

Psychometric Evaluation of a New Patient-Reported Outcome (PRO) Symptom Diary for Varicose Veins: VVSymQ[®] Instrument

David D. I. Wright¹ · Jean Paty² · Diane M. Turner-Bowker² · Andrew Bradbury³

Published online: 25 March 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract

Objective To evaluate the psychometric properties of the VVSymQ[®] instrument, a new 5-item patient-reported outcome (PRO) measure for symptoms of varicose veins.

Method The VVSymQ[®] electronic daily diary was administered to outpatients who received routine treatment for varicose veins ($N = 40$). Compliance with diary administration and item score variability, reliability, construct validity, sensitivity to change, and clinically meaningful change were evaluated.

Results Patients completed >97 % of scheduled diary assessments (at screening, baseline, and week 8). The

VVSymQ[®] instrument captured patients' pre-treatment symptoms (all VVSymQ[®] symptoms were endorsed by ≥ 75 % of patients at baseline), and the change post-treatment (mean change in score -6.1), with a large Cohen effect size (1.6). Test–retest reliability was high (intraclass correlation coefficient 0.96); internal consistency was good (Cronbach's $\alpha \geq 0.76$; baseline, week 8). VVSymQ[®] scores were more strongly associated with PRO scores that reflect symptoms and symptom impact (the Venous Insufficiency Epidemiological and Economic Study—Quality of Life/Symptoms [VEINES-QOL/Sym] instrument and the Chronic Venous Insufficiency Quality-of-Life Questionnaire [CIVIQ-20]) than with PRO scores that reflect appearance (the Patient Self-Assessment of Appearance of Visible Varicose Veins [PA-V³]) or clinician-reported outcome scores (the Clinical–Etiology–Anatomy–Pathophysiology [CEAP] Classification of Venous Disorders and Venous Clinical Severity Score [VCSS]), demonstrating construct validity. Patients reporting that symptoms were “moderately” or “much improved” on the Patient Global Impression of Change (PGIC) anchor (i.e., >97 % of patients) had mean improvements of -6.3 VVSymQ[®] points, while a cumulative distribution curve showed that 50 % of patients improved by ≥ -5.8 points; thus, a score change of approximately -6 demonstrated a clinically meaningful change in this study. The clinically meaningful change in the VVSymQ[®] score was greater in patients with a greater baseline VVSymQ[®] symptom burden, and the VVSymQ[®] instrument captured clinically meaningful treatment benefit even in patients with a low baseline symptom burden.

Conclusion The 5-item VVSymQ[®] instrument is a brief, psychometrically sound, useful tool for evaluating patient-reported varicose veins symptoms.

✉ David D. I. Wright
david.wright@btgplc.com

Jean Paty
Jean.Paty@quintiles.com

¹ BTG International, Ltd., Lakeview, Riverside Way, Watchmoor Park, Camberley, Surrey GU15 3YL, UK

² ERT, Pittsburgh, PA, USA

³ University of Birmingham, Heart of England NHS Foundation Trust, Birmingham, UK

Key Points for Decision Makers

Until now, there has been no patient-reported outcomes (PRO) instrument for evaluating symptoms alone in patients with varicose veins that has undergone development according to the US Food and Drug Administration PRO Guidance.

This psychometric evaluation of the 5-item VVSymQ[®] electronic daily diary in patients who underwent treatment for varicose veins showed that the instrument was easy to use by patients, is reliable and valid, and captured the change in symptoms after treatment, with a very large effect size.

The VVSymQ[®] instrument is a psychometrically sound, useful tool for evaluating patient-reported symptoms of varicose veins. The instrument may be useful for capturing treatment benefit and monitoring the symptom experience of patients over time in clinical research and practice.

1 Background

Varicose veins are extremely common, affecting up to 73 % of women and up to 56 % of men [1]. Varicose veins and the associated chronic venous insufficiency have a substantial impact on patient quality of life (QOL) [2, 3] and are among the most common vascular conditions requiring specialist treatment [4]. Signs and symptoms of varicose veins are important to patients [Paty et al. Content Validity for the VVSymQ[®] Instrument: A New Patient Reported Outcome Measure for the Assessment of Varicose Veins Symptoms (Patient, manuscript in review, 2016), 5] and are markers of treatment benefit. Treatments for varicose veins can benefit patients by improving the appearance of leg veins and by reducing symptoms and symptom impact on health-related QOL. Since symptoms (e.g., pain, burning, swelling) are not observable by clinicians, they are best measured by querying patients directly, using patient-reported outcome (PRO) instruments [6].

Various PRO instruments have been used to assess patients who have varicose veins, including generic PRO measures (e.g., the SF-36 Health Survey), which do not address venous-specific symptoms and impacts, and condition-specific instruments, such as the Aberdeen Varicose Veins Questionnaire (AVVQ) [7], the Venous Insufficiency Epidemiological and Economic Study—Quality of Life/Symptoms (VEINES-QOL/Sym) instrument [8], the Specific Quality of Life and Outcome Response—Venous Questionnaire (SQOR-V) [9], and the Chronic Venous

Insufficiency Quality-of-Life Questionnaire (CIVIQ-20) [10]. Although they are widely used, none of these existing measures focus solely on varicose vein symptom assessment, and they have not followed best practices for instrument development and validation (e.g., see the US Food and Drug Administration [FDA] PRO Guidance [6]), having been developed prior to publication of the FDA PRO Guidance.

A new electronic PRO daily symptom diary, the VVSymQ^{®1} instrument, was developed to assess the key symptoms of superficial venous incompetence of the great saphenous vein system that are important and relevant to patients, including heaviness, achiness, swelling, throbbing, and itching (Paty et al. Content validity for the VVSymQ instrument [Patient, manuscript in review, 2016]) [5]. Development and validation of the VVSymQ[®] instrument followed established instrument development guidelines [6, 11]. Previous work on the VVSymQ[®] instrument established the content validity of the measure and demonstrated patient understanding of the instructions, items, and response options [5]. This article reports results from a psychometric evaluation of the VVSymQ[®] instrument, using data from a single-center study of patients with varicose veins.

2 Objectives

The specific objectives of this research were to (1) examine whether the 5-item VVSymQ[®] instrument appropriately reflects patients' experience of varicose vein symptoms; (2) evaluate the quantitative psychometric properties (item distributions, reliability, validity, ability to detect change) of the VVSymQ[®] instrument; (3) identify responder definitions that can be used to determine if a patient has experienced a clinically meaningful change on the instrument; and (4) evaluate the administrative feasibility of the instrument as an electronic daily diary.

3 Methods

3.1 Study Design

The psychometric performance of the VVSymQ[®] instrument was evaluated in a single-center study that was designed to evaluate its measurement properties. The evaluation was conducted in the context of treating patients with the site's standard of care for varicose veins. The single-center site was selected for this study on the basis of its experience with and access to the targeted sample. The screening population included all patients available to

¹ VVSymQ[®] is a registered trademark of Provensis Ltd., a BTG International group company.

