
Article:
Aber, A. orcid.org/0000-0002-5707-2430, Poku, E. orcid.org/0000-0001-6549-5081, Phillips, P. orcid.org/0000-0003-0241-1527 et al. (6 more authors) (2017) Systematic review of patient-reported outcome measures in patients with varicose veins. British Journal of Surgery, 104 (11). pp. 1424-1432. ISSN 0007-1323

https://doi.org/10.1002/bjs.10639

Systematic review of patient-reported outcome measures in patients with varicose veins

A. Aber[1], E. Poku[1], P. Phillips[1], M. Essat[1], H. Buckley Woods[1], S. Palfreyman[3], E. Kaltenthaler[1], G. Jones[2] and J. Michaels[1]

[1]School of Health and Related Research, University of Sheffield, Sheffield, and [2]Department of Psychology, School of Social Sciences, Leeds Beckett University, Leeds, UK, and [3]Faculty of Nursing, University of Alberta, Edmonton, Alberta, Canada

Correspondence to: Mr A. Aber, School of Health and Related Research, University of Sheffield, Regent Court, 30 Regent Street, Sheffield S1 4DA, UK (e-mail: a.aber@sheffield.ac.uk)

ahmedaber@gmail.com

Preliminary findings presented to the Vascular Society Annual Scientific Meeting, Manchester, UK, November 2016

Background: Varicose veins can affect quality of life. Patient-reported outcome measures (PROMs) provide a direct report from the patient about the impact of the disease without interpretation from clinicians or anyone else. The aim of this study was to examine the quality of the psychometric evidence of PROMs used in patients with varicose veins.

Methods: A systematic review was undertaken to identify studies that reported the psychometric properties of generic and disease-specific PROMs in patients with varicose veins. Literature searches were conducted in databases including MEDLINE, up to July 2016. The psychometric criteria used to assess these studies were adapted from published recommendations in accordance with US Food and Drug Administration guidance.

Results: Nine studies were included which reported on aspects of the development and/or validation of one generic (36-Item Short Form Survey, SF-36®) and three disease-specific (Aberdeen Varicose Vein Questionnaire, AVVQ; Varicose Veins Symptoms Questionnaire, VVSymQ®; Specific Quality-of-life and Outcome Response – Venous, SQOR-V) PROMS. The evidence from included studies provided data to support the construct validity, internal consistency and responsiveness of the AVVQ; this instrument also had test–retest reliability (intraclass correlation coefficient 0.59). However, its content validity, including weighting of the AVVQ questions, was biased and based on

the opinion of clinicians, and the instrument had poor acceptability. VVSymQ® displayed good responsiveness and acceptability rates. SF-36® was considered to have satisfactory responsiveness and internal consistency (Cronbach's $\alpha = 0.80$).

Conclusion: There is a scarcity of psychometric evidence for PROMs used in patients with varicose veins. These data suggest that AVVQ and SF-36® are the most rigorously evaluated PROMs in patients with varicose veins.

+A: Introduction

Varicose veins are enlarged lumpy visible veins caused by reflux of blood in the superficial veins of the leg[1]. They are extremely common, affecting more than half of the population in Western Europe and North America[2–4]. Varicose veins can cause symptoms such as pain, aching, swelling, throbbing, cramping, itching and bleeding[5]. Complications include superficial thrombophlebitis, external bleeding, lipodermatosclerosis, eczema and ulceration[6,7]. Traditionally, treatment comprised surgery with stripping of the great saphenous vein and removal of the varicose veins through small incisions (avulsions or phlebectomies). However, in the past decade new less invasive treatments have been developed[8]. In 2009–2010, 35 659 varicose vein procedures were carried out in the National Health Service (NHS)[8].

Patient-reported outcome measures (PROMS) provide a means by which the impact of varicose veins or their treatments on quality of life can be measured. The questionnaires are typically developed from qualitative studies involving patients and clinicians. The items in these questionnaires are then tested for their ability to capture the patient's experience in prospective surveys, using psychometric analyses to explore the relationship of the items with each other and their overall ability to detect change[9]. The NHS PROMS programme has been collecting PROMs data from patients undergoing varicose vein interventions since April 2009 using generic and disease-specific PROMS[10].

