



Universiteit
Leiden

The Netherlands

The aftermath of acute pulmonary embolism: approach to persistent functional limitations

Boon, G.J.A.M.

Citation

Boon, G. J. A. M. (2022, March 1). *The aftermath of acute pulmonary embolism: approach to persistent functional limitations*. Retrieved from <https://hdl.handle.net/1887/3277045>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3277045>

Note: To cite this publication please use the final published version (if applicable).



5

Measuring functional limitations after venous thromboembolism: optimization of the Post-VTE Functional Status (PVFS) Scale

G.J.A.M. Boon, S. Barco, L. Bertolotti, W. Ghanima, M.V. Huisman, S.R. Kahn, S. Noble, P. Prandoni, R.P. Rosovsky, A.K. Sista, B. Siegerink, F.A. Klok

ABSTRACT

Introduction: We recently proposed a scale for assessment of patient-relevant functional limitations following an episode of venous thromboembolism (VTE). Further development of this post-VTE functional status (PVFS) scale is still needed.

Methods: Guided by the input of VTE experts and patients, we refined the PVFS scale and its accompanying manual, and attempted to acquire broad consensus on its use.

Results: A Delphi analysis was performed involving 53 international VTE experts with diverse scientific and clinical backgrounds. In this process, the number of scale grades of the originally proposed PVFS scale was reduced and descriptions of the grades were improved. After these changes, a consensus was reached on the number/definitions of the grades, and method/timing of the scale assessment. The relevance and potential impact of the scale was confirmed in three focus groups totaling 18 VTE patients, who suggested additional changes to the manual, but not to the scale itself. Using the improved manual, the κ -statistics between PVFS scale self-reporting and its assessment via the structured interview was 0.75 (95%CI 0.58-1.0), and 1.0 (95%CI 0.83-1.0) between independent raters of the recorded interview of 16 focus groups members.

Conclusion: We improved the PVFS scale and demonstrated broad consensus on its relevance, optimal grades, and methods of assessing among international VTE experts and patients. The interobserver agreement of scale grade assignment was shown to be good-to-excellent. The PVFS scale may become an important outcome measure of functional impairment for quality of patient care and in future VTE trials.

INTRODUCTION

Along with recurrent venous thromboembolism (VTE), bleeding and mortality, both quality of life and functional limitations are important outcome measures that matter greatly to VTE patients.¹⁻⁷ These two issues are mostly determined by the presence and severity of leg or chest pain, dyspnea, anxiety, post thrombotic panic syndrome, and depression, all of which are important determinants of recovery after VTE.^{4-6,8-21} Persistent functional limitations and/or decreased quality of life after an episode of VTE are framed within the concepts of the post-thrombotic syndrome (PTS) and post-pulmonary embolism (PE) syndrome.^{2,22} PTS manifests as chronic venous insufficiency in a limb affected by deep vein thrombosis (DVT) caused by venous obstruction and/or venous reflux.^{22,23} The post-PE syndrome includes chronic thromboembolic pulmonary hypertension (CTEPH), chronic thromboembolic pulmonary vascular disease (CTED), and persistent right ventricular dysfunction, but also otherwise unexplained new or progressive exercise intolerance after PE.^{2,24-26}

While there are validated questionnaires for the assessment of pain, dyspnea, anxiety and depression, these were mostly not designed to rank patients into meaningful categories and do not target functional outcomes *per se*. The same holds true for measures of quality of life.^{5,14,23,24,27,28} Therefore, quality of life outcomes between treatment strategies are difficult to put into a comprehensible perspective and may not always serve their purpose when used as a main outcome measure in the setting of an experimental study. Given these gaps, we recently proposed the first version of the post-VTE functional status (PVFS) scale, which is meant to be used as a comprehensive measure to quantify the consequences of VTE on functional status. It covers the full spectrum of functional outcomes, ranging from no symptoms to death, and focuses on both limitations in usual duties/activities and changes in lifestyle (**Appendix A**).¹¹ The scale, required further development (**Table 1**). In the current study, we sought to refine the scale and acquire broad consensus on the methods of assigning a PVFS scale grade at a certain time point, guided by the expertise of VTE specialists as well as by patient focus groups. Moreover, we aimed to establish the reproducibility of PVFS scale assessment.

METHODS

Study design

The Delphi method was used to assess expert consensus among a panel of VTE specialists with scientific and clinical expertise in measuring and managing long-term outcomes of VTE across different patient subgroups. A Delphi analysis is a widely used structured

Table 1: Proposed steps of development of performance outcome instruments, and current status of development of the post-VTE functional status scale

