

# Development and measurement properties of the PEM/PESE activity questionnaire (PAQ)

Todd E. Davenport<sup>a,b,\*</sup>, Staci R. Stevens<sup>b</sup>, Jared Stevens<sup>b</sup>, Christopher R. Snell<sup>b</sup> and J. Mark Van Ness<sup>a,b</sup>

<sup>a</sup>*University of the Pacific, Stockton, CA, USA*

<sup>b</sup>*Workwell Foundation, Ripon, CA, USA*

Received 27 September 2022

Accepted 20 January 2023

## Abstract.

**BACKGROUND:** Existing instruments often are inappropriate to measure the effects of post-exertional malaise (PEM) and post-exertional symptom exacerbation (PESE) on activities of daily living (ADLs). A validated questionnaire to measure self-reported ability with ADLs would advance research and clinical practice in conditions like myalgic encephalomyelitis and Long Covid.

**OBJECTIVE:** Determine the measurement properties of the PEM/PESE Activity Questionnaire (PAQ).

**METHODS:** The PAQ is adapted from the Patient Specific Functional Scale. Respondents rated three self-selected ADLs on two 0-100 scales, including current performance compared to (1) a ‘good day’ and (2) before illness. Respondents provided a Burden of Functioning rating on a 0-100 scale, anchored at 0 being the activity took “No time, effort, and resources at all” and 10 being “All of my time, effort, and resources.” Respondents took the PAQ twice, completing a demographic questionnaire after the first PAQ and before the second PAQ. Descriptive statistics and intraclass correlation coefficients were calculated for each scale to assess test-retest reliability. Minimum detectable change outside the 95% confidence interval (MDC<sub>95</sub>) was calculated. Ceiling and floor effects were determined when the MDC<sub>95</sub> for average and function scores crossed 0 and 100, respectively.

**RESULTS:**  $n = 981$  responses were recorded, including  $n = 675$  complete surveys. Test-retest reliability was generally fair to excellent, depending on function and scale. MDC<sub>95</sub> values generally indicated scale responsiveness. Ceiling and floor effects were noted infrequently for specific functions.

**CONCLUSION:** The PAQ is valid, reliable, and sensitive. Additional research may explore measurement properties involving functions that were infrequently selected in this sample.

Keywords: Myalgic encephalomyelitis, chronic fatigue syndrome, post-exertional malaise, health related quality of life, measurement

## 1. Introduction

Myalgic encephalomyelitis (ME) is a cause of disabling fatigue that affects at least 1-4 million adults in the United States [1]. Many people living with ME

have yet to be diagnosed, so the actual prevalence of ME may be much greater than suggested by published estimates [2]. In addition, the novel coronavirus 2019 (COVID-19) pandemic has increased the prevalence of people living with disabling fatigue following illness (i.e., post-acute COVID-19, post-acute sequelae of novel coronavirus 2019, or Long Covid). Data from a recent systematic review of 40 studies on the fre-

---

\*Address for correspondence: Todd E. Davenport, University of the Pacific, Stockton, CA, USA. E-mail: tdavenport@pacific.edu.

quency of Long Covid indicated a pooled prevalence of 0.36 (95% confidence interval; CI: 0.25-0.48), 0.24 (95% CI: 0.13-0.39), 0.29 (95% CI: 0.12-0.57) and 0.51 (95% CI: 0.42-0.59) at 30, 60, and 90 days following a positive index test, respectively [3]. Among individuals with a positive COVID-19 test, approximately 43% of non-hospitalized cases and over half of hospitalized cases report symptoms and signs of Long Covid [3]. The most common symptoms reported across studies were fatigue and dyspnea. These studies do not consider the millions of people who do not have access to tests, or who have tested positive and convalesced outside the medical system. Yet, taken together, data from the epidemiology of ME and Long Covid are compelling that the frequency of prolonged disability and delayed recovery often associated with an initial illness is accelerating over time. This situation raises the importance of valid, reliable, and responsive tools to measure disability for clinical practice and research in ME, Long Covid, and apparently related conditions.

The hallmark clinical feature of ME is post-exertional neuroimmune exhaustion (PENE), which is considered compulsory for the diagnosis according to the International Consensus Criteria [4]. PENE involves an insufficient ability to produce energy on demand at the cellular level to complete desired activities [4]. Signs and symptoms of PENE include marked, rapid physical and/or cognitive fatigability in response to exertion; post-exertional symptom exacerbation; post-exertional exhaustion; prolonged recovery periods following activity; and a low threshold of physical and mental fatigability (lack of stamina) results in a substantial reduction in pre-illness activity level [4]. In addition to PENE, at least one symptom from among neurological impairments, immune dysfunction, and energy production or transportation impairment is required for the diagnosis of ME. Symptoms and signs of ME tend to wax and wane with activity, frequently lasting 3-7 days but may last a week or more [5, 6]. Symptoms consistent with PESE also have been observed in people with Long Covid [7, 8], suggesting co-occurrence of PENE-like symptoms across these etiologies of post-viral fatigue [9, 10]. The rapidly changing and diverse signs and symptoms of PENE often result in severe functional impairment in various aspects of social functioning [11-13]. Fluctuation and severity of PENE that pose unique challenges for disability measurement in individuals living with these signs and symptoms.