The aim of this study was to identify and examine the quality of the psychometric evidence for PROMs used for patients with varicose veins. This study was divided into two parts; initially a systematic review was undertaken to identify the appropriate papers, and then a psychometric assessment was undertaken to assess the quality of the methods used to validate or design these PROMs.

+A: Methods

A systematic review was undertaken and reported in accordance with the general principles recommended in PRISMA statement[11]. The protocol for the systematic review was developed and registered in the PROSPERO international prospective register of systematic reviews before the start of the data extraction[12].

Systematic searches were undertaken in MEDLINE and MEDLINE In-Process, Embase, the Cochrane Library, CINAHL, PROQOLID, PsycINFO and Web of Science. A two-stage search approach was used. The first stage used general terms for PROMs (known generic and condition-specific PROMS) and terms for the condition (varicose veins) to identify studies. These were retrieved, and the title and abstract examined for additional PROM terms used in patients with varicose veins. The second stage incorporated these terms with the preliminary search strategy and a methodological search filter for finding studies on measurement properties. Databases were searched from inception up to July 2016 for search 1 and up to July 2016 for search 2. Searches were supplemented by hand-searching reference lists of relevant reviews and included studies, citation search of included studies and contact with experts in the field. Search strategies are shown in (Appendix S1, supporting information)

+B: Study selection

The titles were reviewed, and the abstracts and full text of the included articles were assessed by at least two reviewers independently. Any disagreements in the selection process were resolved by discussion, with involvement of a third reviewer. Eligible studies included articles published in English of any study design that reported the validation or development of PROMs capturing quality of life, health status or functional limitation in patients with varicose veins in an English-speaking population (Table 1).

+B: Data abstraction

Data relating to study design, patient characteristics, type of treatment, PROM used, methods and outcomes were extracted by one reviewer on to a standardized data extraction form, and independently checked for accuracy by a second. Any discrepancies were resolved by discussion, with involvement of a third reviewer. Where necessary, study authors were contacted for missing information or additional data.

+B: Methodological quality assessment (psychometric evaluation)

The methodological quality assessment in developing the PROMs was based on specific psychometric criteria. Owing to lack of consensus on how to appraise PROMs, the study-specific criteria were adapted from published recommendations[13–16,18] in accordance with the US Food and Drug Administration (FDA) guidance 2009[17]. They were mainly based on the Oxford University PROMs Group guidelines and the COnsensus-based Standards for the selection of health status Measurement

INstruments (COSMIN)[19]. These criteria can be divided into four areas: reliability, validity, responsiveness and acceptability (Table 2). Two independent researchers appraised these psychometric properties for each PROM independently using the following methods of assessment. A rating scale was designed to allocate a mark for each domain: 0, not reported; –, evidence not in favour; +/–, conflicting evidence; and +, evidence in favour. Any disagreements were resolved through discussion or with involvement of a psychometrics expert.

+C: Assessment of reliability

The reliability of a PROM is its ability to produce the same results when measurements are repeated in populations with similar characteristics[20]. The reliability of each identified PROM was assessed by examining the reported data on reproducibility and internal consistency. The reproducibility of an instrument is commonly examined by performing test–retest at different time points. The degree of correlation is examined between the scores at baseline and those at different time points. PROMs should report test–retest using the intraclass correlation or weighted κ score; this should be at least 0.70 for group comparisons[20].

PROMs commonly use more than one item to measure a single dimension that is important to the patient; this is because several related observations can produce a better estimate than one. These items need to be homogeneous; this means that they all measure aspects of a single attribute rather than different ones and are therefore internally consistent[13]. Internal consistency is usually measured using Cronbach's α, which should have a value of more than 0.70 and below 0.90 for the proposed PROM to be psychometrically sound[13,23].

+C: Assessment of validity

Validity is the measure of how well a PROM measures what it is intended to measure. Validity was assessed for each identified PROM by assessing content validity, construct validity and criterion validity. Content validity was measured by examining the relevance of the items in the PROM to their intended use. This was assessed on the basis of whether these items were developed through qualitative studies with patient groups involving clinicians and incorporating published evidence[23]. Criterion validity is concerned with assessing the PROM in question against a standard PROM that provides a benchmark of the true values. The new PROM should demonstrate correlation coefficient scores of more than 0.70. However, in reality this is often very difficult to assess in the absence of such a standard[14,15].