Steps of development	Status of the PVFS scale
1 Review of evidence for functional scales and tools assessing functional status in patients with VTE (<i>first publication</i>)	✓
2 Identification of the key characteristics of the modified Rankin Scale for patients with stroke in order to draft measure and item specifications, and fields of applicability which may be relevant for patients with venous thromboembolism (<i>first publication</i>)	✓
3 Assemble a dedicated multidisciplinary work group (including patients, and physicians, nurses, and representatives of major societies) to achieve consensus on the instrument (<i>current manuscript</i>)	✓
4 Formal rounds of review of the proposed categories of the ordinal scale from the dedicated multidisciplinary work group (<i>current manuscript</i>)	✓
5 Formal assessment of reliability and validity of the scale (<i>current manuscript</i>)	✓
6 Next research topics: - Formal assessment of reliability and feasibility (e.g. logistics and costs) of the scale in clinical trials - Formal evaluations of assessment methods; blinded versus non-blinded raters, and structured interview versus self-report - Assessment of interrater agreement of structured interviews after translation into other languages - Assessment of variability in time of the PVFS scale grades following the intended time points for assessing functional status - Relating quality of life and utilities to functional status, with focus on cultural differences	Ongoing
7 Dissemination and implementation in both research protocols and clinical practice for routinely collected data analyses (quality indicator)	Ongoing

process to achieve consensus through opinions and feedback from a group of informed experts.²⁹⁻³¹ The process is anonymous and focuses on a predefined dilemma, usually lacking empirical evidence. The expert panel is consulted via questionnaires. Responses are analyzed and used to guide the next round of new questions. In the subsequent rounds of questions, experts are provided with the overall results of the previous round of the Delphi study. The optimal level of consensus is usually achieved after two or more rounds.³²

A patient focus group represents a qualitative research method in which a small group of participants discuss a topic chaired by a moderator.³³ It is an appropriate and suitable method to involve patients in the development of scales and other medical instruments for various medical conditions. Following a number of semi-structured questions to help focus the group's discussion, participants can explore issues of concern, propose changes and identify strategies for further exploration of the topic at hand.

In both stages of the study, we aimed to achieve consensus and explore the experts' and patients' view on six issues: 1) the relevance of measuring functional limitations after VTE both in clinical trials and in clinical practice, 2) whether current available tools are sufficiently reliable to assess functional limitations after VTE, 3) the appropriateness of the PVFS scale for measuring functional limitations after VTE, 4) whether the PVFS scale has sensible and representative scale categories, 5) whether the PVFS scale captures both PTS and post-PE syndrome, and 6) how and when the scale should be measured.

Selection of experts and patients

Experts were selected based on the following criteria: 1) leaders in the VTE field, as demonstrated by a strong publication track record and leading roles in scientific societies; or 2) clinical experience in treating patients with post-VTE complications. The panel of physicians and epidemiologists was selected to represent a wide geographic area, to include both sexes, and to cover broad medical specialties, including (pediatric) hematology, cardiology, pulmonology, vascular medicine and vascular surgery. The experts completed the questionnaires anonymously and were unaware of the identity of the other experts involved. Patients were invited to participate in the focus groups via mailings from the Netherlands Thrombosis Foundation (patient association). As with the panel, we aimed to include patients who do or did experience functional limitations after their VTE diagnosis, rather than to represent the whole of VTE patients.. Therefore, we did not apply any selection criteria except for consenting to participation, accepting that selection bias would occur.

Delphi and focus group processes

A multinational steering committee of four members was established to oversee the process (DB, SB, BS and FK). A first version of the Delphi questionnaire was drafted by two members (DB and FK). All members of the steering committee provided feedback on the questionnaire and approved its final version. The first round of our Delphi study consisted of a total of 10 multiple choice questions/statements (**Appendix B**). Each question included a free text box for further elaboration. Additionally, a final open question was included, which allowed the experts to provide any input about the PVFS scale, including its design, and how it may be used. Subsequent rounds were planned until consensus was reached, which had been predefined as a minimum level of agreement of 70%, in line with previous Delphi reports.³⁴⁻³⁷ The Delphi questionnaire was distributed by using an online survey tool (Google Forms). Responses were filed at the experts' discretion until a given deadline date, a total of two reminders were sent before reaching this date.

Table 2: Semi-structured interview questions for the patients focus groups

Engagement questions	Explorative questions	Exit question
Can you tell about yourself and your experience with VTE?	Do you consider assessment of functional limitations after VTE to be relevant? Was this assessed in your case, and how?	Is there anything else that you would like to say about the PVFS scale?
How did the VTE influence your physical and mental condition?	What do you think of the PVFS scale and how can it be used?	
How did the VTE influence your working, family, and social life?	Are the scale categories straightforward and rational? Are any important aspects of functional outcomes after VTE missing?	
Have you been/are you functionally impaired by the VTE?	Is the wording of the scale categories adequate and comprehensible? Is the interview (and corresponding manual) adequate? Is the patient self-report flowchart (and corresponding manual) adequate? How often is the PVFS scale to be assessed, and would you prefer self-report over a structured interview? Do you think it should be implemented in future clinical trials on optimal treatment of VTE?	

For the focus groups, semistructured interview questions were developed by the steering committee. The questions were divided into three parts, namely engagement questions, exploration questions, and exit questions (**Table 2**). The two-hour meetings were conducted in Dutch and chaired by two members of the steering committee. All participants agreed to a voice-recording of the meeting. Extraction of all relevant suggestions and remarks that needed to be implemented in the PVFS scale or its manual was done by the steering committee.

Interrater agreement

Interobserver variability was determined by comparing the self-reported PVFS scale grade to the scale grade identified via the structured interview. One member of the steering committee interviewed the patients from the focus groups in a standardized way according to the updated manual for the structured interview (**Appendix C**). These interviews were recorded on tape. Before the interview was conducted, patients were given the self-report flowchart and corresponding table and were given instructions on how to determine their PVFS scale grade (**Appendix C**). Upon completion of the

interview, the interviewer noted the identified PVFS scale grade, independently and blinded to the scoring determined by the patient. Two additional raters -independently from the interviewer and patient and blinded to both of their ratings- reevaluated the recording and assigned the patients to a scale grade in accordance with the manual for the structured interview.