Measures of self-reported ability and disability may provide important insights regarding functional impacts of PENE. These measures are useful to researchers as study endpoints and for clinicians to monitor the severity of disease and effect of treatment. Despite the importance of measuring functional disability related to PENE, few self-report questionnaires exist for this purpose. Most instruments have discriminative validity to identify post exertional malaise [14], rather than longitudinally assess the functional impact of symptoms. The DePaul Symptom Questionnaire involves a 0-4 rating scale each for the frequency of each symptom and the extent to which each symptom bothers the respondent during the preceding six months [14]. Yet, this symptom inventory does not correlate specific symptoms, frequencies, and severities with deficits in functioning. Previously, we conducted a study to determine the measurement characteristics of the Fatigue Severity Scale, Multidimensional Fatigue Inventory, and Medical Outcome Survey Short Form 36 (Version 2) in people living with ME during and after a two-day cardiopulmonary exercise test [15]. While these instruments demonstrated convergent validity with participants' symptom reports, they also exhibited potential floor effects that call into question the responsiveness of these common surveys. These surveys also do not provide the opportunity for respondents to rate their ability or disability on specific functional tasks. Other measures are being used to measure the functional effects of PENE, such as various questionnaires within the Patient Reported Outcome Measure Information System (PROMIS). However, in addition to measuring function in the context of specific parts of the clinical presentation of PENE, these questionnaires have not yet been validated in people with PENE to date.

Development of a valid, reliable, and responsive functional self-rating instrument remains an important gap to address in clinical practice and research related to PENE. An adequate disability self-rating questionnaire would reflect specific functions affected by post-exertional symptoms and signs, as well as anchoring effectively to avoid floor and ceiling effects. Existing questionnaires that are designed with these purposes in mind may serve as an appropriate template for measuring PENE-related disability. However, this adaptation would need to be formally assessed in the context of a formal study. Therefore, the purpose of this study is to describe the development and preliminary validation of the Post

Exertional Malaise/Post Exertional Symptom Exacerbation (PEM/PESE) Activity Questionnaire.

## 2. Method

### 2.1. Survey development

The PEM/PESE Activity Questionnaire (PAQ) is based on the Patient Specific Functional Scale (PSFS) [16]. For the PSFS, the patient self-selects items with which they are experiencing disability after a change in status. The patient then rates their ability to complete each item on an 11-point scale anchored on each end at 0 (totally unable to perform) to 10 (totally able to perform) [16]. Test-retest reliability is excellent in people with persistent pain (intraclass correlation coefficient; ICC: .97) [16], knee pain (intraclass correlation coefficient; ICC: .84) [17], and neck pain (intraclass correlation coefficient; ICC: .92) [18]. The PSFS also demonstrates adequate convergent and criterion-related validity in people with neck pain [19, 20] and osteoarthritis [21–23].

The PSFS was modified in several substantive ways to create the PAQ:

- The PAQ was created as a computerized format, instead of the pencil-and-paper format of the PSFS. This modification was intended to increase the ease of data gathering and analysis for clinical and research purposes.
- The PAQ was created with slider bars for patients to self-rate ability instead of assigning a 0 to 10 number, in addition to the ability to input a number. This modification was intended to simplify integration of the survey into mobile apps and other applications that require a simplified user interface.
- The 0 to 10 scale of the PSFS was expanded to 0 to 100 for the PAQ to provide for more increments of measurement. This modification was intended to increase precision to optimize the opportunity for change sensitivity.
- Activities in the PAQ were tied to the International Classification of Functioning, Disability, and Health core set for myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) [24]. This modification was intended to enhance content validity and to reduce barriers to participation by people with cognitive dysfunction by providing a parsimonious set of activities for respondents with cognitive dysfunction. Activi-

ties were arranged in a drop-down menu rather than write-in format to further mitigate potential cognitive barriers to participation.

- Measurement anchors were explicitly used for the PAQ because the symptoms and signs of PENE, and therefore its consequent disability, may change frequently. This rapid change may challenge overall interpretation of changes in score. Patients were asked to rate their disability according to two different measurement anchorages, including “How well can you complete this activity in this moment compared to a good day?” and “How well can you complete this activity in this moment compared to before you became ill?”
- A third measurement scale was introduced for the PAQ to measure the time, effort, and resources necessary for the patient to complete the activity at the level of self-reported disability. This modification from the PSFS was intended to capture the phenomenon that patients living with PENE may require them to invest a substantial amount of time, effort, and resources to achieve even a low self-reported level of functioning.

A draft of the PAQ was circulated to a small group of ME/CFS researchers and patients for initial feedback. Minor modifications were made to the item and questionnaire formatting on the platform to enhance readability and comprehension. The PAQ was then introduced for formalized study. University of the Pacific’s Institutional Review Board approved the study protocol (#2021-181).

### 2.2. Participants

Participants were adults with self-reported symptoms of PENE who responded to an internet-based survey between November 7-December 12, 2021. Participant recruitment took place by social media (chiefly, Twitter, Facebook, and Instagram), and the link to the survey was circulated by the first author and in partnership with several patient advocacy societies.

### 2.3. Procedure

Participants were first required to signal informed consent by clicking their agreement with the informed consent document. Participants then answered two questions to screen for eligibility. The first question was age at the time of survey completion. The second was to click all the symptoms

of PENE that applied to the participant. Participants who do not meet eligibility requirements were taken to an exit screen and their data was not collected. Participants who met eligibility requirements were permitted to continue into the study survey.

The study survey continued with a first administration of the PAQ. Participants selected three activities that were meaningful to them using drop-down menus. They then rated their ability to do these three activities on three different scales. Each scale was anchored at 0 and 100, with 0 being “completely unable to perform at all” and 100 being “can perform at the same level as a good day” or “can perform at the same level before I became ill,” respectively. Participants rated their ability to do the task compared to (1) before they became ill and (2) what they would consider a good day. For these scales, lower scores mean more disability. For the third scale, participants rated themselves according to the extent of time, effort, and resources it took to perform the activity at that level. This scale was anchored at 0 being “no time, effort, and resources at all” and 100 being “all of my time, effort, and resources.” In this scale, greater scores mean more time, effort, and resources devoted to attain the level of self-rated disability.

