Evaluation of a new Tunisian version of behçet's disease current activity form[*]

Olfa Harzallah[#], Yassine Meksi, Amira Hamzaoui, Rym Klii, Amira Atig, Silvia Mahjoub

Internal Medicine Department, Fattouma Bourguiba Hospital, Monastir, Tunisia
Email: #harzallaholfa@yahoo.fr

ABSTRACT

Background: Behçet's Syndrome (BS) is characterized by a heterogeneous vessel involvement, a fluctuating natural history and by the absence of biological markers correlated to disease activity that's why objective clinical scores are needed for the assessment of its activity. The Behçet's Disease Clinical Activity Form (BDCAF) is the most recent and widely used clinical activity score. **Objectives:** To perform a cross-cultural adaptation of the Behçet's Disease Current Activity Form (BDCAF) to the Tunisian Dialect (Arabic Language) and to evaluate the metrological characteristics of the Tunisian version (Tu-BDCAF) especially its reliability in BD activity evaluation. **Methods:** Cross-cultural adaptation was done according to the established guidelines. Reliability of Tu-BDCAF was tested among 40 BD patients (mean age: 38 years, sex ratio: 1.37). Patients were questioned by two BD specialists at 20 minutes interval to evaluate inter-observer reproducibility and twice by the same physician at 48 hours interval to assess the intra-observer reproducibility. κ Coefficient was used to test the concordance between qualitative variables and correlation between quantitative variables was evaluated used Pearson coefficient and Bland and Altman graphical method. **Results:** There was a good correlation between global scores calculated by the two physicians on the same day (r = 0.94, p < 0.0001) and also between the scores calculated by the same clinician at different times (r = 0.98, p < 0.0001). κ Coefficient analyses demonstrated a good intra and inter observer reliability for all the Tu-BDCAF items excepted for diarrhea and Clinician's impression. As the original version, Tu-BDCAF is an objective, easy-calculated and reliable index for assessing disease activity in BD. The main limit of the BDCAF score remains the absence of a cut-off point defining BD activity. **Conclusion:** Tu-BDACF is a Tunisian version of the BDCAF score which can be used in routine to assess BD activity but also in international studies and clinical trials.

Keywords: Behçet's Disease; Disease Activity; Disease Activity Score; Reliability of Results

1. INTRODUCTION

Behçet's syndrome (BS) is a multisystem inflammatory disease originally described by Hulusi Behçet, a Turkish dermatologist, in 1937 [1]. It is characterized by recurrent oral and genital ulcers (OU and GU), skin lesions and uveitis [2].

BS is a chronic disease that progresses by unpredictable flares alternating with remission periods. A multisystem disease activity is defined by the occurrence of new symptoms or the worsening of preexisting ones [3].

To date, there is no recognized and validated definition of BS activity. Furthermore, there isn't any laboratory marker correlated to this activity.

Many clinical scores were developed to evaluate the disease activity in patients with BS: the Clinical Activity Index (CAI) [4], the IBDDAM (The Iranian Behçet's Disease current Activity Form) [5] and the BDCAF (the Behçet's Disease Current Activity Form) [6] etc.

The BDCAF, developed and validated in Leeds in 1994 as a result of the combination of the European scheme and the IBDDAM, is the most recent and most widely used index activity in BS.

The aims of our work were first to perform a cross-cultural adaptation of the BDCAF (Behçet's Disease Current Activity Form) to the Tunisian Dialect (Arabic Language) and then to test the reliability of the Tunisian version (Tu-BDCAF) in the disease activity evaluation of BS Tunisian patients.

---

2. PATIENTS AND METHODS

2.1. Patients

40 Behçet's disease (BD) patients were prospectively enrolled from March to June 2012 after signing a written informed consent. All patients fulfilled the "International Study Group of Behçet's Disease Criteria" [7] and were aged over 18 years. Blind and psychiatrically ill patients were excluded from this study.

Patients were recruited from the Internal Medicine Department of Fattouma Bourguiba Hospital of Monastir. The study was approved by the Ethic Committee of our Medical Faculty of Monastir.

2.2. Methods

First, main demographic and clinical patient's characteristics were specified. The study consisted then in two major steps: cross-cultural adaptation to the Tunisian dialect and reliability test of the Tunisian version of BDCAF score (Tu-BDCAF).