Analysis

Descriptive statistics and interobserver variability were calculated using SPSS version 25.0.0 (SPSS, IBM). The interobserver agreement of PVFS scale assessment between the patients and the raters as well as between the raters themselves was assessed by calculating the κ -statistic. The κ -statistic was interpreted as follows: poor (<0.20), fair (0.21–0.40), moderate (0.41–0.60), good (0.61–0.80) or excellent (0.81–1.00).³⁸

RESULTS

Delphi analysis

In April 2019, the first round of the Delphi questionnaire was distributed among 70 international VTE experts. Of those, 53 (76%) consented to participate and completed the questionnaire. The expert panel included VTE specialists practicing in 15 different countries. A total of 48 out of 53 participants completed the full Delphi procedure. Among them, 17 were women. Several specialties were represented, including (pediatric) hematology (n=21), vascular medicine/surgery (n=13), pneumology (n=10), cardiology (n=3), radiology (n=1), psychologists (n=2) and clinical epidemiology (n=3).

Following compilation of the first Delphi round, immediate consensus was reached on 5 of 10 questions/statements. The vast majority agreed that measuring functional outcomes after VTE was relevant for both research purposes (98%) and clinical practice (96%). Also, the panel considered current tools (among others Villalta Score, Venous Clinical Severity Score, New York Heart Association Classification, Modified Medical Research Council Dyspnea Scale, 6-minute walk distance, pulmonary function test and quality of life measures) unsatisfactory reliable for both research purposes (73%) and clinical practice (83%). They agreed that the PVFS scale would to be a potentially helpful tool for these purposes (79% and 71%, respectively). Further, the experts considered the design of the PVFS scale to be optimal (84%) and its manual to be clear and complete (78%).

The Delphi panel, however, clearly indicated that the originally proposed 7-level ordinal scale needed to be improved in regards to two main points: 1) the score should better reflect functional impairment related to PTS and 2) the categories needed to be more distinctive. Moreover, the impact of anxiety and depression needed

to be addressed more explicitly. All comments were discussed within the steering committee, which led to modifications to the scale and manual. Specifically, the following adjustments were made (in addition to linguistic tweaking): 1) the scale was adjusted to be more sensitive to DVT-associated functional limitations by replacing 'symptoms/discomfort' to 'symptoms, pain, or anxiety' in the grade description (in this way, psychological aspects of physical functioning were incorporated as well); 2) specific symptoms/signs (such as dyspnea at rest or venous ulceration) were removed to avoid measuring symptoms rather than their functional impact; 3) one scale grade was removed ("moderately-severe functional limitations") to facilitate distinctive grades in the middle spectrum of severity, and 4) 'death' was considered as a 'D' class instead of 'grade 6' to make it more visually distinctive.

The updated scale and manual were sent out for a second Delphi questionnaire round; the same respondents were able to see which of the multiple-choice options of the first round achieved the highest level of consensus, and how and why adjustments had been made. The second questionnaire consisted of four statements, all of which

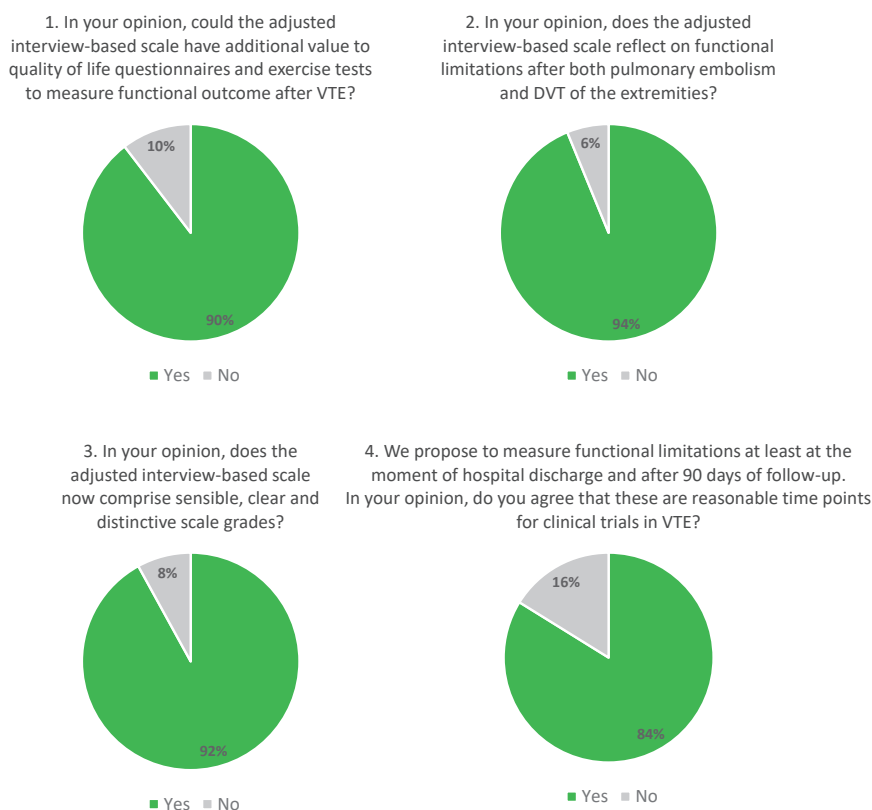


Figure 1: Statements reaching consensus in the second round of the Delphi analysis