Participants then completed a brief demographic questionnaire regarding the respondent’s age, gender, country of response, health history, and their symptom experience with PEM/PESE. This survey was intended to collect data regarding important covariates that could be used for participant and subgrouping in the analysis, as well as to serve as a distractor task. After the demographic questionnaire, participants received a second administration of the PAQ. Activities that participants selected on the first administration were fed forward to the second administration, so the reliability analysis focused on disability ratings instead of task selection.

#### 2.4. Data analysis

Descriptive statistics were calculated, including means, standard deviations, and 95% confidence intervals (95% CIs) for continuous data, and counts and proportions for binomial data. Two-sample *t*-tests assuming equal variances were used to assess between groups differences in continuous data, and Pearson’s chi-square analysis was used to determine the statistical significance of between-group differences in frequencies. Test-retest reliability was calculated using intra-class correlation coefficient

(ICC; two-way random effects model). Minimum detectable change outside the 95% confidence interval was calculated as  $1.96 \times \text{standard error of measure (SEM)} \times \sqrt{2}$ , where SEM was the standard deviation from the first test multiplied by  $\sqrt{(1 - \text{ICC})}$  [25, 26]. Ceiling and floor effects were assessed using the 95% CI calculated around the test 1 and test 2 mean scores. Where the 95% confidence interval around mean scores did not include 0 and 100, ceiling and floor effects were considered unlikely.

### 3. Results

#### 3.1. Survey completion and respondent characteristics

There were 981 responses to the survey. Eight surveys were excluded because the respondent reported no symptoms of PENE. No surveys were excluded due to respondent age under 18 years. Of the 973 remaining responses, 298 were considered incomplete because they did not contain all data from the first PAQ, demographic questionnaire, and second PAQ. Therefore, 675 surveys were available for analysis for an effective survey completion rate of 68.8% (Table 1). Mean age of respondents completing the survey was 47.8 years (95% CI: 46.8–48.7). 568 (84.1%) of respondents completing the survey identified as a woman (including trans woman) and most frequent country of response was the United Kingdom, including Northern Ireland, Scotland, and Wales ( $n=244$ ; 36.1%). The most frequent durations of illness followed a trimodal distribution with peaks at 1–2 years ( $n=149$ ; 22.1%), 6–10 years ( $n=116$ ; 17.2%), and 21 years or more ( $n=155$ ; 23.0%). The most common self-reported onsets were associated with herpesvirus infection ( $n=290$ ; 43.0%) and SARS-COV-2 infection ( $n=178$ ; 26.4%).

A comparison of respondent characteristics for completed vs. incomplete surveys indicated respondent age was statistically similar (Table 1). However, respondents completing the survey more frequently reported post-exertional symptom exacerbation, prolonged recovery periods after activity, and a low threshold of physical and mental fatigability resulting in a substantial reduction in pre-illness activity level ( $p < .001$ ). Respondents who completed and did not complete the survey reported statistically similar frequency of marked rapid physical and/or cognitive fatigability in response to exertion and post-exertional exhaustion.

Table 1  
Respondent characteristics

Demographic/illness characteristics	Completed survey (n = 675)	Incomplete survey (n = 298)	p-value
Age (years)	47.8 (46.8-48.7)	47.8 (46.5-49.2)	.909
Gender			
Woman (including trans woman)	568 (84.1%)	–	–
Man (including trans man)	73 (10.8%)	–	–
Non-binary	13 (1.9%)	–	–
Genderfluid	6 (0.9%)	–	–
Agender	5 (0.7%)	–	–
Other	2 (0.3%)	–	–
Prefer not to say	5 (0.7%)	–	–
Country			
Australia	26 (3.9%)	–	–
Canada	64 (9.5%)	–	–
Germany	15 (2.2%)	–	–
Norway	23 (3.4%)	–	–
Sweden	22 (3.3%)	–	–
The Netherlands	19 (2.8%)	–	–
United Kingdom (including Northern Ireland, Scotland, and Wales)	244 (36.1%)	–	–
United States of America	199 (29.5%)	–	–
All others (18 additional countries)	63 (9.3%)	–	–
Duration of illness			
5 months or less	10 (1.5%)	–	–
6-11 months	48 (7.1%)	–	–
1-2 years	149 (22.1%)	–	–
3-5 years	69 (10.2%)	–	–
6-10 years	116 (17.2%)	–	–
11-15 years	79 (11.7%)	–	–
16-20 years	46 (6.8%)	–	–
21 years or more	155 (23.0%)	–	–
Onset			
<i>Borrelia burgdoferi</i> infection	19 (2.8%)	–	–
Herpesvirus infection	290 (43.0%)	–	–
SARS-COV-2 infection	178 (26.4%)	–	–
Toxic mold	9 (1.3%)	–	–
Veteran of 1990-1991 Gulf War theatre of operations	5 (0.7%)	–	–
Other or not sure	169 (25.2%)	–	–
Self-reported symptoms and signs of post-exertional neuroimmune exhaustion			
Marked, rapid physical and/or cognitive fatigability in response to exertion, which may be minimal such as activities of daily living or simple mental tasks, can be debilitating and cause a relapse	595 (88.1%)	267 (87.3%)	.675
Post-exertional symptom exacerbation (e.g., acute flu-like symptoms, pain, and worsening of other symptoms.)	621 (92.0%)	256 (83.7%)	<.001
Post-exertional exhaustion that may occur immediately after activity or be delayed by hours or days	644 (95.4%)	290 (94.8%)	.747
Recovery period after activity is prolonged, usually taking 24 hours or longer	618 (91.6%)	250 (81.7%)	<.001
Low threshold of physical and mental fatigability (lack of stamina) results in a substantial reduction in pre-illness activity level	644 (95.4%)	267 (87.3%)	<.001
Other self-reported symptoms and signs			
Hard to process information mentally	607 (89.9%)	–	–
Short-term memory loss	612 (90.7%)	–	–
Headaches	503 (74.5%)	–	–
Significant pain	533 (79.0%)	–	–
Disturbed sleep patterns	590 (87.4%)	–	–
Unrefreshing sleep	619 (91.7%)	–	–
Sensory and perceptual impairments	563 (83.4%)	–	–
Motor impairments	506 (75.0%)	–	–
Flu-like symptoms	536 (79.4%)	–	–
Frequent illness or prolonged recovery time	272 (40.3%)	–	–

