



## Review article

## Lack of content overlap and essential dimensions – A review of measures used for post-stroke fatigue



Ingrid Johansen Skogestad<sup>a,b,\*</sup>, Marit Kirkevold<sup>b</sup>, Bent Indredavik<sup>c,d</sup>, Caryl L. Gay<sup>e,f</sup>, Anners Lerdal<sup>b,e</sup>, On behalf of the NORFAST (Norwegian Study of Fatigue After Stroke) Group

<sup>a</sup> Lovisenberg Diaconal Hospital, Oslo, Norway

<sup>b</sup> Department of Nursing Science, Institute of Health and Society, Faculty of Medicine, University of Oslo, Norway

<sup>c</sup> Department of Neuromedicine and Movement Science, NTNU, Trondheim, Norway

<sup>d</sup> Department of Stroke, St. Olavs Hospital, Trondheim, Norway

<sup>e</sup> Department of Research and Development, Lovisenberg Diaconal Hospital, Oslo, Norway

<sup>f</sup> Department of Family Health Care Nursing, University of California, San Francisco, USA

## ARTICLE INFO

## Keywords:

Content validity  
Fatigue  
Mixed methods  
PROM  
Rehabilitation  
Stroke

## ABSTRACT

**Introduction:** Post-stroke fatigue (PSF) is a highly prevalent and disabling condition. A major obstacle in PSF research is the lack of consensus on how to assess and diagnose fatigue after stroke. A wide variety of patient reported outcome measures (PROMs) are currently being used, none of which are developed specifically for stroke patients. The objectives of this study are to evaluate content validity in individual fatigue PROMs, and to identify similarities and differences through cross-comparison of PROMs.

**Methods:** We used a novel mixed-methods approach to evaluate content validity in fatigue PROMs. First, we performed a qualitative content analysis of items in eleven fatigue PROMs used in stroke populations, and then we used descriptive statistics and a similarity coefficient to investigate similarities and differences across instruments.

**Results:** The analysis of 156 items in eleven PROMs revealed 83 different items each representing a distinct attribute of fatigue. The results show that currently used fatigue PROMs omit important PSF-specific items, do not take into account the multidimensional nature of PSF and lack content overlap.

**Summary:** The wide variety of items and lack of overlap between fatigue PROMs illuminates the need for researchers to report why a specific PROM was used. PROMs that capture the specific experiences of patients with PSF are also needed to advance research on PSF and its etiology and treatment.

## 1. Introduction

Post-stroke fatigue (PSF) affects 50% of stroke survivors. This disabling condition has negative impacts on patients' rehabilitation, functioning and return to work [1–4]. PSF is often described as a feeling of physical or mental exhaustion, which may develop in connection with routine activities, following an acute stroke [5,6]. Despite recent efforts, no effective PSF treatment exists [7]. Lack of consensus on how to assess and diagnose fatigue after stroke is a major obstacle in PSF research [8]. Currently, PSF researchers and clinicians use a wide variety of patient reported outcome measures (PROMs), none of which are developed specifically for stroke [9,10].

Despite evidence that PSF is a multidimensional phenomenon, this is generally not reflected in the PROMs used in research [8]. The

selection of PROMs is often based on psychometric testing such as evaluation of internal structure, reliability and responsiveness [11]. Based on this limited evidence, researchers conclude that the instruments are valid and reliable. However, testing of validity and reliability involves evaluation of a multitude of measurement properties. The first, and the one with most clinical impact, is content validity [12–14]. Content validity is an assessment of whether the content of PROMs reflects the construct that is to be measured, and refers to the relevance, comprehensiveness and comprehensibility of PROMs in relation to the construct, target population and context of use [13]. To achieve accurate interpretations of research results related to PSF, it is essential that PROMs include stroke-relevant items and cover relevant PSF dimensions [11]. There are three major concerns regarding item content when assessing PSF with a PROM not developed specifically for stroke

\* Corresponding author at: Lovisenberggata 17, N-0456 Oslo, Norway.