achieved consensus (**Figure 1**): 89% of 48 respondents agreed that the adjusted scale had additional value to quality of life questionnaires and exercise tests to measure functional outcome after VTE; 94% agreed that the adjusted scale reflected functional limitations after both DVT and PE; 92% agreed that the adjusted scale comprised sensible, clear and distinctive scale grades, and 83% agreed that the scale could best be assessed at the time of VTE diagnosis and after 90 days, leaving explicit room for longer follow-up depending on the clinical setting or objective of a clinical trial. The respondents advised to mention the possibility of measuring the pre-VTE functional status, which was then incorporated as an optional item in order to obtain a true picture of change in functional status after the VTE event.

Patient focus groups

A total of 18 patients responded to the invitation and participated in one of three patient focus group sessions. Their ages ranged between 21 and 70 years, and two were men. Their background was diverse: the group included physicians, nurses, teachers, journalists and housecleaners. Several patients had administrative jobs, and one was retired at the time of the VTE diagnosis. Most had been diagnosed with both DVT and acute PE (n=9), and 11 had recurrent VTE. None of the patients were known with CTEPH or severe PTS with leg ulcers. The mean time since last VTE event varied from 6 to 108 months, with most patients (67%) between 0 and 2 years after their event.

All participants highlighted their appreciation of focused attention to functional limitations after VTE. All 18 considered it a relevant topic as they had (and most of them still) suffered greatly from the long-term impact of their VTE in regards to their professional and personal life. Several had lost their jobs or had to reduce the intensity of their work. Furthermore, several marriages and relations were broken, and most of them considered the VTE a traumatic experience. The quote "I had a different life before than after the VTE diagnosis" by one of the focus group participants was heartedly endorsed by the other participants. Many still faced anxiety due to the possibility of recurrent events. In general, they recognized the lack of attention from their treating physician to aspects of their recovery other than management of the anticoagulant treatment. The lack of a status scale such as the PVFS scale was agreed on to be an unmet clinical need. The general consensus was that the introduction of the PVFS scale could help address persistent functional limitations in the (outpatient) clinic, but also to help explain their functional status to their families and relatives. Also, the patients generally agreed that the scale reflected functional limitations after both DVT and PE. Most patients could imagine self-reporting the scale via a mobile application, at fixed time points during their follow-up care but also on their own initiative to better capture good and/or bad weeks. This latter option would give patients a sense of more control of their treatment.

Table 3: Final post-VTE functional status scale as agreed upon by the Delphi panel and patient focus groups

PVFS scale grade	Description
0	No functional limitations All usual duties/activities at home or at work can be carried out at the same level of intensity. Symptoms, pain and anxiety are absent.
1	Negligible functional limitations All usual duties/activities at home or at work can be carried out at the same level of intensity, despite some symptoms, pain, or anxiety.
2	Slight functional limitations Some usual duties/activities at home or at work are carried out at a lower level of intensity or are occasionally avoided due to symptoms, pain, or anxiety.
3	Moderate functional limitations Usual duties/activities at home or at work have been structurally modified (reduced) due to symptoms, pain, or anxiety.
4	Severe functional limitations Assistance needed in activities of daily living due to symptoms, pain, or anxiety: nursing care and attention are required.
D	Death Death occurred before the scheduled assessment.

Note: Providing a reference value (pre-VTE grade) is optional and should refer to the functional status 1 month prior to the VTE diagnosis. Full manual for structured interview and patient self-report is provided in **Appendix C**.

At least half of patients reported that they had some reservations toward the self-report flowchart and table. The main concern raised was with the distinction between moderate and severe functional limitations (scale grade 2 and 3). Textual changes were suggested to make it clearer that grade 2 involves being able to do all ones' duties/activities, even at a slower pace or extended over a longer period of time, and that grade 3 indicates the inability to perform a particular duty/activity. The manual to the structured interview was adjusted accordingly. Moreover, it was suggested not to actually provide a 'grade name' to the limitations themselves (e.g. moderate limitations), because of its subjectiveness, but rather, to stick to describing the limitations. The scale (after adjustments suggested and approved by the Delphi panel) was considered to be adequate, and no further changes were suggested or incorporated by the patients. The final scale is shown in **Table 3**; the final patient self-report flowchart and corresponding table are shown in **Figure 2** and **Table 4**.

Interrater agreement

Structured interviews were conducted and recorded for 16 focus group participants. There was full agreement between the patient self-reported scale grade and the grade assigned by the interviewer in 14/16 patients (88%) for a kappa statistic of 0.75 (95%CI 0.58-1.0). The two discrepancies were in patients who rated themselves a grade 2 while the interviewer categorized them as grade 3. Two independent raters blinded to the grading by the patient and the interviewer evaluated the recorded interview post-hoc, were both in full agreement with each other and with the interviewer, for a kappa statistic of 1.0 (95%CI 0.83-1.0).