2.2.1. Cross-Cultural Adaptation

1) Instrument:

BDCAF was described and evaluated by Bahakta *et al.* [6].

It's an index based on the presence of clinical signs related solely to BD. Symptoms occurring only a month before the evaluation day are taken into account.

Fatigue, headache, OU, GU, erythema nodosum, superficial thrombophlebitis, pustules, arthralgia, arthritis, nausea, vomiting, abdominal pain and bloody diarrhea are scored between 0 and 4 according to the duration of the symptom in the preceding 4 weeks.

The scoring system is different for ocular, central nervous system and large vessels involvement.

An eye activity is considered if the patient has a red or painful eye and/or a blurred vision. The patient is referred then to an ophthalmologist which will determine the eye score (Behçet's ocular index).

Five and 4 questions explore respectively central nervous system and large vessels involvement (thromboses and/or aneurysms).

Eye, CNS and large vessels involvement are considered as dichotomous variables.

This tool also includes 3 visual scales (VS). Two of them are addressed to evaluate the patient's perception of his well-being (the evaluation day and during the previous 28 days). The third scale appraises the clinician's perception of the overall disease activity.

These 3 scales are scored according to an ordinal system from 0 to 6.

The clinician's presence is required as well to evaluate the disease activity (VS) and to specify the intended treatment modifications as well to help the patient responding to the questionnaire and giving explanations when needed.

2) Translation

The cross-cultural adaptation of the questionnaire was performed according to the guidelines of Beaton *et al.* [8].

Two Tunisian BD experts (Internal Medicine physicians) translated the BDCAF from English to Tunisian dialect. Two independent translations were hence obtained: Tu_1 and Tu_2 that were then combined into one version Tu-1,2. This version was next back translated twice (BT_1 and BT_2) to English by two non medical English language teachers unaware of the purpose of the study.

Translations and back translations were all reviewed by clinicians and English Teachers and finalized into one Tunisian BDCAF form called Tu-BDCAF.

A pre-test phase was planned. This phase consists in applying the Tu-BDCAF to 15 patients. If more than 15% of questions are not understood, these questions are changed and the pre-test phase is repeated until reaching a final version.

2.2.2. Reliability Test

Patients were questioned by two BD specialists (Tu-$BDCAF_1$ and Tu-$BDCAF_2$) at 20 min interval to evaluate inter-observers reliability and twice by the same physician at 24 hours interval (Tu-$BDCAF_1$ and Tu-$BDCAF_3$) to assess the intra-observer reliability.

2.3. Statistical Methods

κ Coefficient was used to test the concordance between qualitative variables.

The agreement level was rated according to κ Coefficient value as follows

κ coefficient	Concordance level
κ > 0.6	Good
0.41 < κ < 0.6	Moderate
<0.4	Fair

Correlation between quantitative variables was evaluated using two methods:

The intraclass correlation coefficient (Pearson coefficient): the concordance was considered good if this coefficient was ≥ 0.7.

The Bland and Altman method

Our study's data were collected and analyzed using the 18th version of Statistical Package for the Social Sciences (SPSS) software.

3. RESULTS

Patients were 23 males and 17 females with a mean age of 38 years (18 - 70 years). Their disease was progressing, on average, since 11.5 +/− 9.48 years (1 - 30 years).

Patients filled the form, on average, in 3 minutes 36 seconds (2 - 4 min 20 s). All questions were understood by all the patients.

The mean Tu-BDCAF score was 3.23 +/-3 (0 − 8) and Tu-BDCAF 2 was on average at 3.18+/−2 (0 - 7). Only 30 patients returned the second day to fill the Tu-BDCAF and the mean Tu-BDCAF3 was 3.20+/−2.1 (0 - 7).

The inter-observers ($r = 0.94$; $p < 0.0001$) as well as the intra-observer concordance of the global Tu-BDCAF score ($r = 0.98$; $p < 0.0001$) was excellent (**Table 1**). This result was confirmed by the graphic method of Bland and Altman (**Figures 1** and **2**).

Table 2 summarizes the results of κ Coefficient analyses for all Tu-BDCAF items.

κ Coefficient analyses demonstrated a moderate to good intra and inter observer's agreement for all the Tu-BDCAF items excepted for diarrhea and Clinician's impression.

