Current Status, Goals, and Research Agenda for Outcome Measures Development in Behçet Syndrome: Report from OMERACT 2014

Gulen Hatemi, **Yesim Ozguler**, **Haner Direskeneli**, **Alfred Mahr**, **Ahmet Gul**, **Virna Levi**, **Sibel Z. Aydin**, **Gonca Mumcu**, **Ozlem Sertel-Berk**, **Randall M. Stevens**, **Hasan Yazici**, and **Peter A. Merkel**

Division of Rheumatology, Department of Internal Medicine, Istanbul University Cerrahpasa Faculty of Medicine, Istanbul; Division of Rheumatology, Marmara University Faculty of Medicine, Istanbul, Turkey; Department of Internal Medicine, University Paris Diderot, Paris, France; Division of Rheumatology, Koc University Faculty of Medicine, Istanbul; Faculty of Health Sciences, Marmara University; Department of Psychology, Istanbul University, Istanbul, Turkey; Celgene Corporation, Warren, New Jersey; Rutgers, Robert Wood Johnson Medical School, Department of Rheumatology, New Brunswick, New Jersey; Division of Rheumatology and Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, Pennsylvania, USA

Abstract

Objective—There is an unmet need for reliable, validated, and widely accepted outcomes and outcome measures for use in clinical trials in Behçet syndrome (BS). Our report summarizes initial steps taken by the Outcome Measures in Rheumatology (OMERACT) vasculitis working group toward developing a core set of outcome measures for BS according to the OMERACT methodology, including the OMERACT Filter 2.0, and discussions during the first meeting of the BS working group held during OMERACT 12 (2014).

Methods—During OMERACT 12, some of the important challenges in developing outcomes for BS were outlined and discussed, and a research agenda was drafted.

Results—Among topics discussed were the advantages and disadvantages of a composite measure for BS that evaluates several organs/organ systems; bringing patients and physicians together for discussions about how to assess disease activity; use of organ-specific measures developed for other diseases; and the inclusion of generic, disease-specific, or organ-specific

Address correspondence to Dr. P.A. Merkel, Section of Rheumatology, University of Pennsylvania, White Building, 5th Floor, 3400 Spruce St., Philadelphia, Pennsylvania 19104, USA. pmerkel@upenn.edu.

G. Hatemi, MD, Associate Professor of Medicine, Division of Rheumatology, Department of Internal Medicine, Istanbul University Cerrahpasa Faculty of Medicine; Y. Ozguler, MD, Fellow in Rheumatology, Division of Rheumatology, Istanbul University Cerrahpasa Faculty of Medicine; H. Direskeneli, MD, Professor of Rheumatology, Division of Rheumatology, Marmara University Faculty of Medicine; A. Mahr, MD, PhD, Professor of Internal Medicine, Department of Internal Medicine, University Paris Diderot; A. Gul, MD, Professor of Rheumatology, Division of Rheumatology, Department of Internal Medicine, Istanbul University; V. Levi, BA, Patient Research Partner, Istanbul; S.Z. Aydin, MD, Associate Professor of Rheumatology, Division of Rheumatology, Koc University Faculty of Medicine; G. Mumcu, DDS, PhD, Professor in Health Management, Faculty of Health Sciences, Marmara University; O. Sertel-Berk, MA, Associate Professor of Psychology, Department of Psychology, Istanbul University; R.M. Stevens, MD, Celgene Corporation, and Clinical Professor of Medicine, Rutgers, Robert Wood Johnson Medical School, Department of Rheumatology; H. Yazici, MD, Professor of Rheumatology, Division of Rheumatology, Department of Internal Medicine, Istanbul University Cerrahpasa Faculty of Medicine; P.A. Merkel, MD, MPH, Professor of Medicine and Epidemiology, Division of Rheumatology and Department of Medicine and Department of Biostatistics and Epidemiology, University of Pennsylvania.

measures. The importance of incorporating patients' perspectives, concerns, and ideas into outcome measure development was emphasized.