(Continued next page)

Table 1  
(Continued)

Demographic/illness characteristics	Completed survey (n = 675)	Incomplete survey (n = 298)	p-value
Symptom behavior			
Frequency of change for the symptom most frequently experienced			
By the minute	74 (11.0%)	–	–
By the hour	308 (45.6%)	–	–
By the day	150 (22.2%)	–	–
By the week	38 (5.6%)	–	–
By the month	5 (0.7%)	–	–
By the year	3 (0.4%)	–	–
It depends	60 (8.9%)	–	–
I'm not sure	34 (5.0%)	–	–
Frequency of change for the symptom least frequently experienced			
By the minute	4 (0.6%)	–	–
By the hour	14 (2.1%)	–	–
By the day	77 (11.4%)	–	–
By the week	115 (17.0%)	–	–
By the month	154 (22.8%)	–	–
By the year	87 (12.9%)	–	–
It depends	125 (18.5%)	–	–
I'm not sure	97 (14.4%)	–	–

### 3.2. Reported symptoms and signs

Self-reported symptoms and signs of PENE were frequently reported by the sample (Table 1), including marked, rapid physical and/or cognitive fatigability in response to exertion ( $n = 595$ ; 88.1%), PESE ( $n = 621$ ; 92.0%), post exertional exhaustion ( $n = 644$ ; 95.4%), prolonged recovery period ( $n = 618$ ; 91.6%), and lack of stamina ( $n = 644$ ; 95.4%). In addition, most self-reported symptoms and signs consistent with the International Classification Criteria for Myalgic Encephalomyelitis were reported by three-quarters or more of the sample. These symptoms included difficulty processing information mentally ( $n = 607$ ; 89.9%), short-term memory loss ( $n = 612$ ; 90.7%); headaches ( $n = 503$ ; 74.5%); significant pain ( $n = 533$ ; 79.0%); disturbed sleep patterns ( $n = 590$ ; 87.4%); unrefreshing sleep ( $n = 619$ ; 91.7%); sensory and perceptual impairments ( $n = 563$ ; 83.4%); motor impairments ( $n = 506$ ; 75.0%); and flu-like symptoms ( $n = 536$ ; 79.4%). The symptom most frequently experienced was most frequently reported to change by the hour or day ( $n = 382$ ; 56.6%). In contrast, the symptom least frequently experienced was most often reported to change by the week or month ( $n = 269$ ; 39.8%).

### 3.3. Selected functions

All functions were selected by at least four respondents (Table 2). The most common functions

selected were “carry out my daily routine” ( $n = 253$ ; 12.5%), “maintain employment for financial reward” ( $n = 166$ ; 8.3%), “wash myself” ( $n = 154$ ; 7.6%), “walk short distances” ( $n = 151$ ; 7.5%), and “do recreational and leisure activities” ( $n = 123$ ; 6.1%). On the first measurement, “compared to a good day” mean scores ranged from 24.7 (95% CI: 25.9-34.0) for “be economically self-sufficient” to 92.8 (95% CI: 81.6-104.1) for “lay down.” On the second measurement, “compared to a good day” mean scores ranged from 23.5 (95% CI: 16.9-30.2) for “care for my hair” to 71.5 (95% CI: 40.8-102.2) for “lay down.” These scores suggest low self-reported functioning even with common activities of daily living involving normal metabolic expenditure, compared to the respondent’s assessment of a good day.

Mean responses for the “compared to before illness” prompt generally were lower than “compared to a good day” prompt at each measurement point. First measurement “compared to before illness” mean scores ranged from 14.8 (95% CI: 9.8-19.9) for “sit up” and 14.8 (95% CI: -7.0-36.5) for “care for my skin” to 80.7 (95% CI: 53.4-107.9) for “lay down.” On the second measurement, mean scores ranged from 14.0 (95% CI: -6.3-34.3) for “care for my skin” to 84.3 (95% CI: 64.6-104.1) for “lay down.” These scores indicate a low degree of self-reported function for many common activities of daily living, compared to pre-illness level of functioning.

Burden of functioning mean scores ranged from 61.3 (95% CI: 32.5-90.1) for “lay down” to 85.5