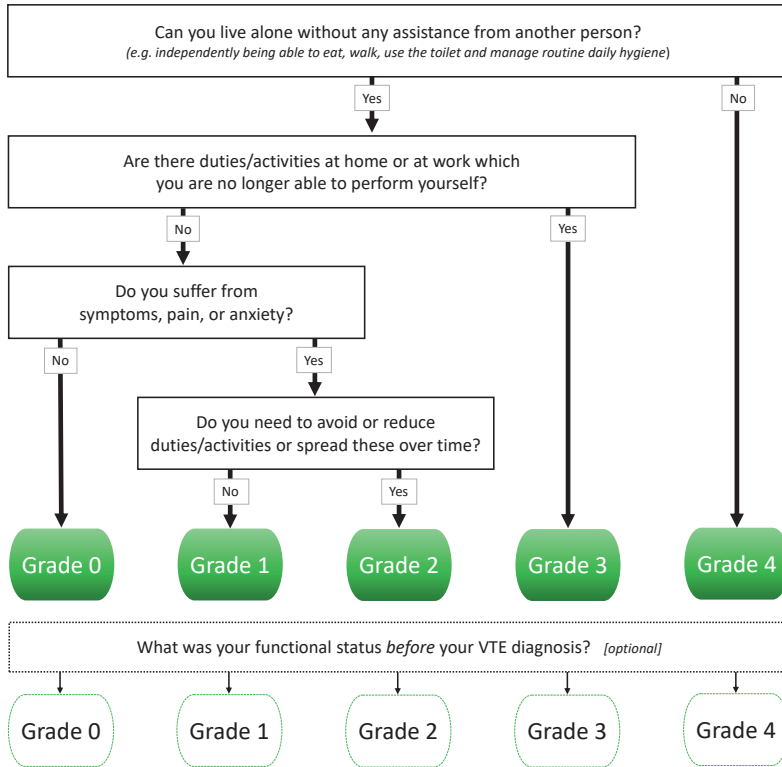


Figure 2: Flowchart for patient self-report of the post-VTE functional status scale (full manual for structured interview and patient self-report is provided in **Appendix C**)

Table 4: Table accompanying the flowchart for patient self-report of the post-VTE functional status scale

How much are you currently affected in your everyday life by the VTE? Please indicate which one of the following statements applies to you most. Please tick only one box at a time.	<input type="checkbox"/>	Corresponding PVFS scale grade if the box is ticked
I have no limitations in my everyday life and no symptoms, pain, or anxiety related to the VTE.	<input type="checkbox"/>	0
I have negligible limitations in my everyday life as I can perform all usual duties/activities, although I still have persistent symptoms, pain, or anxiety.	<input type="checkbox"/>	1
I suffer from limitations in my everyday life as I occasionally need to avoid or reduce usual duties/activities or need to spread these over time due to symptoms, pain, or anxiety. I am, however, able to perform all activities without any assistance.	<input type="checkbox"/>	2
I suffer from limitations in my everyday life as I am not able to perform all usual duties/activities due to symptoms, pain, or anxiety. I am, however, able to take care of myself without any assistance.	<input type="checkbox"/>	3
I suffer from severe limitations in my everyday life: I am not able to take care of myself and therefore I am dependent on nursing care and/or assistance from another person due to symptoms, pain, or anxiety.	<input type="checkbox"/>	4

Note: Full manual for structured interview and patient self-report is provided in **Appendix C**.

DISCUSSION

The concept of the PVFS scale was endorsed by a large panel of international VTE experts as well as by VTE patients of diverse ages and backgrounds. The results of our Delphi consensus indicate that the PVFS scale may help to address an unmet clinical need. Moreover, we were able to improve the original proposed scale, refine the optimal assessment method, and establish good-to-excellent interobserver agreement between different medical professionals as well as between the structured interview and the patient self-report. The development of the scale is in an advanced stage, and it can be now used in clinical practice and implemented in clinical trials (**Table 1**).

In the field of stroke research, the modified Rankin Scale (mRS) -by which the authors were inspired when proposing and developing the PVFS scale- has achieved a key position as an important, and often primary, outcome of seminal trials that have subsequently shaped the current standard of care^{39,40}. With the introduction of the PVFS scale, VTE trials can now start to include an overall outcome measure that captures the broad range of physical and psychological long-term complications of VTE and its treatment, expressed in meaningful categories that are linked to quality of life with both social and economic impact (e.g. healthcare costs and societal costs). This scale could be used to identify patients with slower than expected recovery after VTE. In such cases, the culprit symptom can be identified and targeted, although this latter is beyond the scope of the scale itself. Another example of how the score may be used is to establish the optimal duration of treatment in unprovoked VTE, in which the balance between the impact of relatively frequent recurrences of VTE and less frequent -but more impactful- occurrences of bleeding complications is still a matter of debate and research.⁴¹⁻⁴⁴ A third example where the PVFS scale may help to determine conclusively the optimal treatment strategy is the dilemma of the benefit of early pharmacomechanical catheter-directed thrombolysis in iliofemoral DVT, which has been associated with better quality of life in some studies, but not with less PTS.^{45,46} Based on our results presented here, the PVFS scale has been included as a secondary outcome in four clinical trials scheduled to start on short notice: PEITHO-3 (PHRCN_16-0580), SAFE-SSPE (NCT04263038), L-TRRiP (ZonMw 848017007) and ARIVA (NCT04128956). We will learn more about the value of the PVFS scale by analysing the results of the PVFS scale assessments within those and other clinical trials on correlation between PVFS scale grades and health-related quality of life or more specific components of PTS and the post-PE syndrome (e.g. pain or dyspnea). In the clinical setting, the PVFS scale may help health care providers to monitor the recovery of individual patients, may better identify complications such as CTEPH at an early stage and help explaining the impact of persistent symptoms to the treating physician and relatives.

