Survey and Validation of the Criteria for Behçet's Disease Recently Used in Korea: a Suggestion for Modification of the International Study Group Criteria

The criteria employed for Korean Behçet's disease (BD) from January 1990 to December 2000 have been investigated, and the diagnostic validity for those criteria was determined. For the generation of a modified set of preliminary criteria from the International Study Group (ISG) criteria, the diagnostic values for individual feature of BD were calculated. The criteria by the Behçet's Disease Research Committee of Japan appeared to be widely employed with the ISG criteria. However, because the ISG criteria revealed a relatively valid outcome in Korea, the application of this criteria will be needed for the universal unification until the criteria with better performance comes out. On the other hand, the modified set of preliminary criteria that consisted of the clinical items with better results seemed to improve some pitfalls of the ISG criteria. Although that criteria showed better performance than the preexisting criteria, it should be necessary to validate its effectiveness in other areas.

Key Words : *Behçet's Syndrome; Diagnosis; Criteria; Sensitivity and Specificity*

Hyun Kyu Chang, Seung Yong Kim*

Department of Internal Medicine, Dankook University Hospital, Cheonan; Department of Internal Medicine*, University of Ulsan Asan-Kangnung Hospital, Kangnung, Korea

Address for correspondence
Hyun Kyu Chang, M.D.
Division of Rheumatology, Department of Internal Medicine, Dankook University Hospital, College of Medicine, Dankook University, 16-5 Anseo-dong, Cheonan 330-715, Korea
Tel : +82-41-550-3915, Fax : +82-41-556-3256
E-mail : hanks22@dankook.ac.kr

INTRODUCTION

Behçet's disease (BD) is a chronic inflammatory multisystemic disorder with unknown etiology. Because the pathognomonic clinical features and tools are absent, the diagnosis of BD mainly relies on the characteristic clinical features and the judgement of the experienced physician. Before the introduction of the International Study Group (ISG) criteria (1), several sets of classification or diagnosis criteria had widely been employed (2-5), which had interfered with the comparison or collaboration of the clinical studies between centers or countries. Therefore, the ISG criteria was developed to ensure the uniformities of patients in the clinical studies, rather than the diagnosis of the individual case (6).

In spite of the internationally agreed criteria, the ones used in Korea for the clinical studies of BD still have not been unified. In addition, this country had not taken part in the original survey for the ISG criteria. Even though this criteria has shown the relatively valid performance in different regions (7-10), some problems have been raised. Although patients with BD in the original survey for the ISG criteria were recruited from the 12 centers in 7 countries, most patients (86.7%) were derived from the countries with high positive rate of the pathergy reaction (PR) such as Iran, Turkey, and Japan (6),

and the sensitivity of that criteria could be low in the ethnic areas where the positive PR is less prevalent (8). Another problem is that the ISG criteria may not classify or diagnose acute cases without recurrent oral ulcerations (ROUs) over 3 times in a 1-yr, patients who do not have the oral ulcerations, or patients in whom the oral ulcerations manifest later (8, 11, 12).

Therefore, this study was undertaken to investigate the criteria for BD commonly used in Korea after the introduction of the ISG criteria, to assess the diagnostic validity for those criteria, and to propose a modified set of preliminary criteria with better performance from the ISG criteria.

MATERIALS AND METHODS

From January 1990 to December 2000, the criteria applied to the clinical studies for Korean patients with BD have been investigated with the exception of review articles or case reports.

Seventy unselected patients with BD were recruited at a tertiary referral center from March 1997 to March 2001, in whom the diagnosis was established by a rheumatologist familiar with BD. Patients with only orogenital ulceration were excluded. In addition, 70 patients, matched age and sex, who

had similar features to BD (diagnoses based on internationally accepted criteria) or recurrent aphthous stomatitis were studied as controls. Patients with BD and controls were ethnically homogenous Koreans.