Table 2  
Activities selected by respondents

Activity	Reported frequency	First measurement			Second measurement		
		Compared to a good day	Compared to before illness	Burden of functioning	Compared to a good day	Compared to before illness	Burden of functioning
Carry out my daily routine	253 (12.5%)	43.4 (40.7-46.1)	27.1 (24.4-29.9)	82.4 (80.3-84.6)	41.2 (38.4-44.1)	26.9 (24.4-29.3)	81.4 (79.3-83.5)
Lay down	6 (0.3%)	92.8 (81.6-104.1)	80.7 (53.4-107.9)	61.3 (32.5-90.1)	71.5 (40.8-102.2)	84.3 (64.6-104.1)	53.0 (11.4-94.6)
Sit up	31 (1.5%)	36.4 (25.5-47.2)	14.8 (9.8-19.9)	72.6 (64.6-80.7)	35.6 (24.4-46.9)	14.5 (8.9-20.2)	70.9 (62.3-79.5)
Stand up	45 (2.2%)	39.5 (32.1-46.9)	21.3 (15.0-27.6)	76.8 (71.0-82.6)	35.2 (28.3-42.0)	20.5 (14.0-27.0)	75.0 (68.5-81.4)
Manage stress and other psychological demands	105 (5.2%)	39.2 (34.9-43.4)	27.9 (23.7-32.1)	75.3 (71.5-79.2)	36.2 (32.2-40.2)	28.3 (24.4-32.2)	75.6 (71.8-79.4)
Walk short distances	151 (7.5%)	32.8 (29.1-36.5)	17.3 (14.4-20.3)	78.7 (75.8-81.6)	32.6 (28.8-36.3)	17.6 (14.8-20.5)	78.6 (75.8-81.4)
Walk long distances	112 (5.5%)	27.2 (22.2-32.1)	17.8 (13.8-21.8)	81.3 (76.8-85.7)	27.7 (22.7-32.7)	17.6 (13.6-21.5)	82.2 (77.8-86.5)
Care for my hair	39 (1.9%)	29.5 (21.8-37.2)	18.7 (13.3-24.1)	79.9 (72.8-87.1)	23.5 (16.9-30.2)	19.1 (12.7-25.4)	79.1 (71.7-86.5)
Wash myself	154 (7.6%)	32.7 (29.0-36.4)	20.1 (17.0-23.1)	76.3 (73.0-79.6)	29.0 (25.6-32.4)	19.5 (16.6-22.4)	78.4 (75.3-81.6)
Dress myself	19 (0.9%)	47.6 (34.1-61.2)	23.8 (11.4-36.3)	62.1 (50.8-73.3)	36.7 (22.5-51.0)	25.1 (14.5-35.6)	65.4 (54.8-76.0)
Take care of my own health	79 (3.9%)	41.3 (36.6-46.0)	29.1 (24.5-33.8)	77.8 (73.6-82.1)	37.7 (32.4-43.0)	28.4 (23.6-33.1)	79.2 (74.7-83.6)
Shop and gather daily necessities	81 (4.0%)	32.0 (26.8-37.2)	20.2 (15.5-24.9)	78.4 (73.5-83.3)	30.6 (25.5-35.7)	21.6 (17.4-25.7)	77.5 (72.3-82.8)
Prepare meals	197 (9.7%)	35.8 (32.4-39.1)	22.6 (19.7-25.6)	77.1 (74.4-79.8)	33.3 (30.0-36.6)	21.5 (18.7-24.3)	78.2 (75.5-80.9)
Do housework	116 (5.7%)	27.5 (28.1-48.9)	19.6 (15.8-23.4)	79.1 (75.2-82.9)	28.4 (24.2-32.6)	19.9 (16.3-23.4)	79.5 (75.7-83.2)
Assist others	23 (1.1%)	38.5 (35.6-43.6)	28.5 (18.4-38.6)	77.1 (67.6-86.7)	38.4 (27.7-49.1)	24.9 (15.9-33.8)	82.8 (75.9-89.8)
Engage in basic interpersonal interactions	114 (5.6%)	39.6 (28.6-37.2)	22.7 (19.2-26.2)	76.1 (72.4-79.8)	35.9 (31.9-39.8)	23.0 (19.4-26.6)	76.7 (73.1-80.4)
Maintain employment for financial reward	166 (8.3%)	32.9 (17.0-32.4)	23.1 (19.4-26.7)	84.0 (80.4-87.7)	31.0 (26.6-35.3)	22.1 (18.5-25.6)	82.8 (79.0-86.6)
Be economically self-sufficient	53 (2.6%)	24.7 (25.9-34.0)	17.2 (10.6-23.9)	81.2 (73.1-89.2)	24.9 (17.2-32.7)	18.9 (11.9-26.0)	81.6 (73.6-89.7)
Do recreational and leisure activities	123 (6.1%)	30.0 (33.7-43.0)	17.4 (14.7-20.2)	81.4 (77.7-85.1)	32.0 (27.7-36.2)	20.0 (16.9-23.0)	81.2 (77.7-84.8)
Engage in informal or casual gatherings with others	97 (4.8%)	38.3 (49.0-65.2)	22.4 (18.9-25.9)	77.5 (73.5-81.6)	39.6 (34.7-44.5)	25.4 (21.5-29.3)	78.7 (75.1-82.2)
Use the toilet	34 (1.7%)	57.1 (29.0-54.3)	37.8 (26.8-48.8)	74.2 (67.0-81.4)	52.9 (43.5-62.3)	38.1 (26.9-49.3)	71.6 (62.9-80.3)
Care for my teeth	21 (1.0%)	41.7 (-0.5-83.0)	30.9 (18.4-43.4)	66.0 (55.4-76.6)	43.0 (30.8-55.2)	30.3 (16.4-44.2)	67.9 (56.9-78.9)
Care for my skin	4 (0.2%)	41.3 (28.1-48.9)	14.8 (-7.0-36.5)	85.5 (65.6-105.4)	42.3 (-9.64-94.1)	14.0 (-6.3-34.3)	88.8 (80.2-97.3)

(65.6-105.4) for “care for my skin” on the first measurement. On the second measurement, burden of functioning scores ranged from 53.0 (95% CI: 11.4-94.6) for “lay down” to 88.8 (95% CI: 80.2-97.3) for “care for my skin.” These scores suggest a great extent of time, resources, and effort to attain a relatively low level of self-reported functioning with common daily tasks.