The main strength of this study is the broad consensus reached among a large sample of international experts from various backgrounds, as well as the clear endorsement of the PVFS scale by the patient focus groups. The scale and its manual were refined to a point with good-to-excellent interrater agreement, either self-reported or assessed via a structured interview, underlining its validity when used as an outcome of clinical trials. Although patient self-reporting is probably the most practical approach for collecting PVFS scale outcome data, the structured interview is the preferred mode of assessment from a scientific point of view until the value of self-reported data has been established by future research.

This study is limited by an obvious selection bias in the patient focus groups: only patients with moderate to severe presentations of PTS and/or post-PE syndrome responded to our call, leading to a probable overestimation of the relevance for VTE patients as a whole group. Of note, this bias is inherent to any outcome measure since generally the majority of patients would never meet a given endpoint. Another potential limitation lies in using focus groups to obtain qualitative data. There is a risk of data bias if more forthright participants dominate the discussions. This risk was ameliorated by using an experienced and unprejudiced researcher to facilitate discussions and ensure all participants were involved. Also, while we involved a broad international panel of VTE experts and developed the first versions of the scale in the English language, only Dutch patients participated in the focus groups and the interobserver assessment was only tested in the Dutch language. Hence, the scale and its manual for the structured interview and patient self-report need to be evaluated in other countries and languages. Still, because of the simplicity of the PVFS scale, we do not expect much different scale performance or interobserver agreement when translated into other languages.

In conclusion, we demonstrated broad consensus on the relevance and methods of assessment of the PVFS scale among international VTE experts and patients. Based on their comments and suggestions, the scale and its manual were improved after which the interobserver agreement of scale assessment in our study was good-to-excellent. These findings suggest that the PVFS scale can be integrated as a relevant outcome measure in future clinical trials as well as in daily clinical practice to monitor patient recovery.

REFERENCES

1. Klok FA, Zondag W, van Kralingen KW, et al. Patient outcomes after acute pulmonary embolism. A pooled survival analysis of different adverse events. *American Journal of Respiratory and Critical Care Medicine* 2010; 181(5): 501-6.
2. Klok FA, van der Hulle T, den Exter PL, Lankeit M, Huisman MV, Konstantinides S. The post-PE syndrome: a new concept for chronic complications of pulmonary embolism. *Blood Reviews* 2014; 28(6): 221-6.
3. Sista AK, Miller LE, Kahn SR, Kline JA. Persistent right ventricular dysfunction, functional capacity limitation, exercise intolerance, and quality of life impairment following pulmonary embolism: Systematic review with meta-analysis. *Vascular Medicine* 2017; 22(1): 37-43.
4. Klok FA, van Kralingen KW, van Dijk AP, et al. Quality of life in long-term survivors of acute pulmonary embolism. *Chest* 2010; 138(6): 1432-40.
5. Kahn SR, Ducruet T, Lamping DL, et al. Prospective evaluation of health-related quality of life in patients with deep venous thrombosis. *Archives of Internal Medicine* 2005; 165(10): 1173-8.
6. Hunter R, Noble S, Lewis S, Bennett P. Long-term psychosocial impact of venous thromboembolism: a qualitative study in the community. *BMJ Open* 2019; 9(2): e024805.
7. Huisman MV, Barco S, Cannegieter SC, et al. Pulmonary embolism. *Nature Reviews Disease Primers* 2018; 4: 18028.
8. Lukas PS, Krummenacher R, Biasiutti FD, Begre S, Znoj H, von Kanel R. Association of fatigue and psychological distress with quality of life in patients with a previous venous thromboembolic event. *Thrombosis and Haemostasis* 2009; 102(6): 1219-26.
9. Liu CP, Li XM, Chen HW, et al. Depression, anxiety and influencing factors in patients with acute pulmonary embolism. *Chinese Medical Journal* 2011; 124(16): 2438-42.
10. Keller K, Tesche C, Gerhold-Ay A, et al. Quality of life and functional limitations after pulmonary embolism and its prognostic relevance. *Journal of Thrombosis and Haemostasis* 2019; 17(11): 1923-1934.
11. Klok FA, Barco S, Siegerink B. Measuring functional limitations after venous thromboembolism: A call to action. *Thrombosis Research* 2019; 178: 59-62.
12. Kachroo S, Boyd D, Bookhart BK, et al. Quality of life and economic costs associated with postthrombotic syndrome. *American Journal of Health-System Pharmacy* 2012; 69(7): 567-72.
13. van Es J, den Exter PL, Kaptein AA, et al. Quality of life after pulmonary embolism as assessed with SF-36 and PEmb-QoL. *Thrombosis Research* 2013; 132(5): 500-5.
14. Hogg K, Kimpton M, Carrier M, Coyle D, Forgie M, Wells P. Estimating quality of life in acute venous thrombosis. *JAMA Internal Medicine* 2013; 173(12): 1067-72.
15. Tavoly M, Utne KK, Jelsness-Jorgensen LP, et al. Health-related quality of life after pulmonary embolism: a cross-sectional study. *BMJ Open* 2016; 6(11): e013086.
16. Lubberts B, Paulino Pereira NR, Kabrhel C, Kuter DJ, DiGiovanni CW. What is the effect of venous thromboembolism and related complications on patient reported health-related quality of life? A meta-analysis. *Thrombosis and Haemostasis* 2016; 116(3): 417-31.
17. Kahn SR, Akaberi A, Granton JT, et al. Quality of Life, Dyspnea, and Functional Exercise Capacity Following a First Episode of Pulmonary Embolism: Results of the ELOPE Cohort Study. *American Journal of Medicine*. 2017; 130(8): 990.e9-.e21.
18. Kahn SR, Hirsch AM, Akaberi A, et al. Functional and Exercise Limitations After a First Episode of Pulmonary Embolism: Results of the ELOPE Prospective Cohort Study. *Chest* 2017; 151(5): 1058-68.

