

Evaluation of the Turkish translation of a disease activity form for Behçet's syndrome

V. Hamuryudan, İ. Fresko, H. Direskeneli¹, M. J. Tenant²,
S. Yurdakul, T. Akoğlu¹ and H. Yazıcı

Division of Rheumatology, Department of Internal Medicine, Cerrahpaşa Medical Faculty, University of İstanbul, ¹Division of Rheumatology, Department of Internal Medicine, Marmara Medical Faculty, University of Marmara, İstanbul, Turkey and ²University of London, London, UK

Abstract

Objective. This study examined the interobserver and intra-observer reliability of the Turkish version of the Behçet's Disease Current Activity Form (BDCAF), which was obtained by a translation and back-translation process.

Methods. Fifty Behçet's syndrome (BS) patients were assessed by four rheumatologists in separate morning and afternoon sessions.

Results. The results showed good intra- and interobserver agreement for the oro-genital ulcers and eye involvement of BS, but there was poor agreement between (kappa score = 0.14) and within observers (range for kappa scores 0.09–0.25) for their overall impression of disease activity. Individual low kappa scores were also noted for erythema nodosum, vascular involvement, central nervous system involvement and gastrointestinal involvement.

Conclusion. These results suggest that the Turkish version of BDCAF may be useful for assessing the classic triad of BS (oro-genital ulceration and eye involvement), but more experience is needed for its other parts.

KEY WORDS: Behçet's syndrome, Disease activity, Behçet's Disease Current Activity Form, Turkish translation, Reliability analysis.

Disease activity in Behçet's syndrome (BS) is difficult to define [1]. The syndrome has a heterogeneous nature of organ involvement, runs a fluctuating disease course, and lacks reliable laboratory indices that are useful in reflecting overall disease activity [2]. Therefore, judgement of disease activity for BS in clinical practice is mainly subjective. Furthermore, there is also no agreement between investigators even for accurately reporting on various manifestations of BS. A recurrent oral ulcer can be scored by its number, size, frequency or the time needed for healing, and it is not known which of these reflect the actual activity [3]. Thus, there is a need for a clinical instrument to assess disease activity in BS.

There have been previous attempts to develop an internationally accepted disease activity measure [3–5]. Recently, a new instrument, the Behçet's Disease Current Activity Form (BDCAF), has been developed at the University of Leeds, UK [6]. This form showed good interobserver reliability in a formal study [6] and is now routinely used in the clinic where it has been developed. However, the marked geographical differences in disease expression of BS [7, 8], as well as

possible ethnic and intercultural differences in disease impact among individuals from different geographical regions, require the evaluation of this form in other countries before recommending its clinical use universally. We therefore translated the activity form into Turkish and tested its usefulness among Turkish patients.

Methods

Patients

The Behçet's Syndrome Research Centre at the Cerrahpaşa Medical Faculty is a multidisciplinary outpatient clinic, which regularly meets every Monday with approximately five new and 60 previously registered patients. Consecutive patients attending this clinic and fulfilling the international criteria [9] were initially screened for eligibility by one rheumatologist who did not take part in the assessment. Patients having at least one active manifestation of BS within the previous 4 weeks and who gave their verbal consent were enrolled. All patients were coded to prevent their identification and were randomized for their order of appearance during the study.

Submitted 2 July 1998; revised version accepted 15 March 1999.

Correspondence to: V. Hamuryudan, Veysipasa sokak 100, Yıl Sitesi I Blok D16, 81190 Uskudar—İstanbul, Turkey.

Instrument

The Behçet's Current Activity Index has been described in detail elsewhere [6]. In brief, this activity form scores (from 0 to 4) the duration of clinical features (oral ulcers, genital ulcers, skin lesions, etc.) which have been present during the 4 weeks prior to the day of assessment. Eye activity is assessed for the presence or absence of blurring of vision or pain or redness in any eye. Additionally, the patients and clinicians were asked to rate their impressions of the overall disease activity within the preceding 4 weeks by indicating on a scale consisting of seven faces with different expressions.

Translation

A physician who was aware of the purpose of the translation initially translated this form into Turkish. Subsequently, a non-medical person translated it again back into English to assess whether substantial differences in meaning were created during the translation. Finally, two rheumatologists discussed the discrepancies in the translations and made their decisions on the final appearance of the form (the English and Turkish versions of the activity form are available on request).

Design of the study

The study was performed on two different days. Every patient was evaluated twice in one day with separate morning and afternoon sessions. Four rheumatologists took part in the evaluation. To overcome a locality bias, two of them came from another institution in Istanbul, which also has a special interest in BS. All four observers had a brief meeting regarding the use of the form before the evaluation began.