Conclusion—The planned research agenda includes conducting a Delphi exercise among physicians from different specialties that are involved in the care of patients with BS and among patients with BS, with the aim of identifying candidate domains and subdomains to be assessed in randomized clinical trials of BS, and candidate items for a composite measure. The ultimate goal of the group is to develop a validated and widely accepted core set of outcomes and outcome measures for use in clinical trials in BS.

Key Indexing Terms

BEHÇET SYNDROME; OUTCOME MEASURES; OUTCOMES

Behçet syndrome (BS) is a form of vasculitis that affects several organs and organ systems including the eyes, arteries, veins, gastrointestinal (GI) system, nervous system, joints, skin, and mucosa. It runs a variable course depending on the patients' age, sex, and disease duration[1]. BS follows a more severe course among young men, with more frequent eye and vascular involvement, and clinical findings usually abate over time[2,3]. This variable disease course prevents the development of a single management strategy, and the treatment has to be individually tailored for each patient. Several clinical trials that address different types of organ involvement have been conducted in BS. However, owing to the diversity and variability in the outcome measures used in these trials, it is difficult to compare research from different centers, combine datasets for additional analysis, or fully incorporate the trial results into healthcare decisions. Thus, there is an unmet need to establish reliable, validated, and widely accepted outcome measures for this complex disorder.

The Outcome Measures in Rheumatology (OMERACT) vasculitis working group has been working to advance outcomes research in BS. Investigators are working together with patient research partners and have begun the process of establishing a research agenda, following the methodology endorsed by OMERACT, with the ultimate goal of creating a core set of data-driven outcome measures for use in clinical trials in BS. Our initial step in this process was to conduct a systematic literature review of outcomes and outcome measures used in trials of BS[4]. Our next step was to perform a survey among experts in BS from different specialties to understand their views regarding the needs for outcomes assessment in BS. We also conducted a number of individual patient interviews with the aim of better understanding patients' perspectives.

In this report we summarize the work we conducted prior to OMERACT 12 (2014), the discussions during the BS special interest group (SIG) held at OMERACT 2014, and our future plans regarding the development of outcomes and outcome measures for use in both clinical research and patient care, and ultimately, produce a core set of outcome measures endorsed by OMERACT for use in clinical trials.

Current Status

To understand the current status of outcome measures in BS and to identify the domains that need to be addressed by outcome measures, we conducted a systematic review of outcome measures used in randomized controlled trials, nonrandomized clinical trials, longitudinal cohorts, case series, biomarker studies, and genetic association studies of BS[4]. The main findings of this systematic review were that few of these measures were properly validated or widely used, and there was a lack of standardized definitions for key concepts such as response, relapse, or remission. We identified 139 outcomes or outcome measures used in a total of 249 manuscripts. These outcomes or outcome measures can be grouped into 3 categories: (1) Behçet-specific outcome measures such as the Behçet's Disease Current Activity Index and the Behçet's Disease Quality of Life Measure; (2) generic outcome measures that have been used for various rheumatologic and nonrheumatologic conditions such as the Medical Outcomes Study Short Form-36 (SF-36); and (3) organ-specific outcome measures. Some of the single organ measures were developed for other diseases, such as the Crohn's Disease Activity Index, or a multiple sclerosis functional compound scale, and were not properly validated for BS. The wide range of approaches used to study BS, as shown by this systematic review, makes it difficult to compare the results of different trials and hampers development of standardized approaches to trial design that would be acceptable to investigators, patients, the biomedical industry, payers, and medical regulatory authorities.

Survey Among Experts

We conducted an Internet-based survey to understand the needs of researchers who actively work in this field, to obtain ideas regarding the current status of outcomes in BS, and to generate ratings of the domains of importance for clinical trials (Figure 1). The survey included 11 questions, and 35 out of 51 experts (69%) responded to the survey. The respondents represented various specialties: 31 rheumatologists, 7 dermatologists, 6 ophthalmologists, 3 gastroenterologists, 3 internists, and 1 dentist from 13 different countries (France, Germany, Iran, Italy, Japan, Korea, Morocco, Portugal, Switzerland, Tunisia, Turkey, the United Kingdom, and the United States).