Table 1 Schedule of study assessments

Action	Screening and baseline visit 1, day -28 (at least 14 days before day 0)	Treatment visit 2, day 0	Post-treatment visit 3, week 8 (days 56–63)
Informed consent, calculation of score derived from question 1 on the Modified VEINES-QOL/Sym, demographics, inclusion and exclusion criteria, CEAP assessment, training of patient to use electronic diary	✓		
Patient self-assessments			
Daily diary (daily assessment of varicose vein symptoms and activity in the electronic diary)	✓ ^a		✓ ^a
Modified VEINES-QOL/Sym	✓	✓ ^b	✓ ^c
CIVIQ-20	✓	✓ ^b	✓ ^c
PA-V ³	✓	✓ ^b	✓ ^c
PGIC in symptoms			✓ ^c
Trained health care professional assessment			
VCSS	✓	✓ ^b	✓ ^c
Return electronic diary			✓
Varicose vein treatment		✓	

CEAP Clinical–Etiology–Anatomy–Pathophysiology Classification of Venous Disorders, CIVIQ-20 Chronic Venous Disease Quality-of-Life Questionnaire, PA-V³ Patient Self-Assessment of Appearance of Visible Varicose Veins, PGIC Patient Global Impression of Change, VCSS Venous Clinical Severity Score, VEINES-QOL/Sym Venous Insufficiency Epidemiological and Economic Study—Quality of Life/Symptoms

^a Patients used the daily diary for at least 14 days immediately prior to visit 1 and for 10 days immediately prior to visit 3, with the last day being the day before the visit

^b Had to have been completed before treatment

^c Had to have been completed after post-treatment electronic diary collection

participating investigators who had a clinical diagnosis of great saphenous vein incompetence (varicose veins) and were scheduled for treatment with ultrasound-guided foam sclerotherapy. Patients received compensation of up to £150 for their time and travel. Patients came to the clinic on three occasions (Table 1). At each visit, patients completed three PRO instrument questionnaires (the Modified VEINES-QOL/Sym, CIVIQ-20, and Patient Self-Assessment of Appearance of Visible Varicose Veins [PA-V³]), and clinicians completed a clinician-reported outcome (ClinRO) instrument (the Venous Clinical Severity Score [VCSS]). Clinicians also completed the Clinical–Etiology–Anatomy–Pathophysiology (CEAP) Classification of Venous Disorders at visit 1. Data from these measures were used to evaluate the construct validity of the VVSymQ[®] instrument. Patients completed the VVSymQ[®] instrument as part of a larger electronic daily diary (evening report) for approximately 14 days between visits 1 and 2 (where week 1 of electronic daily diary use was considered the screening period and week 2 was considered the baseline period). At visit 2, patients received treatment for their varicose veins. Immediately prior to visit 3 (8 weeks after treatment), patients completed the evening report, using the electronic daily diary, for approximately 10 days. This was considered the post-treatment period.

3.2 Patient Population

Adult outpatients (aged ≥ 18 years) with physician-diagnosed saphenofemoral junction incompetence who were scheduled to receive treatment in the UK (ultrasound-guided sodium tetradecyl sulfate foam sclerotherapy) for great saphenous vein incompetence (varicose veins) in one leg were eligible to participate in the study. Patients were required to be symptomatic, with a screening symptom score of ≥ 7 as derived from question 1 on the Modified VEINES-QOL/Sym instrument. Patients were excluded from the study if they were generally unable to complete an electronic daily diary in accordance with the protocol; had participated in any other investigational pharmaceutical product or device study within 3 months prior to visit 1; or had a current venous leg ulcer in either leg.

3.3 Measures

3.3.1 Patient Instruments

3.3.1.1 VVSymQ[®] Instrument The VVSymQ[®] is a 5-item PRO instrument that measures heaviness, achiness, swelling, throbbing, and itching associated with varicose veins

[5, 11]. All items assess symptom duration and use a 6-point Likert-type response scale ranging from “none of the time” (score 0 points) to “all of the time” (score 5 points). The VVSymQ[®] instrument yields a daily sum score that ranges from 0 to 25, with higher scores indicating greater symptom duration. The VVSymQ[®] score is the average of the daily scores over 7 days.

3.3.1.2 Daily Diary for Varicose Vein Symptoms and Activity The VVSymQ[®] instrument was administered as part of a larger electronic questionnaire, the Daily Diary for Varicose Veins—Symptoms and Activity (hereafter referred to as the “daily diary”) (see Fig. 1) [5, 11]. The daily diary alerted each evening to remind the patient to complete the items. In addition to the 5-item VVSymQ[®] instrument, patients completed 15 other items as part of the evening report (a total of 20 items). This included four items to assess additional varicose vein symptoms (heating/burning sensation, tingling sensation, night cramps, and restless legs) on a 6-point duration-based response scale; nine items to assess all of the varicose vein symptoms on a 0–10 intensity-based numeric rating scale; and two items to assess daily activity (overall activity level, time spent sitting or standing without moving around), using a 6-point response scale.

The following scores were calculated from the electronic daily diary data:



Fig. 1 Sample screenshot from the electronic Daily Diary for Varicose Veins—Symptoms and Activity

5-item duration: sum of items 1–5 (range 0–25) [VVSymQ[®] instrument].
 7-item duration: sum of items 1–7 (range 0–35).
 9-item duration: sum of items 1–9 (range 0–45).
 5-item intensity: sum of items 1–5 (range 0–50).
 7-item intensity: sum of items 1–7 (range 0–70).
 9-item intensity: sum of items 1–9 (range 0–90).

3.3.1.3 CIVIQ-20 The Chronic Venous Insufficiency Quality-of-Life Questionnaire (CIVIQ-20) is a 20-item PRO instrument that assesses symptoms, actions and activities, and feelings over the past 4 weeks [10]. The CIVIQ-20 uses a 5-point Likert-type response scale ranging from 0 (no trouble, minimal problem) to 5 (greatest intensity or trouble). The overall score ranges from 0 to 100, with high scores indicating greater discomfort or trouble. The CIVIQ-20 was completed by patients at visits 1, 2, and 3, and was used in tests of the construct validity of the VVSymQ[®] instrument.

3.3.1.4 Modified VEINES-QOL/Sym The Modified Venous Insufficiency Epidemiological and Economic Study—Quality of Life/Symptoms (Modified VEINES-QOL/Sym) is a 26-item PRO instrument that assesses QOL impact and symptoms in individuals who have varicose veins [8, 12]. Responses are rated on 2- to 7-point response scales of intensity, frequency, or agreement. For this study, the original VEINES-QOL/Sym recall was modified from a 1-month recall period to a 1-week recall period. A 1-week recall period was used because the authors were concerned that patients may not be able to reliably recall varicose vein symptom experience over a month-long period. Recent research has indicated that a decision regarding recall for PROs is dependent on the anticipated attributes of a disease [13], and the FDA PRO Guidance recommends “items with short recall periods or items that ask patients to describe their current or recent state” for symptom reporting [6].

