these patients having been allocated to the ‘diseased’ cohort by an experienced clinician earlier in the process and their own acknowledgement that BD can occur in the absence of such ulcers. They drew upon patients with oral ulcers from the same centres who were deemed to have connective tissue disease, which may show similar features to BD, to form a comparison group. A reportedly random 60% of the ‘Behçet’s’ cohort were nominated as the ‘training sample’ and data on their features of disease were weighted and assigned log-likelihood ratios, deriving the ISG diagnostic criteria. These criteria were then validated in the remaining 40% of the ‘Behçet’s’ cohort and compared to previously suggested diagnostic criteria, delivering a reported 95% sensitivity and 98% specificity of their original validation cohort. The ISG criteria became and have since remained the mainstay of diagnosis of BD. Yet the majority of subsequent validation studies have repeatedly found the ISG criteria to have a lower sensitivity compared to other proposed criteria and poor clinical utility for the diagnosis of individuals [7, 8].

The International Criteria for Behçet’s Disease

In 2006, in an effort to achieve improved clinical sensitivity in the diagnosis of BD, a new set of diagnostic criteria were proposed by an international team of Behçet’s experts: The International Criteria for Behçet’s Disease (ICBD) [9–11]. These criteria expanded the 5 items of the ISG criteria to 6, including vascular manifestations, and assigned a weighted point value system. Later in 2010, these criteria were themselves revised, neurologic signs were added and the pathergy test was considered an extra criterion to be used if conducted and positive, given its high specificity (table 2) [2].
In large validation studies comparing various diagnostic criteria for BD, the ICBD have consistently come out in front with improved sensitivity and have been proposed as a replacement for the ISG criteria for diagnosis and classification. In three independent cohorts of patients from China, Iran and Germany, the ICBD performed with greater sensitivity (up to 96.5%) with a good balance of specificity (97.3%), these metrics trading off against lower levels of accuracy (74.4–85.5%) [12].

In 2010, Davatchi et al. [13] evaluated the ICBD’s performance in Iran on a large cohort from their BD registry (6,128 patients), with 3,400 controls. They reported that the ICBD outperformed the ISG criteria on all fronts, with a higher sensitivity (98.2%), specificity (95.6%) and accuracy (97.3%), while comparatively the ISG criteria held low sensitivity (78.1%), high specificity (98.4%) but only 85.5% accuracy. Compared to the ISG criteria, the ICBD possessed better discriminatory power with a higher level of optimisation (difference between sensitivity and specificity 2.6 vs. 20.3%), with a reduced number of false negatives but without compromising on specificity.

The ICBD were further validated recently in a diagnostic accuracy study when the same prospectively collected international series of patients and controls on which the original ICBD were built was re-analysed; it involved 27 countries with 2,556 patients who were clinically diagnosed with BD and 1,163 controls with BD-mimicking disease or presenting with at least one major BD sign [8]. Similar to the methodology involved in the conception of the ISG criteria, the study leading to the development of the ICBD relied upon entrance into the ‘diseased’ arm to be clinically determined by an ‘expert in each relevant centre’. Patient characteristics in the cases and controls were again identified and scored through logistic regression, precipitating a scoring system. This was then compared against their own validation sample and existing criteria. If conducted, positive and assigned a point, the pathergy test improved the estimated sensitivity of the ICBD from 95.5 to 98.5% while only minimally reducing specificity from 92.1 to 91.6%.

The revised ICBD share some of the methodological problems with the ISG criteria, however on the whole appear the better tool for both diagnosis and classification of BD, where a greater sensitivity reduces the encounter-to-diagnosis lag time, fewer patients are missed and more are treated earlier.

### Diagnosis Based on Treatment

In clinical practice the goal of achieving diagnosis is to direct treatment so to reduce the severity of clinical manifestations and prevent complications. Historically, BD treatment regimens have been based on this objective [14]. Benefits of treatment can certainly be achieved in the absence of a formal diagnosis identifying a condition by name, and such practice is not infrequent in medicine. Indeed, a broad spectrum of relatively nonspecific anti-inflammatory and immunosuppressive agents, often in combination with each other or with colchicine, form the basis of BD treatment. Accordingly, perhaps a more useful approach would be broad, more sensitive and less specific diagnostic criteria that raise the suspicion of diagnosis and prompt earlier treatment. Better yet, an approach that identifies characteristics that may determine a patient’s response to a specific treatment and hence guide targeted therapy could be considered. Such an approach has not yet been taken, despite the considerable number of diagnostic criteria previously proposed.

### Diagnosis Based on Aetiology

Perhaps the failure to find a good objective test reflects the complexity of BD. Although many theories of aetio-pathogenesis have been proposed, the aetiology behind BD remains unknown [15]. Most theories centre on immune-mediated mechanisms in response to an infectious or environmental agent in genetically susceptible individuals [16]. HLA-B*5 has been implicated in genetic sus-
ceptibility, with more than 60% of patients with BD testing positive for HLA-B*51, however this circumstantial evidence does not denote causality [17]. In addition to genetic testing, research efforts should now be directed towards understanding the pathophysiology of BD, to confirm whether we are actually dealing with a single nosological entity or rather a syndrome, as has been suggested [18, 19]. This would hopefully precipitate the creation of a more objective and accurate diagnostic test with improved discriminatory powers, which could be used alone or in conjunction with clinical criteria and facilitate more targeted treatment.

Conclusion

While diligent efforts of many researchers and clinicians of the Behçet’s community have lead to numerous proposals of diagnostic criteria over the years, at times the scientific reporting of BD has occurred less frequently than the rate of proposals of diagnostic criteria. One of the critical reasons for developing objective diagnostic criteria is to establish consistency of diagnosis and thus an evidence base to strengthen the medical community’s understanding of a disease. Frequently altering diagnostic criteria for limited marginal objective gain are of limited value. The revised ICBD have now been systematically validated, demonstrating good discriminatory properties with improved sensitivity and clinical utility compared to previous criteria, regardless of country, and we agree with Davatchi et al. [8] that they should be used in place of the ISG criteria. The international Behçet’s community must now direct its efforts towards gaining improved understanding of BD aetiology so to develop an objective test for the disease.

In the meantime, given the potential for significant morbidity and mortality, a reasonable index of suspicion of BD should be maintained by clinicians when caring for a patient with recurrent genital or oral aphthosis, in conjunction with ocular lesions. It is imperative to remember that the use of diagnostic criteria should not supplant clinical judgement, while ongoing suspicion of the diagnosis should be maintained even if not initially fitting, in light of the commonly variable course of BD.

Disclosure Statement

There are no conflicts of interest, financial or otherwise, held by the authors, and no acknowledgements are to be made.

References