Table 1. Inter-observers and intra-observer concordance of the global BDCAF score.

Correlation	r	P
Tu-BDCAF1 - Tu-BDCAF2	0.94	<0.0001
Tu-BDCAF1 - Tu-BDCAF3	0.98	<0.0001

r: intra-class correlation coefficien (Pearson coefficient) Tu-BDCAF$_1$: Tu-BDCAF score calculated by observer 1, Tu-BDCAF$_2$: Tu-BDCAF score calculated by observer 2, Tu-BDCAF3: Tu-BDCAF score calculated by observer 1 at 24 hours interval.
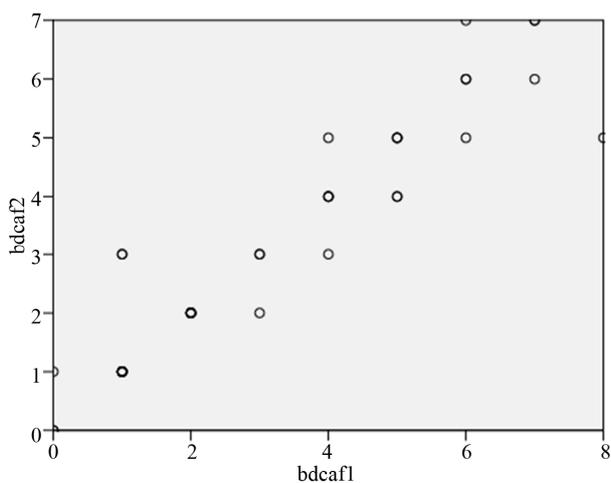

Figure 1. Concordance between Tu-BDCAF1 and Tu-BDCAF2 (intra-observer reliability of the global BDCAF score).
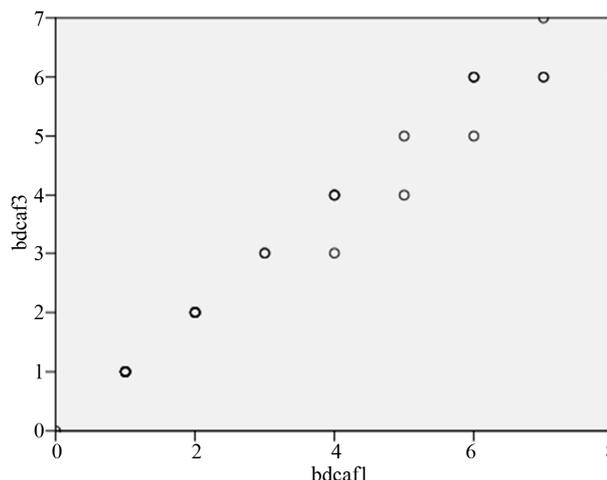

Figure 2. Concordance between Tu-BDCAF$_1$ and Tu-BDCAF$_3$ (Inter-observers reliability of Tu-BDCAF global score).

Table 2. Intra and inter-observers agreement for each item of the Tu-BDCAF score.

κ Coefficient Item	Inter-observers (n = 40)	Intra-observer (n = 30)
Patient well-being (last month)	0.58	0.28
Patient well-being (evaluation's day)	0.55	0.74
Clinician's impression	0.28	0.68
Headache	0.59	0.89
Oral aphtosis	0.92	1
Genital aphtosis	0.68	0.85
Erythema nodosum	0.56	0.77
Necrotizing Pseudofolliculitis	0.73	0.81
Arthralgias	0.78	0.84
Arthritis	0.48	0.34
Nauseas /Itching	0.56	0.51
Diarrhea	0.25	0.71
Oular involvement	0.79	0.93
Neurological involvement	0.81	1
Large vessels involvement	0.75	1

4. COMMENTS

BD is a disorder characterized by a heterogeneous clinical involvement, a fluctuating natural history and by the absence of biological markers correlated to disease activity. Hence, BD activity can't be assessed meaning one simple judgment criterion and composite indices are needed. In this aim, several scores were elaborated. Yazici [4] developed the Clinical Activity Index (CAI) in 1984. This score evaluates clinical symptoms present at the time of

evaluation. In 1991, Davatchi [5] suggested a clinical index, the Iranian Behçet's disease Dynamic Measure (IBDDAM), based on the recall of BD related symptoms, one year prior to the disease activity evaluation. In the same year, Chamberlain, Barnes and Silman [9] developed the European Scheme which incorporates features of the CAI.