The instrument was completed by patients at visits 1, 2, and 3. The QOL domain scores range from 0 to 100 (higher scores indicate better QOL). The VEINES-QOL score was used in tests of the construct validity of the VVSymQ[®] instrument.

3.3.1.5 PA-V³ The Patient Self-Assessment of Appearance of Visible Varicose Veins (PA-V³) is a single-item PRO instrument that assesses the appearance of varicose veins in each leg separately, using a 5-point Likert-type response scale ranging from “not at all noticeable” to “extremely noticeable” [14]. Before treatment, the patient rated the appearance of both legs. After treatment, the patient assessed only the treated leg. In the analysis of

measurement properties, the ratings only for the treated leg were analyzed. The PA-V³ was completed by patients at visits 1, 2, and 3, and was used in tests of the construct validity of the VVSymQ[®] instrument.

3.3.1.6 PGIC The Patient Global Impression of Change (PGIC), a single-item PRO with a 7-point Likert-type response scale ranging from “much worse” to “much improved,” was administered to patients at visit 3 to assess their overall impression of change in varicose vein symptoms in the treated leg over time. The PGIC was used in sensitivity analyses and to establish a responder definition for the VVSymQ[®] instrument.

3.3.2 Clinician Instruments

3.3.2.1 CEAP Classification of Venous Disorders The Clinical–Etiology–Anatomy–Pathophysiology (CEAP) Classification of Venous Disorders is a ClinRO instrument used to characterize the form and severity of chronic venous disease. The clinical portion of the CEAP has seven grades of severity, with higher grades generally indicating worse severity. Only patients with a CEAP clinical grade of C2 (varicose veins) through C5 (skin changes with healed ulceration) were eligible for inclusion in this study. The CEAP was completed by clinicians during visit 1 to classify varicose vein severity and was used in tests of the construct validity of the VVSymQ[®] instrument.

3.3.2.2 VCSS The Venous Clinical Severity Score (VCSS) is a ClinRO instrument that includes a clinician-administered, single-item PRO pain assessment and clinician ratings of the patient’s superficial veins, venous edema, skin pigmentation, inflammation, induration, active ulcer number, active ulcer duration, active ulcer size, and use of compression therapy. The VCSS yields an overall score ranging from 0 to 30, with higher scores representing greater severity. The VCSS was completed by clinicians at visits 1, 2, and 3, and was used in tests of the construct validity of the VVSymQ[®] instrument.

3.4 Procedure

Patients received training to complete the daily diary (VVSymQ[®] instrument and 15 other items) on a hand-held electronic device, and they completed the other PRO measures (the CIVIQ-20, Modified VEINES-QOL/Sym, PA-V³, and PGIC) on paper. Clinicians completed the CEAP and VCSS. The schedule of assessments is shown in Table 1. Paper measures were transcribed onto electronic case report forms.

3.5 Ethics Statement

The study protocol was approved by the Black Country Research Ethics Committee (REC), England, and all patients provided written informed consent prior to participating in the study.

3.6 Handling of Data

The screening population included all patients who signed informed consent, and the study population included all patients who provided data for at least one post-screening assessment.

The daily VVSymQ[®] scores provided by the patient during the week preceding a scheduled study visit were averaged and used as the representative score for the patient’s visit. Four completed days (consecutive or non-consecutive) were necessary to derive a VVSymQ[®] score, otherwise the data were considered missing for that week. The 7-day average VVSymQ[®] score (for baseline and for week 8) was the average of the summed score (including imputed scores; see below) across 7 days. The same approach was taken for the other scores based on daily diary data.

For missing VVSymQ[®] data, if at least 4 but not all 7 days had all of the comprising items scored, missing items from any given day were imputed on the basis of the average of the non-missing scores for that item across the 7 days. The electronic daily diary did not allow skipping of any of the 20 questions.

For missing Modified VEINES-QOL/Sym data, if more than 50 % of the items on a particular subscale were missing, the score for that subscale was set to missing, otherwise the subscale score was calculated as the sum of the scores of the items present, multiplied by the ratio of the maximum possible number of all items to the number of items present.

For the CIVIQ-20, mean imputation methods for missing items occurred if 50 % or more of the subscale items were present.

3.7 Analyses

The psychometric performance of the VVSymQ[®] instrument was evaluated using standard analytic procedures and measurement review criteria developed by the Scientific Advisory Committee of the Medical Outcomes Trust [15] and further elaborated by the US FDA [6, 16].

3.7.1 Administrative Feasibility

The administrative feasibility of electronic daily diary data collection was analyzed in terms of the completeness of the

data obtained (compliance) based on the actual number of completed assessments compared with the expected number of assessments. Three periods were identified for analysis of the electronic daily diary data: screening (days -14 to -8); baseline (days -7 to -1 , with day 0 being the treatment day); and post-treatment (the seven calendar days preceding visit 3, or week 8). For each patient, an individual compliance rate (percentage) was calculated (number of completed diary entries relative to total number of all scheduled entries in the particular reporting period), and then the mean percentage across all patients in that reporting period was calculated to obtain the compliance rate.

3.7.2 Scoring Evaluation

To determine if a 5-item duration-based VVSymQ[®] score appropriately reflects the patients' experience of symptoms, given that nine symptoms from the broader daily diary were assessed on both duration and intensity response scales, Pearson r correlations were examined between the VVSymQ[®] score and scores for the 7- and 9-item duration-based symptom instrument versions, and also between the VVSymQ[®] score and scores for the 5-, 7-, and 9-item intensity-based symptom instrument versions at baseline, and for the changes in scores from baseline to week 8.

3.7.3 Item Distribution and Descriptive Statistics

VVSymQ[®] item scores at baseline and post-treatment (week 8) were assessed through an examination of frequency and descriptive statistics. Item floor or ceiling effects were concluded for baseline data if $>50\%$ of patients reported no symptom duration ("none of the time") or the greatest symptom duration ("all of the time"). Descriptive statistics (i.e., the mean, standard deviation [SD], median, minimum, and maximum) were calculated for baseline and week 8 VVSymQ[®] scores, and for changes in VVSymQ[®] scores at week 8.

3.7.4 Reliability

The extent of agreement between scores obtained on different days and during different periods was evaluated in terms of distributions of difference scores between assessments for different days, and using both nonparametric measures (kappa values) and parametric measures (intraclass correlation coefficients [ICCs]). ICCs were used to assess whether the VVSymQ[®] instrument yielded reproducible scores during a stable period (i.e., from screening to baseline, when only a minimal change or no change in the condition was expected). Values of 0.75 or higher for ICCs are generally considered satisfactory (see, for example, Portney and Watkins [17], Fleiss [18], and

Gwaltney et al. [19]). Internal consistency of the data was evaluated using Cronbach's alpha; values ≥ 0.7 are considered good to excellent (>0.9).