+C: Assessment of responsiveness

This is defined as the ability of a PROM to detect clinically important change over time, if a true change exists. The PROM should be able to distinguish between clinically important changes and measurement error. Responsiveness of a measure can be calculated using methods such as use of standardized response means, t test, effect size and Guyatt's responsiveness ratio[21,22,24].

+B: Assessment of acceptability and floor or ceiling effect

Acceptability is measured by the completeness of the data. For a PROM to show a good level of acceptability, 80 per cent or more of the data should be complete when the PROM is administered to the patients[19]. A floor or ceiling effect is considered if 15 per cent of respondents are achieving the lowest or the highest score on the instrument.

+A: Results

A total of 3647 records were identified; following detailed examination, nine studies[25–33] (reporting on 4 PROMs) were included (Fig. 1). PROMs that were not specific for varicose veins and examined chronic venous disease in general were excluded; examples of these are the ChronIc Venous Insufficiency quality of life Questionnaire (CIVIQ) 20 and CIVIQ-14, both chronic venous disease PROMS, and the Venous Insufficiency Epidemiologic and Economic Study – Quality of Life/Symptoms (VEINES-QOL/Sym), a PROM validated in patients with deep venous thrombosis and venous leg ulcers.

All the included studies assessed the psychometric properties and suitability of the suggested PROMs in patients with varicose veins (Table 3). The studies were prospective in design, and were undertaken in the UK and USA. They were published between 1992 and 2016. The majority of the studies were of a small to moderate size with the number of patients ranging from 40[33] to 1700[24,25]. Patients aged between 16 and 86 years were recruited in the included studies, with the proportion of men ranging from 24 per cent[24] to 47.6 per cent[27].

+B: Patient-reported outcomes measurement data and psychometric evaluation

Overall, data relating to the development and psychometric evaluation of one generic PROM and three condition-specific PROMs for patients with varicose veins were available. The only generic PROM evaluated was the 36-Item Short Form Health Survey (SF-36®)[25,27]. The condition-specific PROMs were the Aberdeen Varicose Vein Questionnaire (AVVQ)[26,28–30], the Varicose Veins Symptoms Questionnaire (VVSymQ®) and the Specific Quality-of-life and Outcome Response (SQOR-V)[31,32].

The protocol regarding timing of PROMs differed between the studies. The shortest follow-up was immediately following the intervention and the longest was 12 months after treatment. The rigour of the psychometric assessment of the PROMs was variable. The AVVQ was the only instrument evaluated in detail, with assessment of all the important psychometric domains were assessed (Table 4.)

+C: Short Form Health Survey 36

Garratt and colleagues[25–28] assessed aspects of the psychometric validity of this generic instrument in patients with varicose veins. In a study of 1700 patients, including 314 with varicose veins, the SF-36® was examined for its suitability as a PROM for patients treated in the NHS. The internal consistency was assessed using two techniques, item scale correlation and Cronbach's α. The first method examined the extent to which an item was related to the rest of the scale, whereas Cronbach's α measured the overall correlation between items in the scale. The correlation for all items was above the 0.4, providing evidence of internal consistency. The Cronbach's α value exceeded 0.8 and satisfied the criteria for internal consistency. The response rate for SF-36® in this study at baseline was 75.5 per cent, showing some evidence of acceptability for this PROM; however, this dropped to 67.5 per cent after 1 year. The construct validity assessment used ordinary least regression to estimate the effect on each scale in the PROM of varicose veins, age, sex and socioeconomic status of the participants. The impact of varicose veins was significant only on the physical functioning scale. The responsiveness of SF-36® was assessed in the same population after 12 months, with results showing good responsiveness for this PROM. The standardized response mean was used to measure this property, and patients with varicose veins had a significantly higher level of improvement across the SF-36® scales at 1 year than those not referred for treatment.