Detailed data on the presence or absence of the clinical features of BD in each individual with BD or controls were prospectively recorded on a standard form and were entered into a computer database. The sensitivity, specificity, and accuracy for the criteria commonly used in Korea were determined: sensitivity is defined as the proportion of subjects with the disease who have a positive test result, specificity is the proportion of subjects without the disease who have a negative test result, and accuracy is the proportion of all test results, both positive and negative, which are correct. For the generation of a modified set of preliminary criteria, the log-likelihood ratio and the expected weight of evidence for the individual clinical feature of BD were calculated with the same method that the ISG criteria was formulated (6, 13).

RESULTS

Fifty-seven clinical articles for BD have been published during the investigated period in Korea. Among these studies, 51 papers have employed 59 criteria of BD, while the remaining 6 have not made mention of the criteria. Thirty-four articles have used the criteria by the Behçet's Disease Research Committee of Japan (Japanese criteria) (2, 14), 21 by the ISG, and 4 by others, of which 3 employed Lehner and Barnes (15) and 1 Mason and Barnes (3). Among the articles that have used the Japanese criteria, 18 included all types, such as complete, incomplete, suspected, and possible types, for the classification or diagnosis of BD in their clinical studies. Twelve comprised the other 3 types except for the possible type and the remaining 4 comprised only two types, the complete and incomplete types. The ISG criteria were first employed in 1994 in Korea (16), and although this criteria has been increasingly used year after year, the Japanese criteria has still been popularly adopted (Fig. 1).

There were no differences of the demographic findings between patients with BD and controls (Table 1). The diagnoses of patients in the controls were presented Table 2, and these were similar to controls of other studies (1, 8). However, patients with inflammatory bowel diseases who might be difficult to be differentiated from BD were supplemented in the controls of the current study.

Table 3 showed sensitivity and specificity, log-likelihood ratio, and the expected weight of evidence of each clinical feature in patients with BD. As for the expected weight of evidence, genital ulcerations revealed the greatest value, and skin lesions, PR, ROUs, HLA-B51, ocular lesions, and ileocecal ulcerations in order showed the relatively good performance. However, other features, such as CNS lesions, vascular lesions, peripheral arthritis, and epididymitis, disclosed negligible val-

Table 1. The demographic findings of patients with Behçet's disease and controls

	Behcet disease (n=70)	Control (n=70)
Sex (female %)	68.6	65.7
Age (yr) (mean±SD)	39.1±10.2	38.6±13.0
Disease duration [yr, mean (range)]	5.7 (0.5-40)	5.5 (0.6-37)

Table 2. The diagnoses of patients in the controls

Diagnosis	Number (%)
Systemic lupus erythematosus	20 (28.6)
Spondylarthropathy	16 (22.9)
Sjögren syndrome	12 (17.1)
Ulcerative colitis	8 (11.4)
Crohn's disease	1 (1.4)
Recurrent aphthous stomatitis	13 (18.6)

Table 3. The measures of sensitivity, specificity, log-likelihood ratio and the expected weight of evidence of each clinical feature of Behçet's disease

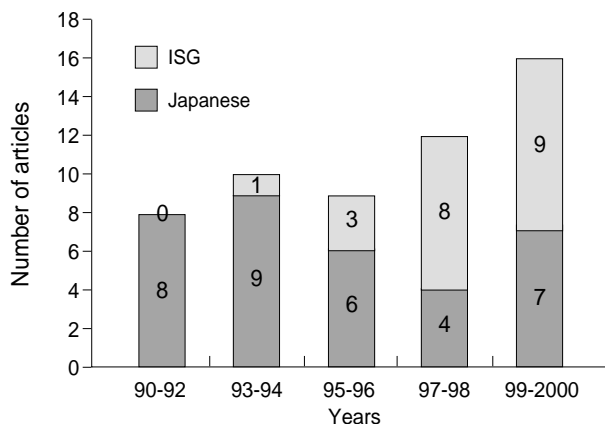| Features	Sensitivity	Specificity	Log-likelihood ratio Feature present	Log-likelihood ratio Feature absent	Expected weight of evidence
Genital ulcers	77.1	98.6	40.1	-14.6	34.2
Skin lesions	85.7	74.3	12.0	-16.5	12.7
Pathergy reaction	34.3	97.1	24.7	-3.9	11.1
Recurrent oral ulcers	97.1	57.1	8.2	-29.8	8.9
HLA-B51	50.0	83.3	11.0	-5.1	8.1
Ocular lesions	21.4	95.7	16.0	-2.0	5.1
Ileocecal ulcers	15.7	98.6	24.2	-1.6	5.1
CNS lesions	4.3	97.1	3.9	-0.1	0.3
Vascular lesions	5.7	95.7	2.8	-0.1	0.3
Peripheral arthritis	22.9	75.7	-0.6	0.2	0
Epididymitis	2.9	98.6	7.3	-0.2	0.4