3.7.5 Construct Validity

Construct validity (how well an instrument measures the constructs it was designed to measure) was evaluated through an examination of Pearson r correlation coefficients for the relationship between VVSymQ[®] scores and scores on the VEINES-QOL, CIVIQ-20, PA-V³, CEAP, and VCSS (for baseline and week 8, and for changes from baseline to week 8). Pearson r correlation coefficients ≥ 0.70 indicate a strong relationship between variables [20]. The VVSymQ[®] scores were expected to be more strongly associated with the PRO instrument scores that assess symptoms or symptom impact (the VEINES-QOL and CIVIQ-20) than with the PRO instrument score that reflects appearance (the PA-V³) or the ClinRO instrument scores (the CEAP and VCSS). Higher correlation coefficients support convergence between measures (i.e., that they are measuring similar constructs), and lower correlation coefficients indicate divergence between measures (i.e., that they are measuring dissimilar constructs).

3.7.6 Ability to Detect Change

Central to understanding the performance of the VVSymQ[®] instrument is the concept that changes in the clinical condition are reflected in instrument scores. The mean changes in VVSymQ[®] score from baseline to week 8 was evaluated with Cohen effect size statistics (i.e., the change in the mean value from baseline to week 8, expressed as a proportion of the SD of the baseline [pre-treatment] score). Effect size thresholds of 0.2, 0.5, and 0.8 or greater are interpreted as small, moderate, and large effect sizes, respectively [21].

3.7.7 Clinically Meaningful Change

In the PRO context, clinically meaningful change reflects the point at which a change in a score can be interpreted as clinically important from the perspective of the patient, and is used to understand test scores beyond what is provided for by "statistically significant" results. An anchor-based approach was used to assess clinically meaningful changes in VVSymQ[®] scores [22]. Anchor-based methods use an external criterion to categorize patients into groups, each reflecting predetermined change groupings (e.g., no change, large positive change, large negative change). VVSymQ[®] scores (means and SDs) were computed for each level of change reported on the PGIC items. The minimal important difference (MID) is the smallest

difference or change in score (from baseline to week 8) that is important to the patient (e.g., those reporting “a little improved” on the PGIC).

A cumulative distribution curve was also produced for the VVSymQ[®] instrument. Cumulative distribution curves display information on what type of responses contributed to the mean group response and provide more useful data than a simple point estimate of the difference between group mean changes [16], allowing the percentage of responders in the group to be determined at all possible response thresholds (e.g., 25 % improvement, 3-point improvement).

4 Results

Forty-two patients were screened for the study; two patients scored <7 on question 1 of the Modified VEINES-QOL/Sym instrument and, thus, were not enrolled. Forty patients were enrolled, and all of them completed the study. No patients withdrew from the study. Demographic characteristics are shown in Table 2.

4.1 Administrative Feasibility of Electronic Daily Diary

Compliance with completion of the electronic daily diary was extremely high, with over 97 % of scheduled entries

Table 2 Sample demographic characteristics of the study population ($N = 40$)

Characteristic	Values
Age, years	
Mean (SD)	49.7 (13.9)
Median (range)	46.0 (27–75)
Sex, n (%)	
Male	15 (37.5)
Female	25 (62.5)
Race, n (%)	
White	37 (92.5)
Asian	3 (7.5)
CEAP grade, n (%)	
C2	20 (50)
C3	5 (12.5)
C4	13 (32.5)
C5	2 (5)
Baseline VCSS ^a	
Mean (SD)	6.3 (2.42)
Median (range)	5.0 (3–12)

CEAP Clinical–Etiology–Anatomy–Pathophysiology Classification of Venous Disorders, SD standard deviation, VCSS Venous Clinical Severity Score

^a $n = 39$

being completed for each of the three assessment periods (screening, baseline, and post-treatment), and with only 6 of 274 entry days missed.

4.2 Scoring Evaluation

The correlations between the 5-item duration-based VVSymQ[®] score and the other alternative scores (the 7- and 9-item symptom duration scores and the 5-, 7-, and 9-item intensity scores) ranged from 0.91 to 0.98 (at baseline). Also, correlations of change from baseline to week 8 for the 5-item duration-based VVSymQ[®] score and other alternative scores ranged from 0.89 to 0.98. This high level of concordance, in which all configurations of scores were highly redundant, suggests that the instruments are assessing the same underlying construct of symptom experience; thus, the shortest version of the instrument (the 5-item duration-based VVSymQ[®] instrument) was further evaluated in this study for psychometric performance.

4.3 Item Distribution and Descriptive Statistics

All five VVSymQ[®] symptoms were endorsed by at least 75 % of patients in the screening and baseline periods (i.e., days –14 to –8 and days –7 to –1, respectively). VVSymQ[®] items were slightly positively skewed (i.e., toward lower scores), and no ceiling or floor effects were observed at baseline (Table 3). Table 4 presents descriptive statistics for the VVSymQ[®] scores at baseline and at week 8, and the changes in VVSymQ[®] scores at week 8.

Table 3 Distribution of 5-item VVSymQ[®] (7-day average) scores at baseline and post-treatment (week 8)

Parameter	Symptom score, % of patients reporting ^{a,b}					
	0	1	2	3	4	5
Baseline						
Heaviness	26.9	27.6	21.5	14.5	8.7	0.7
Aching	11.3	33.1	24.7	18.9	10.2	1.8
Swelling	37.1	18.5	18.5	10.5	11.3	4.0
Throbbing	34.5	28.4	19.6	13.1	4.0	0.4
Itching	29.8	24.4	28.7	10.5	2.9	3.6
Post-treatment (week 8)						
Heaviness	77.3	21.2	1.5	0.0	0.0	0.0
Aching	64.1	29.3	6.2	0.4	0.0	0.0
Swelling	69.6	21.2	7.0	1.5	0.7	0.0
Throbbing	86.8	10.6	2.6	0.0	0.0	0.0
Itching	78.4	14.3	3.3	2.9	1.1	0.0

^a Score of 0 = none of the time; score of 5 = all of the time

^b The percentages for each symptom score are based on the total number of recorded responses (excluding imputations) during the last 7 calendar days (e.g., days –7 to –1) of each period; each patient contributed as many as 7 observations per diary period

Patients' mean VVSymQ[®] item scores for each symptom at baseline ranged from 1.3 to 1.9 (on a 0–5 scale), indicating that they generally experienced each of the

symptoms “a little of the time” to “some of the time” each day. At the end of the post-treatment period (week 8), patients' mean scores on individual symptoms were reduced to 0.4 or below, indicating that patients were experiencing symptoms between “none of the time” and “a little of the time” each day. The results demonstrate that the duration of symptom reporting reduced with treatment.