+C: Aberdeen Varicose Vein Questionnaire

This disease-specific PROM was developed by Garratt et al.[25], and the items were generated based on questions commonly used to assess patients with varicose veins. The items generated were confirmed by two clinicians and then pretested in patients for relevance and validity[25]. The AVVQ was tested for internal consistency, construct and criterion validity, and acceptability. The result of internal consistency evaluation after removing five questions that did not fulfil the criteria was a Cronbach's α value of 0.72, satisfying the psychometric criterion for this PROM[34]. The construct validity of the instrument was tested using stepwise multiple regression and comparison with the Varicose Vein Severity Score. The regression model confirmed that AVVQ explains a substantial proportion of the non-random variation in the patients' perceived health. The AVVQ showed high acceptability among patients with 76 per cent complete data when the PROM was administered[25]. Comparing to eight scales of the SF-36 in patients with varicose veins assessed the criterion validity of the AVVQ; the AVVQ achieved highly negative correlations with all eight scales of the SF-36[28]. Four of these correlations exceeded 0.4, including physical functioning, pain, social functioning and role limitations. These correlations suggest that AVVQ can pick up adverse effects of varicose veins better than the generic PROM SF-36®. The test–retest reliability assessment of this PROM showed an intraclass correlation coefficient of above 0.7 in all domains except one, in which patients reported no change in symptoms after 1 year. The responsiveness of the AVVQ to changes in health over time was assessed by administering the questionnaire to the same respondents after 1 year[28]. In an analysis

of standardized response means over 1 year, all items showed improvement, especially for patients who received treatment; patients not referred to a specialist had lower perceived health compared with the general population[28].

Lattimer and colleagues[30] attempted to examine the responsiveness of the AVVQ in patients receiving endogenous laser ablation or foam sclerotherapy for varicose veins as part of an RCT. The patients included in the study all had primary disease with no previous intervention. The Wilcoxon signed-rank test was used to compare differences within the same group before and after intervention. Spearman's ρ was used to assessed the correlation between the severity of symptoms and AVVQ outcomes. The study reported improved AVVQ score after 3 weeks and 3 months of follow-up[29,30].

Varicose Veins Symptoms Questionnaire

This electronic PROM was developed in accordance with the FDA guidance[16]. This included qualitative studies that involved patients to generate the five items in the PROM, all related to symptoms alone. The psychometric properties were examined as part of two RCTs (VANISH-1 and VANISH-2) evaluating microfoam ablation with varying does of polidocanol endovenous microfoam in patients with varicose veins[31,33]. The test–retest reliability was examined using intraclass correlation coefficients to assess whether VVSymQ® yielded a reproducible score in patients exhibiting no change in health status. The reported intraclass correlation coefficient was 0.75, demonstrating acceptable test–retest reliability. Cronbach's α value was 0.76 showing good internal consistency of the items included in the PROM. The construct validity was evaluated through Pearson correlation analyses; the score from the PROM showed correlations with reported clinical outcomes[31]. The VVSymQ® score captured meaningful clinical change and treatment impact, with an effect size of 1.6 when the scores were compared between baseline and 6 weeks after intervention. This electronic PROM had between 86.1 and 97 per cent data completion, reflecting good acceptability among the patients[31,33].

+C: Specific Quality-of-life and Outcome Response – Venous

This instrument consists of 46 items divided into five domains: physical discomfort, appearance, restriction in movement, emotional problems and threat to health. All patients in the study[32] underwent radiofrequency ablation. The performance of the PROM was tested against the AVVQ and other clinical outcomes. The scores from the AVVQ and SQOR-V showed strong positive correlation with a Spearman coefficient of 0.702 ($P < 0.001$). Responsiveness was tested at 6 weeks, with poor results for SQOR-V in some patient groups compared with the AVVQ. The acceptability, as measured by the completeness of the data, was weak (67 per cent complete data) [32].

+A: Discussion

This study identified PROMs that have undergone validation in patients with varicose veins, and assessed the methodology of psychometric validation in accordance with FDA guidance, Oxford PROMS group guidelines and COSMIN[13–19]. Patient-reported outcome is an important core outcome recommended to be collected as part of service analysis and clinical studies[35–37]. Clinicians and researchers are faced with a dilemma when deciding on the instrument that measures this outcome. In the UK NHS, the measures used to collect data on PROMs for patients undergoing surgical management for varicose veins are the AVVQ and EuroQoL Five Dimensions (EQ-5D™; EuroQol Group, Rotterdam, The Netherlands[38].