19. Sista AK, Klok FA. Late outcomes of pulmonary embolism: The post-PE syndrome. *Thrombosis Research* 2018; 164: 157-62.
20. Tavoly M, Wik HS, Sirnes PA, et al. The impact of post-pulmonary embolism syndrome and its possible determinants. *Thrombosis Research* 2018; 171: 84-91.
21. Hunter R, Lewis S, Noble S, Rance J, Bennett PD. "Post-thrombotic panic syndrome": A thematic analysis of the experience of venous thromboembolism. *British Journal of Health Psychology* 2017; 22(1): 8-25.
22. Kahn SR, Comerota AJ, Cushman M, et al. The postthrombotic syndrome: evidence-based prevention, diagnosis, and treatment strategies: a scientific statement from the American Heart Association. *Circulation* 2014; 130(18): 1636-61.
23. Wik HS, Enden TR, Ghanima W, Engeseth M, Kahn SR, Sandset PM. Diagnostic scales for the post-thrombotic syndrome. *Thrombosis Research* 2018; 164: 110-5.
24. Klok FA, Barco S. Follow-up after acute Pulmonary Embolism. *Hamostaseologie* 2018; 38(1): 22-32.
25. Klok FA, Couturaud F, Delcroix M, Humbert M. Diagnosis of chronic thromboembolic pulmonary hypertension after acute pulmonary embolism. *European Respiratory Journal* 2020;55(6):2000189.
26. Klok FA, Delcroix M, Bogaard HJ. Chronic thromboembolic pulmonary hypertension from the perspective of patients with pulmonary embolism. *Journal of Thrombosis and Haemostasis* 2018; 16(6): 1040-51.
27. Klok FA, Cohn DM, Middeldorp S, et al. Quality of life after pulmonary embolism: validation of the PEmb-QoL Questionnaire. *Journal of Thrombosis and Haemostasis* 2010; 8(3): 523-32.
28. Ghanima W, Wik HS, Tavoly M, Enden T, Jelsness-Jorgensen LP. Late consequences of venous thromboembolism: Measuring quality of life after deep vein thrombosis and pulmonary embolism. *Thrombosis Research* 2018; 164: 170-6.
29. Jones J, Hunter D. Consensus methods for medical and health services research. *British Medical Journal* 1995; 311(7001): 376-80.
30. Humphrey-Murto S, Varpio L, Wood TJ, et al. The Use of the Delphi and Other Consensus Group Methods in Medical Education Research: A Review. *Academic Medicine* 2017; 92(10): 1491-8.
31. Waggoner J, Carline JD, Durning SJ. Is There a Consensus on Consensus Methodology? Descriptions and Recommendations for Future Consensus Research. *Academic Medicine* 2016; 91(5): 663-8.
32. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing* 2000; 32(4): 1008-15.
33. Kitzinger J. Qualitative research. Introducing focus groups. *British Medical Journal* 1995; 311(7000): 299-302.
34. Mahler DA, Selecky PA, Harrod CG, et al. American College of Chest Physicians consensus statement on the management of dyspnea in patients with advanced lung or heart disease. *Chest* 2010; 137(3): 674-91.
35. Blasi F, Concia E, Del Prato B, et al. The most appropriate therapeutic strategy for acute lower respiratory tract infections: a Delphi-based approach. *Journal of Chemotherapy* 2017; 29(5): 274-86.
36. Zafar SY, Currow DC, Cherny N, Strasser F, Fowler R, Abernethy AP. Consensus-based standards for best supportive care in clinical trials in advanced cancer. *Lancet Oncology* 2012; 13(2): e77-82.