Table 4 Descriptive statistics for 5-item VVSymQ[®] scores (7-day average) of the study population ($N = 40$) at baseline and post-treatment (week 8), and change at week 8

Symptom	Score, mean (SD)		
	Baseline	Week 8	Change from baseline
Heaviness	1.6 (1.21)	0.3 (0.37)	−1.3 (1.15)
Achiness	1.9 (1.11)	0.4 (0.45)	−1.5 (1.07)
Swelling	1.5 (1.47)	0.4 (0.67)	−1.1 (1.15)
Throbbing	1.3 (0.99)	0.2 (0.32)	−1.1 (0.90)
Itching	1.4 (1.16)	0.4 (0.69)	−1.1 (0.93)
Total score	7.7 (4.38)	1.6 (1.86)	−6.1 (3.70)

SD standard deviation

4.4 Reliability

The agreement between scores on items from one day to the next was very high for the VVSymQ[®] instrument (Table 5). The scale range for the VVSymQ[®] score is 0–25, and the range of mean scores was approximately 6 units. When the mean scores for days −7 to −2 (during the baseline period) were compared with the mean score for day −1, the modal deviation was 1 scale unit in all cases. There was little or no tendency for scores to differ more markedly, as assessed by absolute agreement, kappa

Table 5 Agreement of VVSymQ[®] scores (0–25) of the study population ($N = 40$) during screening and at baseline

Score	Patients, n (%)									
	Day −7	Day −6	Day −5	Day −4	Day −3	Day −2	Day −1	Screening days −7 to −1	Baseline days −14 to −8	
Comparison with day −1										
Exact match	2 (5.0)	5 (12.5)	11 (27.5)	11 (27.5)	6 (15.0)	7 (17.5)		17 (42.5)	5 (12.5)	
Differ by >0 to ≤1	14 (35.0)	11 (27.5)	12 (30.0)	11 (27.5)	10 (25.0)	15 (37.5)		11 (27.5)	11 (27.5)	
Differ by >1 to ≤2	13 (32.5)	8 (20.0)	5 (12.5)	5 (12.5)	10 (25.0)	6 (15.0)		6 (15.0)	14 (35.0)	
Differ by >2 to ≤3	4 (10.0)	4 (10.0)	6 (15.0)	4 (10.0)	7 (17.5)	2 (5.0)		4 (10.0)	6 (15.0)	
Differ by >3 to ≤4	2 (5.0)	8 (20.0)	4 (10.0)	2 (5.0)	3 (7.5)	5 (12.5)			3 (7.5)	
Differ by >4 to ≤5	2 (5.0)	3 (7.5)		3 (7.5)		2 (5.0)		2 (5.0)		
Differ by >5	3 (7.5)	1 (2.5)	2 (5.0)	4 (10.0)	4 (10.0)	3 (7.5)			1 (2.5)	
Simple kappa	−0.013	0.073	0.216	0.227	0.094	0.122		0.385	0.073	
Weighted kappa	0.559	0.574	0.663	0.592	0.560	0.585		0.773	0.612	
Intraclass correlation coefficient	0.767	0.826	0.859	0.775	0.797	0.785		0.919	0.867	
Comparison with screening days (−7 to −1)										
Exact match	13 (32.5)	13 (32.5)	12 (30.0)	14 (35.0)	18 (45.0)	13 (32.5)	17 (42.5)		15 (37.5)	
Differ by >0 to ≤1	13 (32.5)	10 (25.0)	12 (30.0)	15 (37.5)	9 (22.5)	17 (42.5)	11 (27.5)		12 (30.0)	
Differ by >1 to ≤2	3 (7.5)	7 (17.5)	10 (25.0)	2 (5.0)	5 (12.5)	4 (10.0)	6 (15.0)		11 (27.5)	
Differ by >2 to ≤3	4 (10.0)	7 (17.5)	4 (10.0)	4 (10.0)	4 (10.0)	1 (2.5)	4 (10.0)		2 (5.0)	
Differ by >3 to ≤4	5 (12.5)	2 (5.0)	2 (5.0)	1 (2.5)	3 (7.5)	2 (5.0)				
Differ by >4 to ≤5		1 (2.5)		3 (7.5)	1 (2.5)		2 (5.0)			
Differ by >5	2 (5.0)			1 (2.5)		3 (7.5)				
Simple kappa	0.280	0.276	0.248	0.299	0.403	0.271	0.385		0.316	
Weighted kappa	0.673	0.729	0.741	0.705	0.757	0.714	0.773		0.788	
Intraclass correlation coefficient	0.843	0.923	0.923	0.872	0.933	0.890	0.919		0.960	

When reproducibility is evaluated, scores are expected to remain stable for the same respondents over a period of time when minimal or no change in the condition is expected. Table 5 presents the VVSymQ[®] score agreement during this study's screening-to-baseline 7-day period. The frequency of exact matches and differences in score are presented in relation to patients' scores for day −1. For example, two patients scored exactly the same on the VVSymQ[®] on days −7 and −1, while three patients had a score difference of >5 on the VVSymQ[®] on days −7 and −1

Table 6 Correlations of baseline scores, post-treatment (week 8) scores, and change scores (from baseline to week 8) between the VVSymQ® instrument and other clinical assessments

Assessment	Pearson correlation coefficient, <i>r</i>									
	VVSymQ® instrument (diary) ^a	VEINES-QOL/Sym	CIVIQ-20			VCSS	CEAP ^a	PA-V ³	PGIC ^b	
			Psychological	Pain	Physical					Social
Baseline										
VVSymQ® Instrument ^c	1.0000	-0.7268***	0.4237**	0.6010***	0.3139	0.3607*	0.5175***	0.1438	-0.0519	0.3180*
VEINES-QOL/Sym	1.0000	1.0000	-0.8075***	-0.7692***	-0.4402**	-0.3628*	-0.7937***	-0.2200	0.0584	-0.4670**
CIVIQ-20			1.0000	0.6506***	0.5391***	0.3328*	0.8875***	0.1953	0.0077	0.4949**
Psychological			1.0000	1.0000	0.5003**	0.4620**	0.7998***	0.1760	-0.0378	0.2693
Pain			1.0000	1.0000	1.0000	0.7737***	0.8127***	0.1827	0.0036	0.2249
Physical			1.0000	1.0000	1.0000	1.0000	0.6765***	0.2286	0.1060	0.0797
Social			1.0000	1.0000	1.0000	1.0000	1.0000	0.2355	0.0151	0.4002*
Total			1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.5563***	0.1308
VCSS			1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0403
CEAP ^a			1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
PA-V ³			1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Post-treatment (week 8)										
VVSymQ® Instrument ^c	1.0000	-0.7548***	0.5981***	0.5356***	0.4003*	0.4671**	0.5868***	0.4558**		0.3042
VEINES-QOL/Sym	1.0000	1.0000	-0.7877***	-0.7023***	-0.6103***	-0.6690***	-0.8011***	-0.2203		-0.3295*
CIVIQ-20			1.0000	0.6523***	0.7150***	0.7378***	0.9246***	0.2473		0.4385**
Psychological			1.0000	1.0000	0.7523***	0.7109***	0.8500***	0.1860		0.2355
Pain			1.0000	1.0000	1.0000	0.7514***	0.8839***	0.1404		0.3687*
Physical			1.0000	1.0000	1.0000	1.0000	0.8705***	-0.0334		0.2470
Social			1.0000	1.0000	1.0000	1.0000	1.0000	0.1884		0.3919*
Total			1.0000	1.0000	1.0000	1.0000	1.0000	1.0000		0.2514
VCSS			1.0000	1.0000	1.0000	1.0000	1.0000	1.0000		1.0000
PA-V ³			1.0000	1.0000	1.0000	1.0000	1.0000	1.0000		1.0000
Changes from baseline to week 8										
VVSymQ® Instrument ^c	1.0000	-0.6702***	0.3402*	0.6504*	0.3400*	0.3273*	0.4847**	-0.0376		0.1743
VEINES-QOL/Sym	1.0000	1.0000	-0.7091***	-0.7509***	-0.5443***	-0.5128***	-0.7818***	-0.1746		-0.3508*
CIVIQ-20			1.0000	0.5961***	0.5127***	0.4788**	0.8672***	0.1741		0.4455**
Psychological			1.0000	1.0000	0.5159***	0.4948**	0.7766***	0.1686		0.3624*
Pain			1.0000	1.0000	1.0000	0.8366***	0.8237***	0.1205		0.1532
Physical			1.0000	1.0000	1.0000	1.0000	0.7932***	0.1757		0.0714
Social			1.0000	1.0000	1.0000	1.0000	1.0000	0.1947		0.3585*
Total			1.0000	1.0000	1.0000	1.0000	1.0000	1.0000		-0.2419
VCSS			1.0000	1.0000	1.0000	1.0000	1.0000	1.0000		-0.0297