This review identified only one generic measure (SF-36®) and three disease-specific instruments (AVVQ, VVSymQ®, SQOPR-V) that have undergone psychometric assessment in patients with varicose veins. The evidence suggests that the SF-36® exhibits good internal consistency and acceptability among patients with varicose veins, with some evidence of construct validity and responsiveness. The AVVQ had good test–retest reliability, construct and criterion validity, and responsiveness. However, the evidence for the content validity was weak, and clinicians and researchers generated the items with limited input from patients; the weighting of the items was based on the judgement of two clinicians. VVSymQ® had good internal consistency, test–retest reliability, construct, content and criterion validity, and responsiveness. The acceptability of the VVSymQ® was better than that of the AVVQ and SF-36®; this is in part because it is an electronic questionnaire; however, the only domain in this instrument is symptoms.

The main strength of this study was the use of comprehensive search strategies to identify all relevant papers that reported on psychometric validation of PROMs for patients with varicose veins. The psychometric assessment domains in this study were based on different but overlapping psychometric evaluation criteria[16,17,19,38]. The main limitation of the analysis was the heterogeneity of the patients included in the studies as well as the different protocols for administering the PROMs. Furthermore, the content validity of the disease-specific measures was based on information limited to either that gathered by consulting patients about items generated by researchers and clinicians, or data from small qualitative research studies, with no systematic review of the qualitative evidence[25,27–31,33]. None of the studies included in the review provided any information on how they dealt with missing data.

The only generic PROM with psychometric evidence to support its use in patients with varicose veins was the SF-36®; no data on the EQ-5D™ were found. The AVVQ was the most evaluated disease-specific PROM, with five studies examining its psychometric validity. Further work is needed to improve the content validity and acceptability of PROMs used in patients with varicose veins. The authors also recommend further research on the use of electronic PROMs based on the acceptability data for the VVSymQ®.

+A: Acknowledgements

This research and publication were funded through a National Institute for Health Research Programme Grant (RP-PG-1210-12009).
Disclosure: The authors declare no conflict of interest.

+A: References

1 Evans CJ, Allan PL, Lee AJ, Bradbury AW, Ruckley CV, Fowkes FG. Prevalence of venous reflux in the general population on duplex scanning: the Edinburgh vein study. J Vasc Surg 1998; **28**: 767–776.

2 Beebe-Dimmer JL, Pfeifer JR, Engle JS, Schottenfeld D. The epidemiology of chronic venous insufficiency and varicose veins. Ann Epidemiol 2005; **15**: 175–184.

3 Maurins U, Hoffmann BH, Lösch C, Jöckel KH, Rabe E, Pannier F. Distribution and prevalence of reflux in the superficial and deep venous system in the general population – results from the Bonn Vein Study, Germany. J Vasc Surg 2008; **48**: 680–687.

4 Michaels JA, Campbell WB, Brazier JE, Macintyre JB, Palfreyman SJ, Ratcliffe J et al. Randomised clinical trial, observational study and assessment of cost-effectiveness of the treatment of varicose veins (REACTIV trial). Health Technol Assess 2006; **10**: 1–196, iii–iv.

5 Michaels JA, Campbell B, King B, Palfreyman SJ, Shackley P, Stevenson M. Randomized controlled trial and cost-effectiveness analysis of silver-donating antimicrobial dressings for venous leg ulcers (VULCAN trial). Br J Surg 2009; **96**: 1147–1156.

6 Nijsten T, van den Bos RR, Goldman MP, Kockaert MA, Proebstle TM, Rabe E et al. Minimally invasive techniques in the treatment of saphenous varicose veins. J Am Acad Dermatol 2009; **60**: 110–119.

7 Chen WYJ, Rogers AA. Recent insights into the causes of chronic leg ulceration in venous diseases and implications other types of chronic wounds. Wound Repair Regen 2007; **15**: 434–449.