The BDCAF resulted, after an expert's consensus held in Leeds in 1994, from the incorporation of European and Iranian forms [6].

This index is the most recent and most widely used BD activity score.

To be "safely" used in routine, a clinical score must respond to well codified quality norms. A "good clinical score" must be valid, reliable, sensitive to change and easy to apply.

The BDCAF has already been validated in Leeds in 1994. Its reliability was then tested by Bhakta *et al.* [6] which demonstrated a good inter-observer reliability in overall assessing BD activity.

This score is now widely utilized rather in practice by BD experts than in clinical trials but in order to be internationally used. A cross-cultural adaptation process is required first because of the interethnic and geographic variability of BD and second because of the intercultural difference of patient's perception of the impact of the disease.

The BDCAF was adapted only to the Brazilian Portuguese [10] and to the Turkish languages [11]. We aimed, so, to perform a cross-cultural adaptation of the BDCAF to the Tunisian dialect. Tunisia is a north-African Arabic country. Tunisian people write Arabic but speak Tunisian dialect which is a modified version of Arabic.

To realize this adaptation, we followed the guidelines of Beaton *et al.* [8]. Neves [10] also followed these guidelines to adapt the BDCAF to the Brazilian Portuguese language but that wasn't the case for the Turkish version where only one BD expert and one English translator were implicated [11].

The BDCAF form was easy to understand. We didn't have to perform the pre-test phase included in Beaton's guidelines. The Tu-BDCAF questions were understood by all the patients.

Its application was also easy. It took, on average, 3 minutes 36 seconds to calculate the score. This result confirmed previous findings of Bhakta, Neves and Hamuruydan [6,10,11] who respectively found a mean time completion of BDCAF form at 5 - 10, 4 and 4 minutes.

The follow-up of the cross-cultural adaptation process must be rigorous, because in the contrary case, the obtained tool isn't equivalent to the original one thus preventing an optimal international comparability of results.

Result's comparability of clinical score's use can't also be achieved without a reliable test. We, therefore, planned to analyze the intra- and inter-observer reliability of the BDCAF Tunisian version.

Mean BDCAF total scores, calculated by observers 1 and 2, were well correlated (r = 0.94, p < 0.0001). Thus, to an excellent inter-observer, we concluded an agreement concerning the overall BDCAF score. This finding was also confirmed for the intra-observer agreement with an excellent correlation of the two BDCAF scores calculated by the same physician at the 48-hour interval (3.23 +/−3 and 3.23 +/−3, r = 0.98, p < 0.0001).

Global BDCAF scores weren't calculated in previous works and the intra- and inter-observers' agreement of the overall score was not tested [6,10,11].

The reliability of BDCAF score was also analyzed by testing the intra- and inter-observers' agreement for each item of the score. Agreement of observers was good for oral and genital ulcers, NPF, arthrlagia, ocular, neuronlogical and large vessels involvement. It was moderate for other parameters (patient well-being the last 28 days, patient well-being the evaluation's day, headache, erythema nodosum, arthritis, nausea/itching). The inter-observers' agreement was fair only for diarrhea and for the clinician's impression of the disease activity.

Intra-observer's agreement was fair only for the patient well-being and for arthritis.

The comparison of our findings with those of Bahakta [6], Neves [10] and Hamuruydan [11] is illustrated at **Tables 3** (inter-observers' agreement) and **4** (intra-observers' agreement). Only inter-observers' agreement was analyzed in the evaluation of the original version [6].

For all versions, best κ coefficients were obtained for oral and genital ulcers.

Oral ulcer is a nearly constant symptom in BS and when associated to genital ulcer, highly suggests this diagnosis. Ulcers, in BD, are the leading reason for seeking healthcare and are often invalidating and so easy to recall even if they were progressing since a month. All these facts could explain the excellent concordance of observers about these symptoms and their imputability to BS. A clinical sign isn't taken into account in BDCAF score unless it's attributable to BS.