Table 6 continued

Assessment	Pearson correlation coefficient, <i>r</i>											
	VVSymQ [®] instrument (diary) ^a	VEINES-QOL/Sym	CIVIQ-20	Psychological	Pain	Physical	Social	Total	VCSS	CEAP ^a	PA-V ³	PGIC ^b
PA-V ³											1.0000	-0.3271*
PGIC												1.0000

CEAP Clinical-Etiology-Anatomy-Pathophysiology Classification of Venous Disorders, CIVIQ-20 Chronic Venous Disease Quality-of-Life Questionnaire, PA-V³ Patient Self-Assessment of Appearance of Visible Varicose Veins, PGIC Patient Global Impression of Change, VCSS Venous Clinical Severity Score, VEINES-QOL/Sym Venous Insufficiency Epidemiological and Economic Study—Quality of Life/Symptoms

* $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$

^a CEAP clinical severity grade range 2–5; available only for baseline

^b The correlation is between changes in VVSymQ[®] scores and PGIC symptom scores at week 8

^c Sum of 5 items: heaviness, achiness, swelling, throbbing, and itching

values, or ICCs, when separated by greater time intervals (e.g., there was little or no difference between the day –2/–1 comparison and the day –7/–1 comparison). Similar results were observed when each of the daily scores was compared with the mean score for days 1–7. Scores for days –14 to –8 showed slightly less agreement with day –1 scores than did scores for days –7 to –1. The ICC of the comparison between the mean VVSymQ[®] score at baseline (days –14 to –8) and screening (days –7 to –1) was 0.96. A similar pattern was observed in the kappa values (non-parametric assessment of agreement). The Cronbach's alpha values for the VVSymQ[®] instrument at baseline and week 8 were 0.78 and 0.76, respectively.

4.5 Construct Validity

Table 6 shows the results from correlation analyses used in the examination of construct validity. VVSymQ[®] scores showed strong correlations with the Modified VEINES-QOL scores at baseline ($r = -0.73$) and week 8 ($r = -0.75$), and for change from baseline to week 8 ($r = -0.67$). The correlations were negative because higher VEINES-QOL scores indicate better health, while higher VVSymQ[®] scores indicate worse symptoms. The CIVIQ-20 showed moderate correlations with the VVSymQ[®] instrument at baseline ($r = 0.52$) and week 8 ($r = 0.59$), and for changes from baseline to week 8 ($r = 0.48$). The CIVIQ-20 subscales with the strongest relationship to symptoms were Pain and Psychological (with r values ranging from 0.54 to 0.64 for Pain and from 0.34 to 0.60 for Psychological for baseline, week 8, and changes at week 8) (Table 6). A weaker correlation was also observed between the VVSymQ[®] score and the PA-V³ appearance score at baseline ($r = 0.32$).

No correlation was observed between the VVSymQ[®] score and the CEAP grade at baseline ($r = -0.05$). A moderate correlation was found between the VVSymQ[®] score and VCSS at week 8 ($r = 0.46$) but not at baseline or for score changes from baseline to week 8.

4.6 Ability to Detect Change

A large reduction in the VVSymQ[®] mean score (–6.1, indicating symptom improvement) was observed between baseline and week 8 (Table 4), resulting in a very large effect size of 1.6.

4.7 Clinically Meaningful Change

Of the 40 patients who completed the PGIC, 35 (87.5 %) reported that their symptoms were “much improved,” four (10 %) reported that their symptoms were “moderately improved,” and one (2.5 %) reported that symptoms were

Table 7 Mean changes in VVSymQ[®] scores (from baseline to week 8) according to Patient Global Impression of Change (PGIC) symptom scores at week 8

PGIC symptom	VVSymQ [®] score change ^a		
	<i>N</i>	Mean (SD)	Median (range)
Much or moderately improved	39	−6.3 (3.63)	−5.9 (−15.2 to −0.7)
Much improved	35	−6.1 (3.69)	−5.6 (−15.2 to −0.7)
Moderately improved	4	−7.6 (3.17)	−7.8 (−10.7 to −4.0)
Less than moderately improved	1	−0.3 (NA)	−0.3 (−0.3 to −0.3)
A little improved	1	−0.3 (NA)	−0.3 (−0.3 to −0.3)
No change	0	NA	NA
A little worse	0	NA	NA
Moderately worse	0	NA	NA
Much worse	0	NA	NA

NA not applicable, SD standard deviation

^a Change from baseline to week 8 (7-day summary of VVSymQ[®] score)

Table 8 Mean changes in VVSymQ[®] scores (from baseline to week 8) of patients who reported “moderately improved” or “much improved” Patient Global Impression of Change (PGIC) symptom scores at week 8, according to baseline VVSymQ[®] scores

Baseline VVSymQ [®] score	VVSymQ [®] score change ^a	
	<i>N</i>	Mean (SD)
≤7	22	−3.84 (1.60)
>7 to 10	8	−7.45 (1.06)
>10	9	−11.12 (3.29)