8 <EPATH>NICE. Varicose Veins. The National Institute for Health and Care Excellence guidelines 2013, July 2013. (https://www.nice.org.uk/guidance/cg168/evidence/varicose-veins-in-the-legs-full-guideline-pdf-191485261 Accessed 15 June 2017)

9 <B>Fayers PM, Machin D. Quality of Life: the Assessment, Analysis and Reporting of Patient Reported Outcomes. Wiley-Blackwell, 2016.

10 <EPATH>NHS PROMs. Patient Reported Outcome Measures (PROMs). NHS Digital 2016, November 2016. (http://content.digital.nhs.uk/proms Accessed 15 June 2017)

11 Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009; **339**: b2700.

12 Poku E, Duncan R, Essat M, Woods H, Phillips P, Goka E et al. Systematic review of patient-reported outcome measures in patients with chronic venous insufficiency. PROSPERO. 2015 Jul; CRD42015024820. (http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42015024820 Accessed 10 June 2017)

13 Fitzpatrick R, Davey C, Buxton M, Jones D. Evaluating patient-based outcome measures for use in clinical trials. Health Technol Assess 1998; **2**: 1–73.

14 Lamping DL, Schroter S, Marquis P, Marrel A, Duprat-Lomon I, Sagnier PP. The community-acquired pneumonia symptom questionnaire – a new, patient-based outcome measure to evaluate symptoms in patients with community-acquired pneumonia. Chest 2002; **122**: 920–929.

15 Morris C, Janssens A, Allard A, Coon JT, Shilling V, Tomlinson R et al. Informing the NHS Outcomes Framework: Evaluating Meaningful Health Outcomes for Children with Neurodisability Using Multiple Methods Including Systematic Review, Qualitative Research, Delphi Survey and Consensus Meeting. Health Services and Delivery Research 2014;2:15: 1- 256

16 Terwee CB, Bot SDM, de Boer MR, van der Windt DA, Knol DL, Dekker J et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007; **60**: 34–42.

17 <B>Food and Drug Administration. Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims: Draft Guidance. Report number 1477–7525. US Department of Health and Human Services Food and Drug Administration, 2009.

18 <B>Department of Health. Guidance on the Routine Collection of Patient Reported Outcome Measures (PROMs). Department of Health: London, 2008.

19 Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010; **63**: 737–745.

20 Brazier J, Deverill M. A checklist for judging preference-based measures of health related quality of life: learning from psychometrics. Health Econ 1999; **8**: 41–51.

21 Deyo RA, Centor RM. Assessing the responsiveness of functional scales to clinical-change – an analogy to diagnostic-test performance. J Chronic Dis 1986; **39**: 897–906.

22 Guyatt GH, Cook DJ. Health-status, quality-of-life, and the individual. JAMA 1994; **272**: 630–631.

23 Zimmerman TF. Health measurement scales: a practical guide to their development and use, 5th edition. World Med Health Policy 2015; **7**: 164–165.

24 Guyatt GH, Deyo RA, Charlson M, Levine MN, Mitchell A. Responsiveness and validity in health-status measurement – a clarification. J Clin Epidemiol 1989; **42**: 403–408.

25 Garratt AM, Macdonald LM, Ruta DA, Russell IT, Buckingham JK, Krukowski ZH. Towards measurement of outcome for patients with varicose veins. Qual Health Care 1993; **2**: 5–10.

26 Garratt AM, Ruta DA, Abdalla MI, Buckingham JK, Russell IT. The SF36 health survey questionnaire: an outcome measure suitable for routine use within the NHS? BMJ 1993; **306**: 1440–1444.

27 Garratt AM, Ruta DA, Abdalla MI, Russell IT. SF 36 health survey questionnaire: II. Responsiveness to changes in health status in four common clinical conditions. Qual Health Care 1994; **3**: 186–192.

28 Garratt AM, Ruta DA, Abdalla MI, Russell IT. Responsiveness of the SF-36 and a condition-specific measure of health for patients with varicose veins. Qual Life Res 1996; **5**: 223–234.

29 Lattimer CR, Kalodiki E, Azzam M, Geroulakos G. The Aberdeen varicose vein questionnaire may be the preferred method of rationing patients for varicose vein surgery. Angiology 2014; **65**: 205–209.