As other versions, observer's agreement in our study was very good concerning arthralgias and NPF (**Table 3**). However, it wasn't the case for arthritis. Intra-observer's agreement, for arthritis, was fair, too. Joint inflammatory signs in BD patient can sometimes go unnoticed explaining probably the fair intra- and inter-observers' agreement. However, our result wasn't found by other studies. It is also possible that the Tunisian formulation of the corresponding question, even if validated by clinicians and English language experts, wasn't the most appropriate. Observer's concordance concerning the existence of a potential gastrointestinal (GI) flare, as other studies [6, 10,11], was fair. Two questions, in the BDCAF, explore the GI system, respectively for the upper and lower GI

wasn't good. That wasn't very surprising giving the subjective character of this BDCAF item.

The BDCAF is a clinical tool validated by BD experts. We demonstrated that Tu-BDCAF, as previous versions, was a score easy to calculate and to apply. And even if its reliability wasn't optimal for all the clinical parameters, it remained a relatively objective index for evaluating BD activity in clinical routine. Nevertheless, we think that the absence of a cut-off point clearly defining BD activity can be a limit to its use in daily practice.

5. ACKNOWLEDGEMENTS

REFERENCES

[1]   Behçet, H. (1937) Uber rezidivierende apthose, durch ein virus verursachte geschwure ammund, amauge und an den genitalien. *Dermatol Wochenschr*, **105**, 1152-1157

[2]   Sakane, T., Takeno, M., Suzuki, N. and Inaba, G. (1999) Behçet's disease. *New England Journal of Medicine*, **21**, 1284-1291.
http://dx.doi.org/10.1056/NEJM199910213411707

[3]   Symmons, D.P. (1995) Disease assessment indices: Activity, damage and severity. *Baillieres Clinical Rheumatology*, **9**, 267-285.
http://dx.doi.org/10.1016/S0950-3579(05)80190-6

[4]   Yazici, H., Tüzün, Y. and Pazarli, H. (1984) Influence of age of onset and patient's sex on the prevalence and severity of manifestations of Behçet's syndrome. *Annals of the Rheumatic Diseases*, **43**, 783-789.
http://dx.doi.org/10.1136/ard.43.6.783

[5]   Davatchi, F., Akabaran, M. and Shahram, F. (1991) Iran Behçet's disease dynamic activity measure. *Hung Rheumatology*, **32**, 10-100.

[6]   Bhakta, B.B., Brennan, P., James, T.E., *et al.* (1999) Behçet's disease: Evaluation of a new instrument to measure clinical activity. *Rheumatology*, **38**, 728-723.
http://dx.doi.org/10.1093/rheumatology/38.8.728

[7]   International Study Group for Behçet's Disease (1990) Criteria for diagnosis of Behçet's disease. *Lancet*, **335**, 1078-1080.

[8]   Beaton, D.E., Bombardier, C., Guillemin, F., *et al.* (2000) Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, **25**, 3186-3191.
http://dx.doi.org/10.1097/00007632-200012150-00014

[9]   Bhakta, B., Hamuryudan, V., Brennan, P., Chamberlain, M.A., Barnes, C. and Silman, A.J. (1993) Assessment of disease activity in Behçet's disease. In: Wechsler, B. and Godeau, P., Eds., *Excerpta Medical International Congress Series* 1037, 611, Elsevier Science Publishers BV, Amsterdam, 235-240.

[10] Neves, F.S., Moraes, J.C.B., Kowalski, S.C., *et al.* (2007) Cross-cultural adaptation of the Behçet's Disease Current Activity Form (BDCAF) to Brazilian Portuguese language. *Clinical Rheumatology*, **26**, 1263-1267.
http://dx.doi.org/10.1007/s10067-006-0484-y

[11] Hamuryudan, V., Fresko, J., Direskeneli, H., *et al.* (1999) Evaluation of the Turkish translation of a disease activity form for Behçet's syndrome. *Rheumatology*, **38**, 734-736.
http://dx.doi.org/10.1093/rheumatology/38.8.734

[12] Cheon, J.H., Han, D.S., Park, J.Y., *et al.* (2011) Development, validation, and responsiveness of a novel disease activity index for intestinal behcet's disease. *Inflammatory Bowel Disease*, **17**, 605-613.
http://dx.doi.org/10.1002/ibd.21313