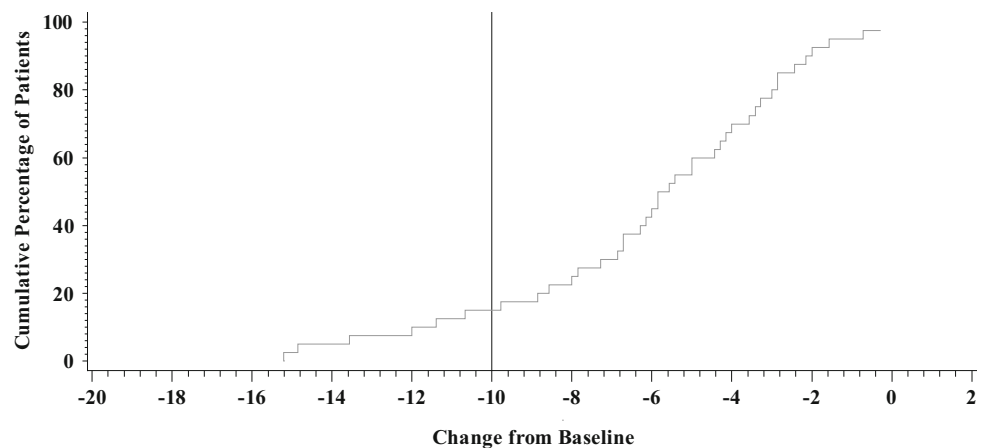
SD standard deviation

^a Change from baseline to week 8 (7-day summary of VVSymQ[®] score)

“a little improved” (Table 7). No patient reported an unchanged or worsened condition. Typically, the threshold for determining clinically meaningful change is based on the change in the score associated with patients who reported that their symptoms were “moderately improved.” However, in this study, there were so few patients in the “moderately improved” group that the threshold was

determined by the mean improvement of patients who were either “moderately improved” or “much improved”. The group of patients who reported on the PGIC that their symptoms were “moderately improved” or “much improved” had a mean improvement of −6.3 points on the VVSymQ[®] instrument, indicating that the upper limit of the clinically meaningful threshold for change in VVSymQ[®] scores was approximately −6.3 for treatment responders (Table 7). Improvements on the VVSymQ[®] instrument varied according to the baseline symptom burden: patients with baseline VVSymQ[®] scores ≤7, 7–10, and >10 who reported that their symptoms were “moderately improved” or “much improved” on the PGIC after treatment had mean improvements in VVSymQ[®] scores of −3.8, −7.5, and −11.1, respectively (Table 8).

In order to provide further information for a responder definition, a cumulative distribution was developed for the VVSymQ[®] instrument and is shown in Fig. 2. The cumulative distribution curve shows that 50 % of patients had an improvement of at least −5.8 points on the VVSymQ[®] instrument.

Fig. 2 Cumulative percentage of patients at week 8 with changes from baseline in 7-day average VVSymQ[®] scores

5 Discussion

Varicose veins are common and significantly impact patients' daily lives. This study examined the psychometric characteristics of a new PRO instrument for varicose vein symptoms, which was developed to address the most relevant varicose vein patient experiences and the limitations of existing PRO and ClinRO instruments. The results from this study demonstrate the sound psychometric performance of the VVSymQ[®] instrument.

The high level of patient compliance with use of the electronic daily diary (including the VVSymQ[®] instrument) indicated that it is easy to use and is not burdensome to complete. Though no universally accepted "gold standard" for compliance exists, rates ≥ 85 to 90 % can be interpreted as strong for clinical trials [17, 18]. In this context, compliance with the daily diary was observed to be excellent and supports an electronic diary approach to measuring symptoms of varicose veins. The results clearly demonstrate the feasibility of utilizing electronic diaries in this patient population.

Scores were highly redundant for duration and intensity instrument versions of varying length; thus, psychometric properties of the shortest version, the 5-item duration-based VVSymQ[®] instrument, were evaluated further. The VVSymQ[®] instrument demonstrated excellent test-retest reliability and good internal consistency reliability, and it related to criterion measures as expected. The VVSymQ[®] instrument was found to be associated, but not redundant, with other PRO measures. Converging scores were observed for the VEINES-QOL, CIVIQ-20, and VVSymQ[®] instruments. The levels of the correlations between the VVSymQ[®] and VEINES-QOL/Sym instruments showed that higher levels of symptoms are related to lower levels of vein disease QOL as measured by symptom complaints and the impact of those symptoms on patients (i.e., the areas comprising the VEINES-QOL score). Of the CIVIQ-20 subscales, the Pain and Psychological subscale scores showed the strongest relationship to the VVSymQ[®] score. This is not surprising, given that the Pain items would be expected to be related to symptoms and the Psychological items ask about the immediate impact of the condition (e.g., how patients feel about themselves). The modest significant relationship observed between the VVSymQ[®] and PA-V³ instruments is not unexpected, given that the same venous condition is causing both symptoms and appearance concerns.

The results suggest that there is no clear relationship between patients' self-reports of varicose vein symptoms (the VVSymQ[®] instrument) and ClinRO measures. Because patients and physicians do not necessarily agree about the impact of a condition, high correlations were not

anticipated, but r values >0.3 are often found between the scores for ClinRO and PRO measures in other conditions (see, for example, Hinchcliff et al. [23] and Mazari et al. [24]). The ClinRO instruments themselves were correlated as expected at baseline, indicating that these measures are reasonably reliable. However, no correlation was found between the VVSymQ[®] score and CEAP clinical grade. No correlation was found between the VVSymQ[®] score and VCSS at baseline and the change scores (from baseline to week 8); however, a moderate correlation was observed between the VVSymQ[®] score and VCSS at week 8. These findings may be attributed to the restricted range in the scores at baseline (due to inclusion criteria requiring patients to have a screening symptom score ≥ 7 as derived from question 1 on the Modified VEINES-QOL/Sym instrument), which can attenuate the correlation between scales and underestimate the actual true correlation. Additionally, the PRO and ClinRO measures may be assessing uniquely different concepts associated with varicose veins. For instance, the low to moderate correlations observed between the VVSymQ[®] score and VCSS suggest that the VCSS may characterize disease severity, yet may not predict patient symptom response [11].

The VVSymQ[®] instrument appears sensitive to changes in the clinical condition over time. Large reductions were observed in VVSymQ[®] scores from baseline to week 8, resulting in a very large effect size of 1.6 [21].

The sample size and diversity may have limited the findings of this study. This study poses a challenge in interpreting the definition of a responder based upon the PGIC anchor. The traditional approach is to compare the sizes of changes in the outcome measures between patients showing no improvement, minimal improvement, etc. However, no patients in this study reported less than minimal improvement, and only one patient reported minimal improvement. Thus, the number of patients at the lower end of the scale was too small to give useful information for the separate categories. The mean change in the 7-day average electronic daily diary VVSymQ[®] score of 6.3 noted for all patients who reported "moderately improved" or "much improved" symptoms on the PGIC can, therefore, be considered an upper bound for determination of a criterion for the definition of a responder. Results from larger studies will allow a more precise estimate to be obtained.

The findings from the PGIC anchor also suggest that the threshold for a clinically meaningful change in the VVSymQ[®] score is higher in patients who had a greater baseline VVSymQ[®] symptom burden, and that the VVSymQ[®] instrument can capture a clinically meaningful treatment benefit even in patients who report a relatively small baseline symptom burden.