### 3.4. Test-retest reliability

Test-retest reliability was generally good to excellent for all functions compared to a good day, compared to before illness, and for the burden of functioning score (Table 3). Overall, ICCs for all items compared to a good day, compared to before illness, and the burden of functioning score were .871 (95% CI: .860-.881), .807 (.791-.821), and .859 (.847-.870), respectively. ICCs for function self-ratings compared to a good day ranged from .229 (95% CI: -.635-.839) to .963 (.546-.998) for “care for my skin.” ICCs for function self-ratings compared to before illness were between .565 (95% CI: .426-.678) for “engage in basic interpersonal interactions” and .997 (.960-1.00) for “care for my skin.” ICCs for the burden of functioning score were between .425 (95% CI: .089-.674) for “sit up” and .977 (95% CI: .961-.987) for “be economically self-sufficient.

### 3.5. Responsiveness

MDC<sub>95</sub> for all items was 36 for “compared to a good day,” 41 for “compared to before illness,” and 23 for the burden of functioning score (Table 3). For the “compared to a good day” prompt, MDC<sub>95</sub> was between 15 for “be economically self-sufficient” and 58 for “dress myself.” MDC<sub>95</sub> ranged from 6 for “care for my skin” to 66 for “engage in basic interpersonal interactions” for the “compared to before illness” prompt. Burden of functioning score MDC<sub>95</sub> ranged from 7 for “be economically self-sufficient” and 114 for “lay down.” Few respondents selected “lay down,” (6; 0.3%) and “care for my skin” (4; 0.2%).

### 3.6. Ceiling and floor effects

Potential floor effects were observed for “care for my teeth” and “care for my skin,” and potential ceiling effects were identified for “lay down.” The 95% CI crossed zero for “care for my teeth” compared to a good day on the first measurement. In addition,

the 95% CI included zero for “care for my skin” compared to before illness on the first measurement, as well as “care for my skin” compared to before illness and compared to a good day on the second measurement. The 95% CI included 100 for the first and second measurements of “lay down” compared to a good day and compared to before illness. Perhaps notably, these items were selected by the least respondents; 6 (0.3%) for “lay down,” 21 (1.0%) for “care for my teeth”, and 4 (0.2%) for “care for my skin.”

## 4. Discussion

This study involved the preliminary development and investigation of the PEM/PESE Activity Questionnaire. The PAQ was created because existing instruments were not specifically created to assess functional deficits associated with PENE. Findings from this study indicated the PAQ is feasible to administer electronically, which is relevant for eventual integration into various telehealth applications. An existing WHO ICF Core Set for myalgic encephalomyelitis was used as the conceptual foundation for the functions in the PAQ to maximize its internal validity. The PAQ overall and most of its individual options for function self-rating were identified to have good to excellent test-retest reliability, adequate responsiveness, and negligible floor and ceiling effects. Two anchors were used for self-rating, including “compared to a good day” and “compared to before illness.” These anchors each may be useful in clinical and research contexts because they provide two separate temporal frames of reference that are meaningful to patients.

Most questionnaires in existence for people living with PENE are focused on quantifying symptoms for the purpose of identifying PENE or PEM. While questionnaires such as the DePaul Symptom Questionnaire demonstrate high reliability, internal consistency, concurrent validity, and discriminative validity [14, 27–29], the ability to use these existing questionnaires to track the specific functional effects of signs and symptoms over time has remained questionable. General health related quality of life instruments such as the Medical Outcome Survey Short Form 36 have been used widely in clinical and research applications involving people living with PENE. However, it has questionable responsiveness, including floor and ceiling effects, in people living

Table 3  
Intraclass correlation coefficients and minimum detectable change outside the 95% confidence interval

Activity	Compared to a good day			Compared to before illness			Burden of functioning		
	ICC	p-value	MDC <sub>95</sub>	ICC	p-value	MDC <sub>95</sub>	ICC	p-value	MDC <sub>95</sub>
Overall	.871 (.860-.881)	<.001	36	.807 (.791-.821)	<.001	41	.859 (.847-.870)	<.001	23
Carry out my daily routine	.851 (.812-.881)	<.001	46	.720 (.655-.775)	<.001	61	.851 (.812-.881)	<.001	29
Lay down	.229 (-.635-.839)	.310	—	.903 (.465-.986)	.003	62	.738 -.038-.959	.029	114
Sit up	.906 (.814-.954)	<.001	31	.814 (.649-.906)	<.001	43	.425 (.089-.674)	.008	31
Stand up	.916 (.853-.953)	<.001	32	.927 (.870-.959)	<.001	26	.893 (.814-.940)	<.001	19
Manage stress and other psychological demands	.750 (.652-.823)	<.001	54	.704 (.593-.789)	<.001	55	.843 (.777-.891)	<.001	31
Walk short distances	.875 (.832-.908)	<.001	32	.723 (.637-.791)	<.001	48	.882 (.840-.913)	<.001	17
Walk long distances	.903 (.862-.932)	<.001	23	.779 (.694-.843)	<.001	36	.939 (.913-.958)	<.001	12
Care for my hair	.706 (.505-.834)	<.001	44	.716 (.520-.840)	<.001	35	.971 (.946-.985)	<.001	9
Wash myself	.875 (.832-.908)	<.001	32	.733 (.650-.798)	<.001	41	.869 (.824-.903)	<.001	20
Dress myself	.807 (.565-.921)	<.001	58	.836 (.624-.934)	<.001	41	.871 (.697-.948)	<.001	24
Take care of my own health	.844 (.766-.897)	<.001	45	.858 (.786-.907)	<.001	39	.756 (.644-.837)	<.001	40
Shop and gather daily necessities	.894 (.840-.931)	<.001	29	.896 (.842-.932)	<.001	27	.887 (.830-.926)	<.001	19
Prepare meals	.840 (.794-.877)	<.001	40	.854 (.811-.888)	<.001	35	.847 (.803-.882)	<.001	25
Do housework	.920 (.887-.944)	<.001	22	.818 (.747-.871)	<.001	34	.769 (.683-.834)	<.001	26
Assist others	.924 (.830-.967)	<.001	29	.905 (.785-.959)	<.001	33	.818 (.618-.918)	<.001	34
Engage in basic interpersonal interactions	.802 (.725-.859)	<.001	49	.565 (.426-.678)	<.001	66	.782 (.699-.844)	<.001	29
Maintain employment for financial reward	.925 (.900-.944)	<.001	25	.835 (.783-.876)	<.001	35	.938 (.917-.954)	<.001	16
Be economically self-sufficient	.952 (.918-.972)	<.001	15	.944 (.905-.967)	<.001	16	.977 (.961-.987)	<.001	7
Do recreational and leisure activities	.888 (.844-.920)	<.001	28	.802 (.729-.857)	<.001	39	.851 (.794-.893)	<.001	19
Engage in informal or casual gatherings with others	.821 (.744-.877)	<.001	45	.872 (.814-.913)	<.001	39	.713 (.599-.799)	<.001	33
Use the toilet	.926 (.857-.962)	<.001	43	.865 (.747-.930)	<.001	54	.893 (.796-.945)	<.001	34
Care for my teeth	.939 (.856-.975)	<.001	29	.865 (.698-.943)	<.001	44	.910 (.793-.963)	<.001	26
Care for my skin	.963 (.546-.998)	.004	22	.997 (.960-1.00)	<.001	6	.614 (-.574-.969)	<.136	25