The survey asked the experts about the endpoints that are relevant to test a drug's efficacy in a clinical trial of BS. Disease activity (100%) and health-related quality of life (97%) were the categories that almost all of the responders agreed on and these were followed by physical function (83%), mortality (74%), disease-related damage (71%), disease severity (66%), fatigue (46%), and overall damage (45%).

The survey also asked about the validity and reliability of 4 different instruments available to assess disease activity in BS: the Behçet's Disease Current Activity Index (BDCAI), the Behçet's Syndrome Activity Scale (BSAS), the Clinical Manifestation Index (CMI), and the Iranian Behçet's Disease Dynamic Activity Measure (IBDDAM)[5,6,7,8]. Less than one-half of the experts agreed that any of these scales are reliable and valid: BDCAI (46%), BSAS (43%), CMI (22%), and IBDDAM (22%).

The majority of respondents (89%) agreed that a new instrument to assess overall disease activity is needed and 97% agreed that such an instrument should be weighted, with different weights for each item, such as oral ulcers, genital ulcers, arthritis, uveitis, and deep vein thrombosis (Figure 2). Seventy-six percent of experts agreed that this new instrument should include patients' input regarding disease activity. The experts did not agree on inclusion of acute-phase reactants in such an instrument (35% agreed, 35% neutral, 30% disagreed).

The experts were also asked to weight potential outcome items on a 10-point scale (0 = no weight, 10 = maximum weight). Receiving the highest scores were neurologic involvement (parenchymal involvement = 9.02 ± 1.77, dural sinus thrombosis = 7.97 ± 2.07); almost all kinds of vascular involvement (pulmonary artery aneurysm = 8.97 ± 1.86, large vessel arteritis = 8.62 ± 2.03, vena cava thrombosis = 8.59 ± 1.88, deep vein thrombosis = 7.75 ± 2.08); uveitis (posterior uveitis = 8.40 ± 1.74, anterior uveitis = 6.27 ± 2.28); and GI involvement (6.97 ± 2.04). The scores for skin and mucosa lesions, arthritis, and fatigue were lower (oral ulcers = 4.97 ± 2.33, genital ulcers = 5.70 ± 2.39, papulopustular lesion = 4.21 ± 1.97, erythema nodosum = 5.13 ± 1.83, arthralgia = 3.16 ± 1.91, arthritis = 5.40 ± 2.05, and fatigue = 3.19 ± 2.05).

The experts were also asked whether organ-specific tools were needed in addition to an overall BS activity instrument. Most of the respondents (92%) agreed on the need for a uveitis activity score followed by a neurologic involvement activity score (82%), vascular activity score (73%), oral ulcer activity score (73%), GI involvement activity score (70%), genital ulcer activity score (59%), and cutaneous involvement activity score (50%; Figure 3).

This survey demonstrated that leaders of the BS research community agree on the need to develop a set of outcome measures that is widely acceptable and properly validated for use in clinical trials in BS.

Patient Interviews

The OMERACT vasculitis working group also conducted pilot interviews with individual patients with BS to understand the effects of their disease on daily activities, physical function, social life, psychological well-being, fears, and concerns regarding the future, as well as to learn about coping strategies. These interviews provided important information regarding patients' perceptions of the disease. Semistructured interviews were conducted with 12 patients with BS (9 men, 3 women, mean age 35 ± 6 yrs) with various types of organ involvement (5 with arthritis, 4 with eye involvement, 4 with vascular involvement, 1 with neurologic involvement, 1 with GI involvement, and 1 with only skin and mucosa involvement). Irrespective of the type of organ involvement, the patients expressed having difficulty working, going to school, and fulfilling household tasks, and experiencing impairment in social life, feelings of inadequacy and dependence on others, fatigue, fear and anxiety about the future, and a depressive mood.