Fig. 1. The annual trend of the criteria recently used in Korea. ISG: Number of articles using the criteria by the International Study Group. Japanese: Number of articles using the criteria by the Behçet's Disease Research Committee of Japan.

Table 4. A suggested preliminary criteria

Clinical Features	Score
1. Recurrent genital ulcerations	2
Painful aphthous ulceration or scarring confidently detected by physician or patient	
The exclusion of genital ulcerations associated with herpes genitalis, chancre, or chancroid	
2. Recurrent oral ulcerations	1
Painful aphthous ulceration confidently detected by physician or patient	
3. Skin lesions	1
a) Erythema nodosum-like lesions confidently detected by physician or patient	
b) Pseudofollicullitis or papulopustular lesions only detected by physician with the exception of lesions related with puberty or corticosteroid therapy	
4. Ocular lesions	1
Anterior uveitis, posterior uveitis or retinal vasculitis diagnosed by ophthalmologist	
5. Pathergy reaction	1
Papule or pustule observed by physician at 48 hr, done by intradermal prick with 20-22 gauge disposable needle	
6. Ileocecal ulcerations	1
The exclusion of inflammatory bowel disease or intestinal tuberculosis	

Behçet's disease can be classified or diagnosed if the score is more than 3. *Finding helpful for diagnosis or classification: HLA-B51

Table 5. The sensitivity, the specificity, and the accuracy of each criteria

Criteria	Sensitivity (%)	Specificity (%)	Accuracy (%)
International Study Group	91.4	97.1	94.3
Japanese	85.7	95.7	90.7
New preliminary criteria	97.1	97.1	97.1

ues. A modified set of preliminary criteria for the classification or diagnosis of BD was established by the clinical features with better results. The genital ulcerations were weighed by the score of 2, and the remaining features were given with the score of 1 with the exception of HLA-B51 not being a clinical feature, showing a lot of the geographic variation (17). HLA-B51 was separately set as a finding to help for classification or diagnosis. BD could be classified or diagnosed if the score is over 3 (Table 4).

Table 5 depicted sensitivity, specificity and accuracy of the three sets of criteria such as the ISG criteria, the Japanese criteria, and the modified set of preliminary criteria, in which the modified criteria revealed the best performance with the accuracy of 97.1%.

DISCUSSION

After the introduction of the ISG criteria, it appeared that the Japanese criteria has still been widely employed, together with the ISG criteria in Korea. Moreover, most studies to adopt the Japanese criteria have included possible type or suspected type, the potential disease entity rather than the definitive diagnosis of BD, which has not only limited the comparative studies or epidemiological research between centers or areas, but also caused the bias of their results. Because the ISG criteria has revealed a valid performance in Korea, the application of this criteria in the clinical studies of BD will be needed

for the universal unification from now on until the criteria with better performance comes out.

The reasons that the Japanese criteria had a lower diagnostic value when compared to the ISG criteria in present study may be as follows. First, some clinical features such as CNS lesions, vascular lesions, peripheral arthritis, and epididymitis among minor criteria of the Japanese criteria were uncommon in our series. Second, the Japanese criteria could not classify or diagnose patients with ROUs, skin lesions, and a positive PR, who were classified into suspected type.