with PENE who are known to have been exposed to heavy physical exertion [15]. Condition-specific health related quality of life instruments for fatigue, such as the Multi-dimensional Fatigue Inventory, also have known problems with responsiveness in people living with PENE who have been exposed to heavy physical exertion.

General and condition-specific health related quality of life instruments have not been validated to measure the time, effort, and resources necessary to attain the self-reported level of function. This is important for functional assessment in PENE, which often requires a significant investment of preparatory activities and self-care to do basic daily tasks. The Burden of Function self-rating contained in the PAQ allows for specific assessment of the ease with which the respondent can complete the task at their self-reported level of function. Considering the ease of functioning alongside the level of functioning may provide an important basis to determine whether functional improvement or worsening may be taking place; even while functional ability/disability may remain the same, the time, effort, and resources may less (suggesting improvement) or more (suggesting worsening) at any given time. Through the interpretation of the Burden of Function score, the PAQ may be more optimized to detect fine changes in functioning than the psychometric properties of the functional self-ratings may indicate.

This study has limitations. Chiefly, the internet-based methodology of this study's design used as sample of convenience that may not reflect the entire population living with PENE. This limitation was addressed by reporting a complete set of descriptive data so the reader may assess for themselves whether an important challenge to generalizability might exist between the data reported in this study and their specific application. All respondents included in this study reported PENE and its associated symptoms. Respondents who completed the survey were as disabled or more disabled by signs and symptoms of PENE than non-respondents, so any response bias that may be present in this study appears to be toward more signs and symptoms, and therefore potentially more functional impairment. The sample also may have over-represented women, which may serve to under-represent the functional impairments experienced by men in this sample. It is unclear whether and how severely participants may have experienced increased PENE resulting from taking this survey. This important question may be the subject of future research.

## 5. Conclusion

Based on this study's results, the PAQ has some promising applications and directions for additional research. The use of a WHO ICF Core Set to map functions provides a foundational basis for this survey in existing terminology and typology for human functioning. This feature will facilitate bioinformatics and software engineering efforts to integrate the PAQ into electronic health records and research databases. In turn, this integration will enhance our ability to phenotype specific functional decrements in people living with conditions involving PENE. Some of the functions were selected infrequently by respondents, so the statistical assessment of their psychometric properties likely are underpowered. Under-selected functions in this study may be the subject of further studies. Future studies also may stratify subjects by illness severity to determine whether important differences in functions selected and psychometric properties exist.

## Ethical approval

The study was approved by the University of the Pacific, Stockton, CA, USA (#2021-181). Informed consent. This study received exempt review approval. Subjects acknowledged informed consent to participate in the study by clicking into the survey.

## Conflict of interest

The authors declare no conflicts of interest.

## Funding

The authors report no funding.

## Acknowledgments

The authors thank the study participants for helping to develop this survey and to advance our collective understanding of disability in people living with post exertional malaise/post-exertional symptom exacerbation.