Statistical analysis

Intra-observer reliability was assessed from the morning and afternoon observations of the same observer using Cohen's kappa, a measure of agreement between two raters [10]. Cohen's kappa is sensitive only to the presence or absence of agreement between observers and not to the relative strength of disagreement, when present. All kappas have a maximum value of one (complete agreement), with zero meaning no agreement beyond that which can be expected by pure chance alone.

Interobserver reliability was assessed between all four observers, in both the morning and afternoon. A generalized k-rater kappa described by Fleiss [11] was used for this purpose, which is more properly a generalization of a measure known as Scott's pi [12]. The generalized k-rater kappa is based on the pragmatic assumption that the probability that an object (patient) is assigned to a particular category (severity of manifestation) does not vary across raters.

Results

Among the 50 patients participating in the study, one did not show up for the afternoon assessment. The mean age of this population (31 males, 18 females) was

TABLE 1. The results of analysis of intra-observer agreement (kappa scores)

	Observer 1	Observer 2	Observer 3	Observer 4
Oral ulcers	0.74	0.81	0.81	0.79
Genital ulcers	0.80	0.90	0.83	0.89
Erythema nodosum/ STM	0.32	0.48	0.33	0.58
Pustules	0.41	0.60	0.65	0.62
Arthralgia	0.73	0.81	0.64	0.74
Arthritis	0.75	0.70	0.71	0.92
Gastrointestinal (I)	0.76	0.81	0.38	0.66
Gastrointestinal (II)	0.77	0.77	0.60	0.92
Pain, left eye	0.94	0.82	0.84	0.72
Blurring, left eye	0.95	0.72	0.74	0.73
Redness, left eye	0.64	0.69	0.80	0.72
Pain, right eye	0.93	0.88	0.89	0.82
Blurring, right eye	0.85	0.84	0.82	0.70
Redness, right eye	0.82	0.86	0.95	0.75
Central nervous system	0.52	0.59	0.26	0.63
Vascular involvement	0.71	0.43	0.49	0.66
Fatigue	0.55	0.75	0.82	0.70
Headache	0.84	0.84	0.75	0.68
Patient's impression of disease activity	0.74	0.69	0.74	0.65
Doctor's impression of disease activity	0.09	0.25	0.13	0.13

STM, superficial thrombophlebitis.

(I), nausea or vomiting or abdominal pain; (II), diarrhoea with altered or frank blood.

TABLE 2. The results of analysis of interobserver agreement (kappa scores)

	Morning session	Afternoon session
Oral ulcers	0.79	0.86
Genital ulcers	0.85	0.86
Erythema nodosum/STM	0.28	0.60
Pustules	0.60	0.60
Arthralgia	0.81	0.65
Arthritis	0.81	0.66
Gastrointestinal (I)	0.65	0.62
Gastrointestinal (II)	0.92	0.82
Pain, left eye	0.83	0.78
Blurring, left eye	0.79	0.72
Redness, left eye	0.94	0.84
Pain, right eye	0.90	0.83
Blurring, right eye	0.77	0.70
Redness, right eye	0.85	0.75
Central nervous system	0.33	0.41
Vascular involvement	0.64	0.58
Fatigue	0.83	0.70
Headache	0.84	0.81
Patient's impression of disease activity	0.79	0.80
Doctor's impression of disease activity	0.14	0.14

STM, superficial thrombophlebitis.

(I), nausea or vomiting or abdominal pain; (II), diarrhoea with altered or frank blood.

33 ± 5.8 (s.d.) yr and they had a disease duration of 4.2 ± 1.9 (s.d.) yr.

The completion of the form took ~4 min on average for each observer. The results of the analysis of intra-

observer agreement are presented in Table 1. The lowest kappa scorings were obtained for the observer's impression of overall disease activity using the seven different faces. Moderately low kappa scorings were also obtained for erythema nodosum (three observers), pustules (one observer), gastrointestinal involvement (one observer), central nervous system (one observer) and vascular involvement (two observers).

Table 2 shows the results of the analysis of interobserver agreement. The lowest kappa values were again obtained for the observer's overall impression of disease activity both in the morning and the afternoon sessions. Low kappa values were also obtained for central nervous system involvement in either session and for erythema nodosum in the morning session.