The prevalence of a positive PR varies according to countries or investigators. The high positive rate of PR has been described in some ethnic areas such as Turkey, Iran, and Japan (18-20), but this was not the case in other countries (21, 22). Furthermore, after the introduction of a disposable needle, the prevalence and intensity of a positive PR in patients with BD have been described to decrease (23). In the original survey for the ISG criteria, most patients with BD were recruited from the areas with a high positive rate of PR, and the PR has become one of the major criteria. Therefore the sensitivity of such criteria could be low in the areas where a positive PR is uncommon. On the other hand, patients with only orogenital ulcerations, of whom most patients may be considered to have a mild BD, do not meet the ISG criteria. If the sensitivity of a positive PR would be greater, a considerable portion of those patients might be classified or diagnosed as BD. Because of the exclusion of patients with only orogenital ulcerations, the diagnostic value of the ISG criteria and the Japanese criteria may be speculated to be higher in the current study.

Although the ISG have considered the ROUs to be a 'conditio sine qua non' for the classification or diagnosis of BD, 3% of patients with definitive BD in whom the oral ulcerations did not manifest were excluded in the original survey (1). Whereas the ROUs have been the most common and important clinical feature, a sizable portion of patients have initially manifested without ROUs with variable rate of 13.5-27% (24-27). In addition, other authors pointed out that the ISG criteria could not classify or diagnose acute cases with severe disease without ROUs over 3 times in a 1-yr, which may delay a proper diagnosis and the prompt initiation of treatment (11, 12, 28). If the sensitivity and specificity of the designated criteria were very high, a classification criteria could be used as a diagnostic criteria, and would apply to every individual case (29-31). Therefore, in the classification or diagnosis of BD, the criteria with better performance to settle the suggested problems may be necessary.

Although our preliminary criteria is in some degree derived from the ISG criteria, there are a few differences between these two criteria. First of all, with reference to the expected weight of evidence, the diagnostic power of the genital ulcerations is doubled when compared to other clinical items. Consequently, the classification or diagnosis of BD could be established by genital ulcerations plus one more other clinical items. However, the genital ulcerations by other causes such as sexually transmitted diseases (genital herpes, chancer, chancroid) should be ruled out. Second, the ROUs are not a 'conditio sine qua non' in our criteria. Moreover, as the frequency (over 3 times in a 1-yr) of them is deleted, the classification or diagnosis of BD could be made even in cases without ROUs or in acute patients that there are not ROUs over 3 times in a 1-yr. Third, the ileocecal ulcerations, one of the serious clinical manifestations, are added with the exclusion of those by other causes such as inflammatory bowel disease and intestinal tuberculosis.

Similarly to our study, the genital ulcerations also showed the greatest diagnostic value in the original study by the ISG (1). As the clinical features of genital ulcerations in BD are quite typical, these lesions can be detectable by an experienced physician through a careful history taking and physical examination. They usually occur on the scrotum and penis in men and on the vulva in women. Genital ulcers resemble oral ulcers but are often larger and deeper, and are more painful and last longer. They often leave scars and recurrences occur less frequently than oral ulcers (19, 20). However, besides the genital ulcerations of BD, there are other causes of genital ulcerations such as sexually transmitted diseases. A clinical diagnosis of genital herpes can be made if typical painful, grouped vesicles or pustules are preceded by a prodrome of stinging or burning. Primary syphilis (chancer) may be easily differentiated from those of BD because this disease has painless, nontender, indurated ulcers frequently accompanied with firm, nontender inguinal adenopathy. Chancroid, another cause of genital ulceration, has a painful ulcer with ragged, undermined borders and associated tender inguinal lymphadenitis. Nevertheless, because some genital ulcerations cannot be diagnosed confidently on clinical grounds alone, the following examinations may be needed: dark-field examination or serologic tests for *Treponema pallidum*, serologic tests for *Herpes simplex* virus, and culture or PCR tests for *Haemophilus ducrei* (32).