It is crucial to incorporate the concerns and ideas of patients with BS into research plans. This pilot project began a process to collect and analyze qualitative data using standard techniques and based on insights gained from patients with BS. In addition to individual

patient interviews, focus groups may also add to our understanding of the patients' perspectives, priorities, and ideas regarding their disease and help develop validated patient-reported outcome measures.

Challenges in Outcome Measures Development for BS

There is a recognized need for a composite measure to assess disease activity in BS that includes patient input and is weighted. Such an index was another topic discussed during the OMERACT BS SIG. Several challenges exist if such an instrument is to be developed. It is not clear how a weighting scheme would be included. Because both the prognostic importance, as indicated by the physician, and the perceived severity, as indicated by the patient, should be considered in weighting each disease manifestation, it may be difficult to incorporate the 2 concepts into 1 tool for a multisystem condition such as BS. For example, oral ulcers may cause serious disability and impair the quality of life of the patients but are transient in nature and do not threaten organ function. Alternatively, aneurysms, which may be fatal, are often asymptomatic. Another challenge is that some of the most severe manifestations, such as arterial aneurysms or neurologic involvement, are relatively less frequent. If included in a composite index, these infrequent elements would have to be assessed in all patients, and thus, in a large proportion, unnecessarily. Further, composite measures that identify the most common findings may ignore or undervalue the most serious but less common manifestations of BS. However, if these are not included in the composite measure, then major factors useful for assessing disease activity among the most severely affected patients would be missed. Another potential difficulty with a composite measure is that sensitivity to change may be hard to attain in short-term trials because many vascular, neurologic, and GI manifestations tend to recur or have their maximal effect over months or years. The inability to identify the difference by such a composite measure may cause underrating of the efficacy of newly developed therapies.

Assessment of specific organs is another challenge in BS. Previous data have shown that a drug that is effective for one type of BS manifestation may be ineffective for another[5]. Most, if not all, randomized trials conducted in BS to date have been designed to determine efficacy of an intervention for involvement of a specific organ. The OMERACT group discussed whether organ-specific widely used outcome measures developed and validated for other diseases should be used for the assessment of that organ/organ system in BS. Some examples of such measures include the Crohn's Disease Activity Index[9], the Multiple Sclerosis Functional Compound Scale[10], and the standardization of uveitis nomenclature criteria. Some researchers are content with using these instruments in BS, while others express their concern regarding the shortcomings of these instruments due to the unique features of BS, because these tools have not been properly validated for BS. Further work is required to determine whether these measures are fit for use in trials of BS, or whether it will be necessary and feasible to develop and validate new single-organ measures specific for BS.

Patient-reported outcome measures (PRO) are critical to characterize the life impact of BS. There is 1 PRO that was developed for assessing quality of life in BS[6] and another for assessing disease activity[7]. However, the validity of these tools needs to be determined, when measured among patients with different types of organ involvement and patients from

different geographical areas. Data from qualitative research in patients with BS will help us understand whether current PRO adequately represent the burden from patients' perspectives.

Application of OMERACT Filter 2.0

We aim to apply OMERACT Filter 2.0 during the process of developing a core set of outcome measures for clinical trials of BS. Four core areas have been identified in Filter 2.0: "death," "life impact," "pathophysiological manifestations," and "resource use"[8]. There may be challenges to the application of Filter 2.0 for a multisystem condition such as BS. However, the proposed methodology for Filter 2.0 enables researchers to modify the domains in each area according to their needs. An example of such a modification may be the assessment of "organ function," one of the proposed domains, by using different instruments in different trials depending on the type of involvement being studied. This way, organ function could be assessed in all trials of BS, but with different tools for eye involvement than for oral ulcers or for arthritis.

One of the important domains within "life impact" is quality of life. Further research is needed to determine whether a generic instrument such as the SF-36, a Behçet-specific instrument such as Behçet's Disease Quality of Life scale[11], or an organ-specific instrument such as the oral health-related quality of life questionnaire[12] would perform well in trials of BS.