Discussion

This study showed differences in the reliability scores of different parts of the disease activity form after its translation into Turkish. Questions directed at the classic triad of BS—oral and genital ulcers and eye involvement—had the highest kappa scores, but there was moderately low agreement in the assessment of, especially, the existence of erythema nodosum and central nervous system involvement.

Our study demonstrated poor agreement between and within observers for their impressions of overall disease activity. However, this result is not surprising [13] and indeed explains the necessity of attempts to develop reliable disease activity measures. This disagreement, we believe, is most probably the consequence of the heterogeneous nature of organ involvement in BS.

Our results are somewhat different from the findings in the English study, which showed a better reliability of the form [6]. The translation process into Turkish differs from the proposed guidelines for cross-cultural adaptation of health-related quality of life measurements only by the absence of a pre-test phase [14], but none of the patients involved in our study volunteered difficulty in understanding the questions. Thus, before considering an essential construction problem for the parts with low agreement, we have to consider the fact that our colleagues from the UK probably had more clinical experience with the use of the activity form, since it was originally developed by them.

This activity form does not have an overall activity score, i.e. a composite index, deduced from the individual scores for different organ systems. The heterogeneous nature of disease expression in BS, both within and between organ systems, makes this difficult to achieve. Furthermore, one can also debate whether such a composite index would jeopardize accuracy in seeking simplicity. On the other hand, this has been achieved in systemic lupus erythematosus, a condition at least as heterogeneous as BS [15].

We conclude that some parts of the BDCAF (oro-

genital ulcers and eye involvement) can be reliably translated into Turkish. However, more experience is needed before all parts of it can be used confidently in daily care and clinical trials. Furthermore, the ultimate test of its usefulness will be to observe its sensitivity and specificity to change over time, and this can only be assessed with its more widespread and dedicated use.

References

1. Chamberlain MA, Noble BA, Behçet's U.K. Study Group. Disease activity in Behçet's disease. In: O'Duffy JD, Kökmen E, eds. Behçet's disease. Basic and clinical aspects. New York: Marcel Dekker, 1991:299–302.
2. Müftüoğlu AÜ, Yazıcı H, Yurdakul S *et al.* Behçet's disease. Relation of serum C-reactive protein and erythrocyte sedimentation rates to disease activity. *Int J Dermatol* 1986;25:235–9.
3. Bhakta B, Hamuryudan V, Brennan P, Chamberlain MA, Barnes CG, Silman AJ. Assessment of disease activity in Behçet's disease. In: Wechsler B, Godeau P, eds. Behçet's disease. Amsterdam: Elsevier Science, 1993:235–40.
4. Davatchi F, Akbaran M, Shahram F *et al.* Iran Behçet's disease dynamic activity index. *Hung Rheumatol* 1991;32(suppl.):FP10–100, 134 (Abstracts of the XIIth European Congress of Rheumatology).
5. Yazıcı H, Tüzün Y, Pazarlı H *et al.* Influence of age of onset and patient's sex on the prevalence and severity of manifestations of Behçet's syndrome. *Ann Rheum Dis* 1984;43:783–9.
6. Bhakta BB, Brennan P, James TE, Chamberlain MA, Noble BA, Silman AJ. Behçet's disease: evaluation of a new instrument to measure clinical activity. *Rheumatology* 1999;38:728–33.
7. Yazıcı H, Chamberlain MA, Tüzün Y, Yurdakul S, Müftüoğlu A. A comparative study of the pathergy reaction among Turkish and British patients with Behçet's disease. *Ann Rheum Dis* 1984;43:74–5.
8. Ehrlich GE. Behçet's disease and the emergence of thalidomide. *Ann Intern Med* 1998;128:494–5.
9. Criteria for diagnosis of Behçet's disease. International Study Group for Behçet's Disease. *Lancet* 1990;335:1078–80.
10. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Measurement* 1960;20:37–46.
11. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971;76:378–82.
12. Scott WA. Reliability of content analysis. The case of nominal scale coding. *Public Opin Q* 1955;19:321–5.
13. Kirwan JR, Chaput de Saintonge DM, Joyce CRB, Currey HLF. Clinical judgment in rheumatoid arthritis. I. Rheumatologists' opinions and the development of 'paper patients'. *Ann Rheum Dis* 1983;42:644–7.
14. Guillemin F, Bombardier C, Beaton D. Cross cultural adaptation of health related quality of life measures: Literature review and proposed guidelines. *J Clin Epidemiol* 1993;46:1417–32.
15. Gladman DD, Goldsmith CH, Urowitz MB *et al.* Crosscultural validation and reliability of 3 disease activity indices in SLE. *J Rheumatol* 1992;19:608–11.