Because the prevalence of some clinical features of BD differs according to areas or countries (19), it may be difficult to make the criteria that embrace all patients with BD in the different areas of the world. For example, the ileocecal ulcerations relatively prevalent in our series and in Japan (20) have been rarely described in Turkey (33). Therefore, even though our preliminary criteria may improve the classification or diagnosis of patients with the intestinal ulcerations, it would be not the case in Turkey. However, this problem might be overcome with an emphasis of the clinical significance of the genital ulcerations. Conclusively, our preliminary criteria may

not only classify or diagnose a mild BD patients with only orogenital ulcerations, patients without ROUs, or acute cases without ROUs over 3 times in a 1-yr, but also increase the sensitivity for the classification or diagnosis of BD in the ethnic areas with low positive rate of the PR. Although our criteria showed greater performance when compared with the existing two criteria in the current study, it should be necessary to validate its effectiveness for patients in other clinics or countries, because the number of patients with BD was small and the validation sample of patients like the original survey by the ISG (6) was lacking in the present study. The validation assay in several centers of the different areas of Korea is going to be under way.

REFERENCES

1. International Study Group for Behçet's Disease (ISGBD). *Criteria for diagnosis of Behçet's disease. Lancet 1990; 335: 1078-80.*
2. Behçet's Disease Research Committee of Japan. *Behçet's disease: guide to diagnosis of Behçet's disease. Jpn J Ophthalmol 1974; 18: 291-4.*
3. Mason RM, Barnes CG. *Behçet's syndrome with arthritis. Ann Rheum Dis 1969; 28: 95-103.*
4. O'Duffy JD. *Criteres proposes pour le diagnostic de la maladie de Behçet et notes therapeutiques. Rev Med 1974; 36: 2371-9.*
5. Dilsen N, Konice M, Aral O. *Our diagnostic criteria for Behçet's disease: an overview. In: Lehner T, Barnes CG, Editors, Recent advances in Behçet's disease. London: Royal Society of Medicine Services: International Congress and Symposium Series 103, 1986: 177-80.*
6. *International Study Group for Behçet's Disease. Evaluation of diagnostic ('classification') criteria in Behçet's disease-towards internationally agreed criteria. Br J Rheumatol 1992; 31: 299-308.*
7. Ferraz MB, Walter SD, Heymann R, Atra E. *Sensitivity and specificity of different diagnostic criteria for Behçet's disease according to the latent class approach. Br J Rheumatol 1995; 34: 932-5.*
8. O'Neill TW, Rigby AS, Silman AJ, Barnes C. *Validation of the International Study Group criteria for Behçet's disease. Br J Rheumatol 1994; 33: 115-7.*
9. Heyman RE, Ferraz MB, Goncalves CR, Atra E. *Evaluation of the International Study Group for Behçet's Disease Criteria in Brazilian patients. Clin Rheumatol 1995; 14: 526-30.*
10. Tunc R, Uluhan A, Melikoğlu M, Ozyazgan Y, Özdoğan H, Yazici H. *A reassessment of the International Study Group criteria for the diagnosis (classification) of Behçet's syndrome. Clin Exp Rheumatol 2001; 19 (Suppl 24): S45-7.*
11. Mignogna MD, Fedele S, Lo Russo L. *International diagnostic criteria and delay of diagnosis in Behçet's disease. J Rheumatol 2000; 27: 2725.*
12. Shimizu S, Chen KR, Ikemoto K, Han-Yaku H. *Abrupt onset of severe Behçet's disease: preceding oral ulceration is not essential for diagnosis. Br J Dermatol 1998; 139: 160-1.*
13. Spiegelhalter DJ. *Statistical methodology for evaluating gastrointestinal symptoms. Clin Gastroenterol 1985; 14: 489-515.*

14. Mizushima Y. *Recent research into Behçet's disease in Japan. Int J Tissue React 1988; 10: 59-65.*

15. Lehner T, Barnes CG. *Criteria for diagnosis and classification of Behçet's syndrome. In: Lehner T, Barnes CG, Editors, Behçet's syndrome. London: Academic Press, 1979: 1-9.*

16. Kim DK, Chang SN, Bang D, Lee ES, Lee S. *Clinical analysis of 40 cases of childhood-onset Behçet's disease. Pediatr Dermatol 1994; 11: 95-101.*