Future Plans and Research Agenda

A consensus meeting on outcome measures during the Behçet's Syndrome Conference was held in September 2014. The conference included experts in BS working in different fields including rheumatology, dermatology, ophthalmology, and neurology, from Europe, Asia, and North America, as well as patient partners and methodologists with expertise in outcome measure development.

We are also planning to conduct a Delphi exercise among physicians from different specialties experienced in the care of patients with BS and among patients with BS, to identify candidate domains and subdomains that should be assessed in BS. This Delphi exercise will also help to identify candidate items for a possible composite measure.

The qualitative work with patients will also continue with interviews and focus groups to complete the initial process of understanding the burden of disease in BS from the patients' perspective.

Finally, a randomized controlled drug trial in BS in which several outcome measures were used was recently completed. We are now planning to analyze the results of this trial with regard to the responsiveness and validity of these outcome measures.

Our goal is to have a provisional draft core set of domains and instruments and an associated next-stage research agenda ready for the OMERACT 13 meeting in 2016.

Acknowledgments

REFERENCES

1. Kural-Seyahi E, Fresko I, Seyahi N, Ozyazgan Y, Mat C, Hamuryudan V, et al. The long-term mortality and morbidity of Behçet syndrome: a 2-decade outcome survey of 387 patients followed at a dedicated center. Medicine. 2003; 82:60–76. [PubMed: 12544711]

2. Yazici H, Tüzün Y, Pazarli H, Yurdakul S, Ozyazgan Y, Ozdogan 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; 43:783–789. [PubMed: 6524980]

3. Hamuryudan V, Hatemi G, Tascilar K, Sut N, Ozyazgan Y, Seyahi E, et al. Prognosis of Behçet's syndrome among men with mucocutaneous involvement at disease onset: long-term outcome of patients enrolled in a controlled trial. Rheumatology. 2010; 49:173–177. [PubMed: 19995858]

4. Hatemi G, Merkel PA, Hamuryudan V, Boers M, Direskeneli H, Aydin SZ, et al. Outcome measures used in clinical trials for Behçet syndrome: a systematic review. J Rheumatol. 2014; 41:599–612. [PubMed: 24488418]

5. Yazici H, Ugurlu S, Seyahi E. Behçet syndrome: is it one condition? Clin Rev Allergy Immunol. 2012; 43:275–280. [PubMed: 22674015]

6. Gilworth G, Chamberlain MA, Bhakta B, Haskard D, Silman A, Tennant A. Development of the BD-QoL: a quality of life measure specific to Behçet's disease. J Rheumatol. 2004; 31:931–937. [PubMed: 15124253]

7. Forbess C, Swearingen C, Yazici Y. Behçet's Syndrome Activity Score (BSAS): a new disease activity assessment tool, composed of patient-derived measures only, is strongly correlated with the Behçet's Disease Current Activity Form (BDCAF) [abstract]. Arthritis Rheum. 2008; 58(Suppl 9):S854.

8. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino MA, et al. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. J Clin Epidemiol. 2014; 67:745–753. [PubMed: 24582946]

9. Best WR, Becktel JM, Singleton JW, Kern F Jr. Development of a Crohn's disease activity index. National Cooperative Crohn's Disease Study. Gastroenterology. 1976; 70:439–444. [PubMed: 1248701]

10. Cutter GR, Baier ML, Rudick RA, Cookfair DL, Fischer JS, Petkau J, et al. Development of a multiple sclerosis functional composite as a clinical trial outcome measure. Brain. 1999; 122(Pt 5): 871–882. [PubMed: 10355672]

11. Gilworth G, Chamberlain MA, Bhakta B, Haskard D, Silman A, Tennant A. Development of the BD-QoL: a quality of life measure specific to Behçet's disease. J Rheumatol. 2004; 31:931–937. [PubMed: 15124253]

12. Locker D, Slade G. Oral health and the quality of life among older adults: the oral health impact profile. J Can Dent Assoc. 1993; 59:830–833. 837–838, 844. [PubMed: 8221283]