6 Conclusion

The 5-item VVSymQ® instrument demonstrated favorable psychometric properties and is a brief, useful assessment for measuring patient-reported symptoms of varicose veins. Understanding the patient's perspective on their varicose vein symptoms can allow for a more informed assessment of treatment efficacy. The VVSymQ® instrument assesses symptoms that are related to the underlying pathophysiology of varicose veins and that are important to patients. This PRO measure may be useful as an efficacy endpoint alongside other measures of disease severity in future clinical trials testing new treatments for varicose veins.

Acknowledgments The authors would like to thank Mr. Gareth Bate (the local study coordinator) and Dr. Emma Thomas, PhD (Senior Manager, Clinical Development, BTG International Ltd.) for their assistance with the smooth running of the study. Tuli Ahmed (JetStream Clinical, LLC, King of Prussia, PA, USA) provided editorial assistance (funded by Biocompatibles, Inc., a BTG International group company). Diane Turner-Bowker drafted the manuscript. David Wright, Jean Paty, Diane Turner-Bowker, and Andrew Bradbury made the following contributions: interpretation of data for the work; critical revision of the manuscript for important intellectual content; approval of the final version of the manuscript to be published; and provision of agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. David Wright is the overall guarantor.

Compliance with Ethical Standards

Funding This study was sponsored by BTG International Ltd.

Conflict of interest David Wright is an employee of BTG International Ltd., the study sponsor. Jean Paty and Diane Turner-Bowker are former employees of ERT, which was contracted by BTG International Ltd. for work on the study and the manuscript; Jean Paty is now an employee of Quintiles (Hawthorne, NY, USA), and Diane Turner-Bowker is now an employee of Adelphi Values (Boston, MA, USA). Andrew Bradbury was the principal investigator of this study, and his site received funding for his participation in the study.

Informed consent Written informed consent was obtained from all participants included in the study.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Beebe-Dimmer JL, Pfeifer JR, Engle JS, Schottenfeld D. The epidemiology of chronic venous insufficiency and varicose veins. *Ann Epidemiol.* 2005;15:175–84.
2. Darvall KA, Bate GR, Adam DJ, Bradbury AW. Generic health-related quality of life is significantly worse in varicose vein patients with lower limb symptoms independent of CEAP clinical grade. *Eur J Vasc Endovasc Surg.* 2012;44:341–4.
3. Kahn SR, M'LAN CE, Lamping DL, for the VEINES Study Group, et al. Relationship between clinical classification of chronic venous disease and patient-reported quality of life: results from an international cohort study. *J Vasc Surg* 2004;39:823–8.
4. Marston WA. Evaluation of varicose veins: what do the clinical signs and symptoms reveal about the underlying disease and need for intervention? *Semin Vasc Surg.* 2010;23:78–84.
5. Paty J. VVSymQ® score and patient profiles: interpreting a new patient-reported outcome (PRO) instrument for great saphenous vein incompetence (GSVI). Presented at the Society for Interventional Radiology 2014 annual meeting, San Diego, California, 22–27 Mar 2014.
6. US Food and Drug Administration. Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims. US Food and Drug Administration. 2009. Available at: <http://www.fda.gov/downloads/Drugs/.../Guidances/UCM193282.pdf>. Accessed 6 Nov 2015.
7. Garratt AM, Macdonald LM, Ruta DA, et al. Towards measurement of outcome for patients with varicose veins. *Qual Health Care.* 1993;2:5–10.
8. Lamping DL, Schroter S, Kurz X, et al. Evaluation of outcomes in chronic venous disorders of the leg: development of a scientifically rigorous, patient-reported measure of symptoms and quality of life. *J Vasc Surg.* 2003;37:410–9.
9. Guex JJ, Zimmet SE, Boussetta S, et al. Construction and validation of a patient-reported outcome dedicated to chronic venous disorders: SQOR-V (Specific Quality of Life and Outcome Response—Venous). *J Mal Vasc.* 2007;32(3):135–47.
10. Launois R, Reboul-Marty J, Henry B. Construction and validation of a quality of life instrument in chronic lower limb venous insufficiency (CIVIQ). *Qual Life Res.* 1996;5:539–54.
11. Paty J, Turner-Bowker DM, Elash CA, Wright D. The VVSymQ® instrument: use of a new patient-reported outcome measure for assessment of varicose vein symptoms. *Phlebology.* 2015. <http://phl.sagepub.com/content/early/2015/07/15/026835515595193.long>. Accessed 18 Feb 2016.
12. Kahn SR, Lamping DL, Ducruet T, et al. VEINES-QOL/Sym instrument was a reliable and valid disease-specific quality of life measure for deep venous thrombosis. *J Clin Epidemiol.* 2006;59:1049–56.
13. Norquist JM, Girman C, Fehnel S, DeMuro-Mercon C, Santanello N. Choice of recall period for patient-reported outcome (PRO) measures: criteria for consideration. *Qual Life Res.* 2012;21:1013–20.
14. Kingsley J, King T, Orfe E, et al. Capturing the visual impact of varicose veins [abstract no. PS138]. Society for Vascular Surgery. 2012. <http://www.vascularweb.org/educationandmeetings/2012%20Vascular%20Annual%20Meeting/programdetail/Pages/Friday,%20June%208%20Abstracts/PS138.aspx>. Accessed 3 June 2015.
15. Lohr KN, Aaronson NK, Alonso J, et al. Evaluating quality-of-life and health status instruments: development of scientific review criteria. *Clin Ther.* 1996;18:979–92.
16. Patrick DL, Burke LB, Powers JH, et al. Patient-reported outcomes to support medical product labeling claims: FDA perspective. *Value Health.* 2007;10(Suppl.):S125–37.
17. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. Norwalk: Appleton & Lange; 1993.
18. Fleiss J. Reliability of measurement: the design and analysis of clinical experiments. New York: Wiley; 1986.
19. Gwaltney CJ, Shields AL, Shiffman S. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. *Value Health.* 2008;11:322–33.

20. Mukaka MM. Statistics corner: a guide to appropriate use of correlation coefficient in medical research. *Malawi Med J.* 2012;24:69–71.
21. Cohen J. *Statistical power analysis for the behavioral sciences.* 2nd ed. Hillsdale: Erlbaum; 1988.
22. Guyatt GH, Osoba D, Wu AW, et al. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc.* 2002;77:371–83.
23. Hinchcliff M, Beaumont JL, Thavarajah K, et al. Validity of two new patient-reported outcome measures in systemic sclerosis: Patient-Reported Outcomes Measurement Information System 29-Item Health Profile and Functional Assessment of Chronic Illness Therapy—Dyspnea short form. *Arthritis Care Res (Hoboken).* 2011;63:1620–8.
24. Mazari FAK, Carradice D, Rahman MN, et al. An analysis of relationship between quality of life indices and clinical improvement following intervention in patients with intermittent claudication due to femoropopliteal disease. *J Vasc Surg.* 2010;52:77–84.