17. Verity DH, Marr JE, Ohno S, Wallace GR, Stanford MR. *Behçet's disease, the silk road and HLA-B51: historical and geographical perspectives. Tissue Antigens 1999; 54: 213-20.*

18. Tüzün Y, Yazici H, Pazarli H, Yalcin B, Yurdakul S, Muftuoglu A. *The usefulness of the nonspecific skin hyperreactivity (the pathergy test) in Behçet's disease in Turkey. Acta Derm Venereol 1979; 59: 77-9.*

19. Kaklamani VG, Vaiopoulos G, Kaklamanis PG. *Behçet's Disease. Semin Arthritis Rheum 1998; 27: 197-217.*

20. Shimizu T, Ehrlich GE, Inaba G, Hayashi K. *Behçet disease (Behçet syndrome). Semin Arthritis Rheum 1979; 8: 223-60.*

21. O'Duffy JD. *Summary of international symposium on Behçet's disease. Istanbul, September 29-30, 1977. J Rheumatol 1978; 5: 229-33.*

22. Davies PG, Fordham JN, Kirwan JR, Barnes CG, Dinning WJ. *The pathergy test and Behçet's syndrome in Britain. Ann Rheum Dis 1984; 43: 70-3.*

23. Dilsen N, Konice M, Aral O, Ocal L, Inanc M, Gul A. *Comparative study of the skin pathergy test with blunt and sharp needles in Behçet's disease: confirmed specificity but decreased sensitivity with sharp needles. Ann Rheum Dis 1993; 52: 823-5.*

24. Güler A, Boyvat A, Türsen Ü. *Clinical manifestations of Behçet's disease: an analysis of 2147 patients. Yonsei Med J 1997; 38: 423-7.*

25. Zouboulis CC, Kotter I, Djawari D, Kirch W, Kohl PK, Ochsendorf FR, Keitel W, Stadler R, Wollina U, Proksch E, Sohnchen R, Weber H, Gollnick HP, Holzle E, Fritz K, Licht T, Orfanos CE. *Epidemiologial features of Adamandiades-Behçet's disease in Germany and in Europe. Yonsei Med J 1997; 38: 411-22.*

26. Bang D, Yoon KH, Chung HG, Choi EH, Lee ES, Lee S. *Epidemiological and clinical features of Behçet's disease in Korea. Yonsei Med J 1997; 38: 428-36.*

27. Gharibdoost F, Davachi F, Shahram F, Akbarian M, Chams C, Chams H, Mansoori P, Nadji A. *Clinical manifestations of Behçet's disease in Iran analysis of 2176 cases. In: Godeau P, Wechsler B Editors, Proceedings of the 6th International Conference on Behçet's Disease. Amsterdam: Elsevier Science Publishers, 1993: 153-8.*

28. Lee S. *Diagnostic criteria of Behçet's disease: probems and suggestions. Yonsei Med J 1997; 38: 365-9.*

29. Fries JF, Hochberg MC, Medsger TA Jr, Hunder GG, Bombardier C. *Criteria for rheumatic disease. Different types and different functions. Arthritis Rheum 1994; 37: 454-62.*

30. Van der Linden S, Valkenburg HA, Cats A. *Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984; 27: 361-8.*

31. Edworthy SM, Zatarain E, McShane DJ, Bloch DA. *Analysis of the 1982 ARA lupus criteria data set by recursive partitioning methodology: new insights into the relative merit of individual criteria. J Rheumatol 1988; 15: 1493-8.*

32. Lucas DJ, Buntin DM. *Approach to the patient with sexually transmitted disease. In: Freedberg IM, Eisen AZ, Wolff K, Austen KF, Goldsmith LA, Katz SI, Fitzpatrick TB, editors, Dermatology in general medicine. New York: McGraw-Hill, 1999; 2547-51.*

33. Yurdakul S, Tuzuner N, Yurdakul I, Hamuryudan V, Yazici H. *Gastrointestinal involvement in Behçet's syndrome: a controlled study. Ann Rheum Dis 1996; 55: 208-10.*