1. What is your full name?
   (e.g. John Q. Smith)

2. What is the name of your institution or company?

3. What is your primary work setting?

   Academic Hospital ○

   Private Practice ○

   Community Hospital ○

   Other ○

4. In a clinical trial on Behçet's syndrome, which endpoints are relevant to test a drug's efficacy?

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Don't Know
Disease activity	○	○	○	○	○	○
Disease severity	○	○	○	○	○	○
Physical Function	○	○	○	○	○	○
Disease-Related Damage	○	○	○	○	○	○
All Damage	○	○	○	○	○	○
Quality of Life	○	○	○	○	○	○
Fatigue	○	○	○	○	○	○
Mortality	○	○	○	○	○	○
Other	○	○	○	○	○	○

If you suggest any other tools please specify what these are in the text box.

5. Do you agree that the following instruments are valid and reliable to measure activity in clinical trials for Behçet's syndrome?

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Don't Know
Behçet Disease Current Activity Index	○	○	○	○	○	○
Behçet's Syndrome Activity Scale	○	○	○	○	○	○
Clinical Manifestations Index	○	○	○	○	○	○
Iranian BD Dynamic Activity Measure	○	○	○	○	○	○

6. Do you think it is pertinent to have a measure of overall disease activity in Behçet's syndrome?

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Don't Know
○	○	○	○	○	○

7. If we were to develop a new overall Behçet's syndrome activity measure, should this measure include patient's input regarding their disease activity?

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Don't Know
○	○	○	○	○	○

8. Should this measure include acute phase reactants such as erythrocyte sedimentation rate and CRP?

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Don't Know
○	○	○	○	○	○

9. Is weighting of each item (different scores for oral ulcer, genital ulcer, arthritis attack, uveitis attack, deep vein thrombosis etc.) important for such an instrument?

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Don't Know
○	○	○	○	○	○

10. Please rate each of the following items on a scale of 0 to 10 to reflect the relative weight you would give the item in a composite disease activity score (0=no weight; 10=maximum weight).

   Oral ulcers
   Genital ulcers
   Papulopustular lesions
   Folliculitis
   Erythema nodosum
   Arthralgia
   Arthritis
   Anterior Uveitis
   Posterior Uveitis
   Superficial thrombophlebitis
   Deep vein thrombosis
   Vena cava thrombosis
   Pulmonary artery aneurisms
   Large vessel arteritis
   Parenchymal CNS involvement
   Aseptic meningitis
   Headache
   Dural sinüs thrombosis
   Gastrointestinal involvement
   Orchitis/epididymitis
   Fatigue
   Other

   If you suggest any other tools please specify what these are in the text box.

11. Apart from an overall Behçet's syndrome activity tool, which organ-specific activity tools do you deem necessary to evaluate the outcome of Behçet's disease?

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Don't Know
Oral ulcer disease activity score	○	○	○	○	○	○
Genital ulcer disease activity score	○	○	○	○	○	○
Cutaneous disease activity score	○	○	○	○	○	○
Uveitis disease activity score	○	○	○	○	○	○
CNS disease activity score	○	○	○	○	○	○
Vascular activity score	○	○	○	○	○	○
Gastrointestinal activity score	○	○	○	○	○	○

If you suggest any other tools please specify what these are in the text box.