30 Lattimer CR, Kalodiki E, Azzam M, Geroulakos G. Responsiveness of individual questions from the venous clinical severity score and the Aberdeen varicose vein questionnaire. Phlebology 2014;**29**: 43–51.

31 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 2016; **31**: 481–488.

32 Shepherd AC, Gohel MS, Lim CS, Davies AH. A study to compare disease-specific quality of life with clinical anatomical and hemodynamic assessments in patients with varicose veins. J Vasc Surg 2011; **53**: 374–382.

33 Wright DD, Paty J, Turner-Bowker DM, Bradbury A. Psychometric evaluation of a new patient-reported outcome (PRO) symptom diary for varicose veins: VVSymQ(®) instrument. Patient 2016; **9**: 335–348.

34 <B>Kline P. A Handbook of Test Construction (Psychology Revivals): Introduction to Psychometric Design. Methuen: London, 1986.

35 <B>Nunnally J, Bernstein J. Psychometric Theory (3rd edn). McGraw-Hill, 1994.

36 Patrick DL, Burke LB, Powers JH, Scott JA, Rock EP, Dawisha S et al. Patient-reported outcomes to support medical product labeling claims: FDA perspective. Value Health 2007; **10**(Suppl 2): S125–S137.

37 What quality of life? The WHOQOL Group. World Health Organization Quality of Life Assessment. World Health Forum 1996; **17**: 354–356.

38 McKenna SP. Measuring patient-reported outcomes: moving beyond misplaced common sense to hard science. BMC Med 2011; **9**: 86.

39 Francis DO, McPheeters ML, Noud M, Penson DF, Feurer ID. Checklist to operationalize measurement characteristics of patient-reported outcome measures. Syst Rev 2016; **5**: 129.

Supporting information

Additional supporting information may be found online in the supporting information tab for this article.

Appendix S1 Search strategy (Word document)

Typesetter: please refer to marked-up figures

Fig. 1 PRISMA diagram showing selection of studies for review of patient-reported outcome measures (PROMs) in patients with varicose veins

Table 1 Criteria for considering eligibility of studies

	Inclusion criteria	Exclusion criteria
Population	A defined population of English-speaking participants with a diagnosis of varicose veins	Undefined population of patients with chronic venous disease or Non-English-speaking patients with varicose veins
Interventions	No intervention or any intervention indicated for varicose veins	
Outcomes	PROMS covering any of the following: generic or preference-based measures e.g. EQ-5D™, SF-6D, SF-36®; directly elicited preference-based measures, e.g. time-trade-off, standard gamble utility values; condition-specific outcome measures; functional outcome measures English version of PROMS	Outcome measures of patient satisfaction or experience, or outcome measures obtained from proxies, carers or health providers Non-English versions of PROMS
Study type	Published validation studies, other than linguistic validation of English versions of relevant PROMS Publication in English	Unpublished studies Studies of linguistic validation of PROMS Review articles, letters, commentaries, abstracts Non-English publications

PROM, patient-reported outcome measure; EQ, EuroQol; SF, Short Form.

Table 2 Psychometric criteria used to assess the quality of the patient-reported outcome measures included in this study

Domain	Criteria
Test–retest reliability	Test–retest: the intraclass correlation/weighted κ score should be ≥ 0.70 for group comparisons and ≥ 0.90 if scores are going to be used for decisions about an individual based on their score[19] The mean difference (paired t test or Wilcoxon signed-rank test) between time points 1 and 2, and the 95% c.i. should also be reported[12,13]
Internal consistency	A Cronbach's α score of ≥ 0.70 is considered good, and it should not exceed ≥ 0.92 for group comparisons as this is taken to indicate that items in the scale could be redundant. Item total correlations should be ≥ 0.20[14,20]
Content validity	This is assessed qualitatively during the development of an instrument. To achieve good content validity, there must be evidence that the instrument has been developed by consulting patients and experts as well as undertaking a literature review[20] Patients should be involved in the development stage and item generation. The opinion of patient representatives should be sought on the constructed scale[12–14]
Construct validity	A correlation coefficient of ≥ 0.60 is taken as strong evidence of construct validity. Authors should make specific directional hypotheses and estimate the strength of correlation before testing[12–14]
Criterion validity	A good argument should be made as to why an instrument is standard and correlation with the standard should be ≥ 0.70[15,16,18,19]
Responsiveness	There are a number of methods to measure responsiveness, including t tests, effect size, standardized response means or responsiveness statistics, Guyatt's responsiveness index. There should be statistically significant changes in score of an expected magnitude[21,22]
Floor and ceiling effects	A floor or ceiling effect is considered if 15% of respondents are achieving the lowest or the highest score on the instrument[12,13]
Acceptability	Acceptability is measured by the completeness of the data supplied; ≥ 80% of the data should be complete[12]