37. Brown AK, O'Connor P J, Roberts TE, Wakefield RJ, Karim Z, Emery P. Recommendations for musculoskeletal ultrasonography by rheumatologists: setting global standards for best practice by expert consensus. *Arthritis and Rheumatism* 2005; 53(1): 83-92.
38. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin* 1968; 70(4): 213-20.
39. Hacke W, Kaste M, Bluhmki E, et al. Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke. *The New England Journal of Medicine* 2008; 359(13): 1317-29.
40. Berkhemer OA, Fransen PS, Beumer D, et al. A randomized trial of intraarterial treatment for acute ischemic stroke. *The New England Journal of Medicine* 2015; 372(1): 11-20.
41. Konstantinides SV, Meyer G, Becattini C, et al. 2019 ESC Guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS). *European Respiratory Journal* 2019; 54(3): 1901647.
42. Klok FA, Hosel V, Clemens A, et al. Prediction of bleeding events in patients with venous thromboembolism on stable anticoagulation treatment. *European Respiratory Journal* 2016; 48(5): 1369-76.
43. van der Wall SJ, van der Pol LM, Ende-Verhaar YM, et al. Fatal recurrent VTE after anticoagulant treatment for unprovoked VTE: a systematic review. *European Respiratory Review* 2018; 27(150).
44. Khan F, Rahman A, Carrier M, et al. Long term risk of symptomatic recurrent venous thromboembolism after discontinuation of anticoagulant treatment for first unprovoked venous thromboembolism event: systematic review and meta-analysis. *British Medical Journal* 2019; 366: l4363.
45. Kahn SR, Julian JA, Kearon C, et al. Quality of life after pharmacomechanical catheter-directed thrombolysis for proximal deep venous thrombosis. *Journal of Vascular Surgery Venous and Lymphatic Disorders* 2020; 8(1): 8-23.e18.
46. Vedantham S, Goldhaber SZ, Julian JA, et al. Pharmacomechanical Catheter-Directed Thrombolysis for Deep-Vein Thrombosis. *The New England Journal of Medicine* 2017; 377(23): 2240-52.

Appendix A: Originally proposed post-VTE functional status scale ¹¹

Scale category	Description
0 No functional limitations	No symptoms/discomfort.
1 No significant functional limitations	Able to carry out all usual activities at the same level of intensity, despite some symptoms/discomfort.
2 Slight functional limitations	Some usual activities are carried out at a lower level of intensity due to symptoms/discomfort.
3 Moderate functional limitations	Able to take care of activities of daily living unassisted, but with modified duties at work/activities at home.
4 Moderate to severe functional limitations	Requires help for activities of daily living or cannot work; breathlessness at rest.
5 Severe functional limitations	Unable to carry on activities of daily living without assistance, breathlessness at rest. Requires nursing care and attention.
6 Death	Death

Note: The Post-VTE functional status scale is assessed during a short structured interview with the patient, either by phone or in the office. Breathlessness at rest is comparable to NYHA class IV and differentiates between classes 1-3 and class 4-5. For class 1 and 2, it is intended to assess whether usual activities *could* be carried out and refers to “usual activities” as any activity that patients used to undertake on a monthly basis or more frequently. Class 3 is intended to capture *any* VTE-associated reason for modification of the usual duties at work/at home. For class 4, it should be indicated that assistance with some activities of daily living, e.g. preparing a meal, personal/toilet hygiene, functional mobility or basic household chores, is *absolutely essential*. For class 5, it should be indicated that assistance with *most/all* activities of daily living is *absolutely essential*. Nursing care here not necessarily administered by a certified nurse, but indicates that patients cannot be left alone for longer periods of time. *When in doubt between two adjacent alternatives on the scale even despite that both appearing equally valid, the worse option should be chosen.*

Appendix B: Questions included in the first round of the Delphi analysis

1. In your opinion, is measuring functional outcome after VTE a relevant outcome?
 - o In studies: yes or no
 - o In clinical practice: yes or noMotivation: (free text)
2. Are current available tools to measure functional outcome after VTE (e.g. questionnaires, pulmonary function tests, exercise tests) sufficiently reliable? If yes, please motivate which tools.
 - o In studies: yes or no
 - o In clinical practice: yes or noMotivation: (free text)
3. Is the proposed scale a potentially helpful tool to measure functional outcome after VTE?
 - o In studies: yes or no
 - o In clinical practice: yes or noMotivation: (free text)
4. Could/should the proposed interview based scale replace existing instruments to measure functional outcome after VTE?
 - o Yes, it could
 - o Yes, it should
 - o NoMotivation: (free text)
5. Is the design of the proposed scale, e.g. ordinal scale, optimal for its purpose?
 - o Yes or no; motivation: (free text)
6. Does the proposed scale include sensible and representative scale categories? If not, please elaborate how this can be improved (descriptions will be discussed in later questions).
 - o Yes or no; motivation: (free text)
7. Does the proposed scale reflect on functional limitations after both pulmonary embolism and deep vein thrombosis of the extremities? If not, please elaborate how this can be improved.
 - o Yes or no; motivation: (free text)
8. Are the proposed categories distinctive and do they not cause any confusion in selecting the appropriate category for a specific patient? If not, please elaborate how this can be improved.
 - o Yes or no; motivation: (free text)
9. Are the descriptions of the proposed scale categories clear and complete? If not, please motivate for each category why you suggest improvements.
 - o For each separate category: yes or no; motivation: (free text)
10. Please read the user instructions of the scale below. Is this explanation clear and complete?
 - o Yes or no; motivation: (free text)
11. Do you have any further comments on this topic for us to consider? (free text)

Appendix C: Manual to the post-VTE functional status scale for physicians and study personnel – including corresponding structured interview and assessment tools

Please find the full manual on <https://osf.io/qgpdv/>