12. Please enter here any comments you may have regarding the issues covered in this survey :
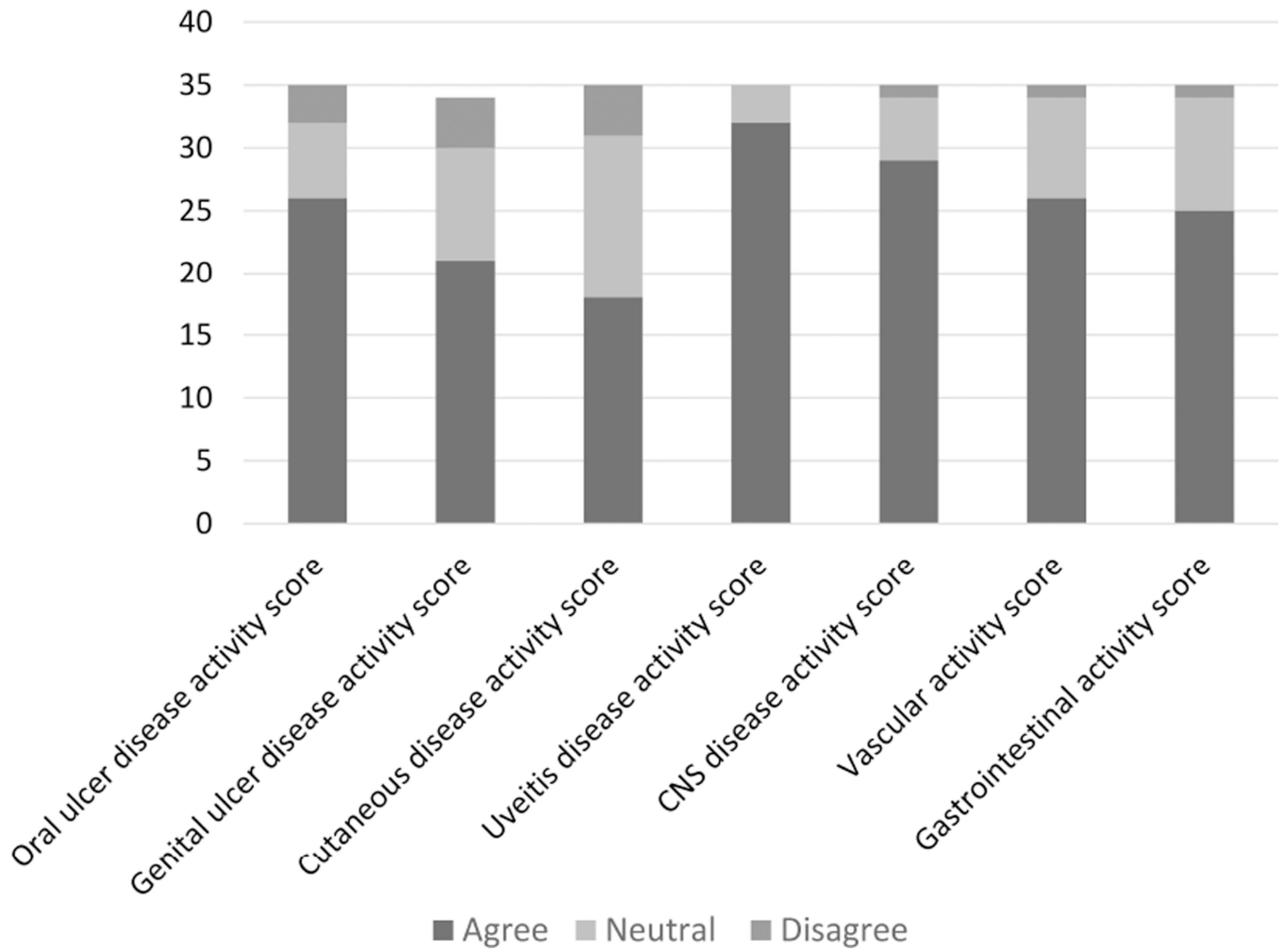
Figure 1.
Questions from the survey of experts in Behçet syndrome.
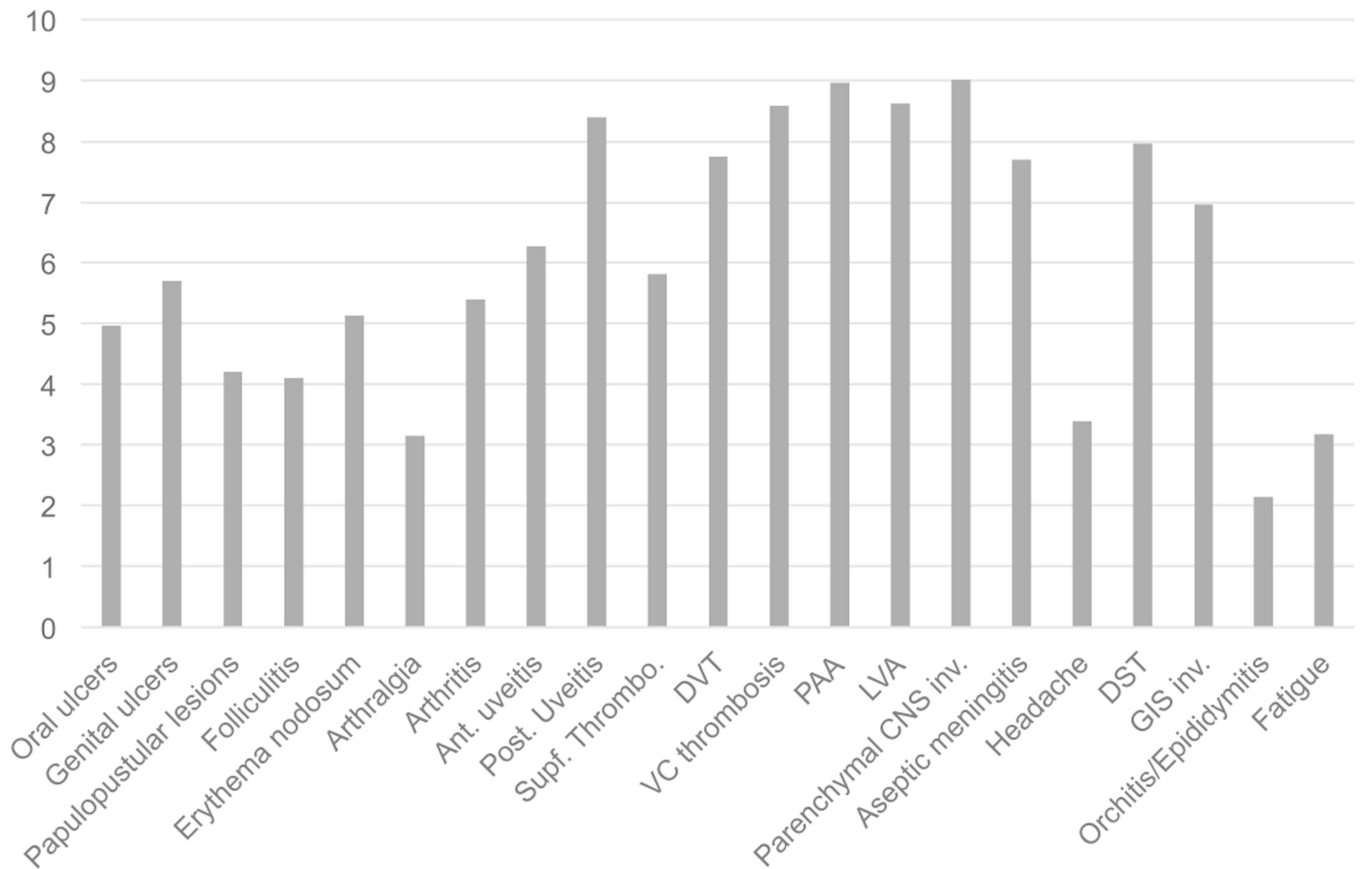
Figure 2.
Experts' ratings for the relative weight of each item in a composite disease activity score.
Ant.: anterior; Post.: posterior; Supf.: superficial; DVT: deep vein thrombosis; VC: vena
cava; PAA: pulmonary artery aneurysm; LVA: large vessel arteritis; CNS: central nervous
system; inv.: involvement; DST: dural sinus thrombosis; GIS: gastrointestinal.

Figure 3.
Experts' ratings for the necessity of each organ-specific activity tool. CNS: central nervous system.