Table 3 Studies reporting validation of patient-reported outcome measures in patients with varicose veins

Reference	Country	Treatment	Type of study	Sample size	Age (years)*	Men (%)	Reported PROM(s)	Timing o PROM(s assessme
Garratt et al.[25]	UK	Usual care	PDVS	373	45.8	24	AVVQ/SF-36®	Administe once
Garratt et al.[26]	UK	Usual care	PDVS	1700	42.7	33.5	SF-36®	2 weeks af baseline
Garratt et al.[27]	UK	Usual care	PDVS	1700	47.9	39.8	SF-36®	Baseline a after 1 ye
Garratt et al.[28]	UK	Usual care	PDVS	373	45.8	46.1	AVVQ/SF-36®	2 weeks a 12 month after basel
Lattimer et al.[29]	UK	EVLA versus UGFS	RCT	100	n.r.	42	AVVQ	Baseline, weeks and months
Lattimer et al.[29]	UK	EVLA versus UGFS	RCT	84	47.5†	47.6	AVVQ	Baseline, weeks and months
Paty et al.[31]	USA	EMA and PEM	RCT	395	49.6	26.78	VVSymQ®	Baseline ar weeks (dai
Shepherd et al.[32]	UK	RFA only	PDVS	317	48.87	28.4	AVVQ, SQOR-V	Baseline ar weeks
Wright et al.[33]	USA	EMA and PEM	RCT	40	49.7	37.5	VVSymQ®	Baseline ar weeks (dai

*Mean values except †median. PROM, patient-reported outcome measure; PDVS, PROM development and validation study; AVVQ, Aberdeen Varicose Vein Questionnaire; SF-36®, 36-Item Short Form Survey; EVLA, endovenous laser ablation; UGFS, ultrasound-guided foam sclerotherapy; n.r., not reported; EMA, endovenous microfoam ablation; PEM, polidocanol endovenous microfoam; VVSymQ®, Varicose Veins Symptoms Questionnaire; RFA, radiofrequency ablation; SQOR-V, Specific Quality-of-life and Outcome Response – Venous.

Table 4 Summary of the psychometric properties of patient-reported outcome measures in patients with VLU

	Psychometric and operational criteria							
Reference	Internal consistency	Test–retest reliability	Content validity	Criterion validity	Construct validity	Responsiveness	Floor/ ceiling effect	Acceptability
Generic PROMS								
SF-36®								
Garratt et al.	+	0	?	0	+/–	+/–	0	+
Garratt et al.[27]	0	0	0	0	0	+	0	+/–
Disease-specific PROMs								
AVVQ								
Garratt et al.	+	0	+/-	+	+	0	0	+/–
Garratt et al.[28]	0	+	0	0	0	+	0	+/–
Shepherd et al.[32]	0	–	0	+	–	+/–	0	+/–
Lattimer et al.[29]	0	0	0	0	0	+	0	0
Lattimer et al.[30]	0	0	0	0	0	?	0	0
VVSymQ®								
Paty et al.[31]	+	0	0	+	0	+	+/–	+
Wright et al.[33]	+	+	+	+/–	+	+	0	+
SQOR-V								
Shepherd et al.[32]	0	–	0	+	–	+/–	0	+/–

0, Not reported (no evaluation completed); –, evidence not in favour; +/-, weak evidence; +, evidence in favour; ?, methodology questionable. PROM, patient-reported outcome measure; SF-36®, 36-Item Short Form Survey; AVVQ, Aberdeen Varicose Vein Questionnaire; VVSymQ®, Varicose Veins Symptoms Questionnaire; SQOR-V, Specific Quality-of-life and Outcome Response –Venous.