## References

- [1] United States Centers for Disease Control and Prevention. Chronic Fatigue Syndrome Atlanta, GA: CDC; 2008 [Available from: <http://www.cdc.gov/cfs/>].
- [2] Reeves WC, Jones JF, Maloney E, Heim C, Hoaglin DC, Boneva RS, et al. Prevalence of chronic fatigue syndrome in metropolitan, urban, and rural Georgia. *Popul Health Metr*. 2007;5:5.
- [3] Chen C, Hauptert SR, Zimmermann L, Shi X, Fritsche LG, Mukherjee B. Global Prevalence of Post COVID-19 Condition or Long COVID: A Meta-Analysis and Systematic Review. *J Infect Dis*. 2022.
- [4] Carruthers BM, van de Sande MI, De Meirleir KL, Klimas NG, Broderick G, Mitchell T, et al. Myalgic encephalomyelitis: International Consensus Criteria. *J Intern Med*. 2011;270(4):327-38.
- [5] Mateo LJ, Chu L, Stevens S, Stevens J, Snell CR, Davenport T, et al. Post-exertional symptoms distinguish Myalgic Encephalomyelitis/Chronic Fatigue Syndrome subjects from healthy controls. *Work*. 2020;66(2):265-75.
- [6] Chu L, Valencia JJ, Garvert DW, Montoya JG. Deconstructing post-exertional malaise in myalgic encephalomyelitis/chronic fatigue syndrome: A patient-centered, cross-sectional survey. *PLoS One*. 2018;13(6):e0197811.
- [7] Davis HE, Assaf GS, McCorkell L, Wei H, Low RJ, Re'em Y, et al. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. *EClinicalMedicine*. 2021;38:101019.
- [8] Twomey R, DeMars J, Franklin K, Culos-Reed SN, Weatherald J, Wrightson JG. Chronic Fatigue and Postexertional Malaise in People Living With Long COVID: An Observational Study. *Phys Ther*. 2022;102(4).
- [9] Jason LA, Islam M, Conroy K, Cotler J, Torres C, Johnson M, et al. COVID-19 Symptoms Over Time: Comparing Long-Haulers to ME/CFS. *Fatigue*. 2021;9(2):59-68.
- [10] Kujawski S, Slomko J, Newton JL, Eaton-Fitch N, Staines DR, Marshall-Gradisnik S, et al. Network Analysis of Symptoms Co-Occurrence in Chronic Fatigue Syndrome. *Int J Environ Res Public Health*. 2021;18(20).
- [11] Drachler Mde L, Leite JC, Hooper L, Hong CS, Pheby D, Nacul L, et al. The expressed needs of people with chronic fatigue syndrome/myalgic encephalomyelitis: a systematic review. *BMC Public Health*. 2009;9:458.
- [12] Kingdon CC, Bowman EW, Curran H, Nacul L, Lacerda EM. Functional status and well-being in people with myalgic encephalomyelitis/chronic fatigue syndrome compared with people with multiple sclerosis and healthy controls. *Pharmacoecon Open*. 2018;2(4):381-92.
- [13] Ziauddeen N, Gurdasani D, O'Hara ME, Hastie C, Roderick P, Yao G, et al. Characteristics and impact of Long Covid: Findings from an online survey. *PLoS One*. 2022;17(3):e0264331.
- [14] Cotler J, Holtzman C, Dudun C, Jason LA. A Brief Questionnaire to Assess Post-Exertional Malaise. *Diagnostics (Basel)*. 2018;8(3).
- [15] Davenport TE, Stevens SR, Baroni K, Van Ness JM, Snell CR. Reliability and validity of Short Form 36 Version 2 to measure health perceptions in a sub-group of individuals with fatigue. *Disabil Rehabil*. 2011;33(25-26):2596-604.
- [16] Stratford P, Gill C, Westaway M, Binkley J. Assessing disability and change on individual patients: a report of a patient specific measure. *Physiotherapy Canada*. 1995;47:258-63.
- [17] Chatman AB, Hyams SP, Neel JM, Binkley JM, Stratford PW, Schomberg A, et al. The Patient-Specific Functional Scale: measurement properties in patients with knee dysfunction. *Phys Ther*. 1997;77(8):820-9.
- [18] Westaway MD, Stratford PW, Binkley JM. The patient-specific functional scale: validation of its use in persons with neck dysfunction. *J Orthop Sports Phys Ther*. 1998;27(5):331-8.
- [19] Cleland JA, Fritz JM, Whitman JM, Palmer JA. The reliability and construct validity of the Neck Disability Index and patient specific functional scale in patients with cervical radiculopathy. *Spine (Phila Pa 1976)*. 2006;31(5):598-602.
- [20] Young IA, Cleland JA, Michener LA, Brown C. Reliability, construct validity, and responsiveness of the neck disability index, patient-specific functional scale, and numeric pain rating scale in patients with cervical radiculopathy. *Am J Phys Med Rehabil*. 2010;89(10):831-9.
- [21] Berghmans DD, Lenssen AF, van Rhijn LW, de Bie RA. The Patient-Specific Functional Scale: Its Reliability and Responsiveness in Patients Undergoing a Total Knee Arthroplasty. *J Orthop Sports Phys Ther*. 2015;45(7):550-6.
- [22] Gill SD, de Morton NA, Mc Burney H. An investigation of the validity of six measures of physical function in people awaiting joint replacement surgery of the hip or knee. *Clin Rehabil*. 2012;26(10):945-51.
- [23] Wright HH, O'Brien V, Valdes K, Koczan B, MacDermid J, Moore E, et al. Relationship of the Patient-Specific Functional Scale to commonly used clinical measures in hand osteoarthritis. *J Hand Ther*. 2017;30(4):538-45.
- [24] Bileviciute-Ljungar I, Schult ML, Borg K, Ekholm J. Preliminary ICF core set for patients with myalgic encephalomyelitis/chronic fatigue syndrome in rehabilitation medicine. *J Rehabil Med*. 2020;52(6):jrm00074.
- [25] DiFabio RP. *Essentials of Rehabilitation Research: A Statistical Guide to Clinical Practice (1st ed.)*. New York City, NY, USA: F.A. Davis; 2012.
- [26] Watkins LG. *Foundations of Clinical Research: Applications to Evidence-Based Practice Fourth Edition (4th ed.)*. New York City, N.Y.: F.A. Davis; 2022.
- [27] Brown AA, Jason LA. Validating a measure of myalgic encephalomyelitis/chronic fatigue syndrome symptomatology. *Fatigue*. 2014;2(3):132-52.
- [28] Jason LA, So S, Brown AA, Sunnquist M, Evans M. Test-Retest Reliability of the DePaul Symptom Questionnaire. *Fatigue*. 2015;3(1):16-32.
- [29] Murdock KW, Wang XS, Shi Q, Cleland CS, Fagundes CP, Vernon SD. The utility of patient-reported outcome measures among patients with myalgic encephalomyelitis/chronic fatigue syndrome. *Qual Life Res*. 2017;26(4):913-21.