E-mail addresses: [ingridjohansen.skogestad@lds.no](mailto:ingridjohansen.skogestad@lds.no) (I.J. Skogestad), [marit.kirkevold@medisin.uio.no](mailto:marit.kirkevold@medisin.uio.no) (M. Kirkevold), [bent.indredavik@ntnu.no](mailto:bent.indredavik@ntnu.no) (B. Indredavik), [caryl.gay@ucsf.edu](mailto:caryl.gay@ucsf.edu) (C.L. Gay), [anners.lerdal@medisin.uio.no](mailto:anners.lerdal@medisin.uio.no) (A. Lerdal).

<https://doi.org/10.1016/j.jpsychores.2019.109759>

Received 29 March 2019; Received in revised form 19 June 2019; Accepted 30 June 2019

0022-3999/© 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

survivors. First, the PROM is likely to include items that are irrelevant for a stroke population. Second, the PROM might fail to identify issues that are a specific feature of PSF [15]. Third, items often refer to the impact of fatigue on physical functioning and participation in everyday activities, not considering that these dimensions might be directly limited by other stroke sequela, thus conflating the fatigue scores [10]. Item content can affect all other measurement properties. Irrelevant items may reduce internal consistency, structural validity, and interpretability of PROMs. Missing dimensions may reduce validity and responsiveness [15]. In addition, the use of different PROMs across studies might affect the generalizability and replicability of PSF research [16].

Based on the lack of a stroke-specific instrument to measure PSF, the overall aims of this study are to explore the item relevance, missing items and comprehensiveness in dimensions of PROMs used in PSF research, and to identify similarities and differences through cross-comparison of PROMs.

## 2. Methods

We performed a systematic literature search to identify all studies using PROMs to assess PSF. We then applied a mixed-methods approach in order to analyze the selected PROMs individually and to do a cross-comparison of items, sub-dimensions and dimensions [17].

### 2.1. Data collection

#### 2.1.1. Search strategy

The systematic literature search was based on PRISMA-guidelines and performed in Medline, EMBASE, Cinahl and PsychINFO up to May 24, 2018. We developed the search strategy and conducted the search in collaboration with specialized research librarians. Search terms contained words equivalent to stroke, fatigue and rehabilitation (Supplemental Table I). There were no limitations regarding publication date, but we only included studies in English, Norwegian, Danish or Swedish.

#### 2.1.2. Eligibility criteria

We included studies of individuals with stroke (both first incidence or recurrent, ischemic or hemorrhagic) that had PSF as a primary outcome, and studies testing psychometric properties of fatigue PROMs in a stroke population. Since the search intent was to investigate the frequency of use of different fatigue PROMs, all study designs (except reviews) were included. We defined the following exclusion criteria: lack of primary data, duplicate studies, < 50% of the study sample affected by stroke, lack of recognized fatigue PROMs, conference abstracts, case reports, and studies not available in full text [12]. In addition, we considered relevant studies from reference lists of the included studies and major PSF reviews. The first author (IJS) scrutinized the titles, abstracts, and full-text articles and considered inclusion based on the above-mentioned criteria. This resulted in 78 included studies (Fig. 1) that used 24 different PROMs for measuring fatigue (Supplemental Table II).

#### 2.1.3. Selection of PROMs

The 24 different PROMs were used in anywhere from one to 39 studies on PSF. Eight PROMs were used in five or more studies and the rest were used in three studies or less. Given the aim to analyze the most commonly used PROMs in a stroke population, the research group (IJS, MK and AL) decided to set the inclusion cut-off at PROMs used in five studies or more. This resulted in inclusion of Fatigue Severity Scale [FSS] [18], Visual Analog Scale-Fatigue [VAS-F], Short Form-36 Vitality [SF-36 Vitality] [19], Checklist Individual Strength Fatigue Sub-scale [CIS] [20], Fatigue Assessment Scale [FAS] [21], Multidimensional Fatigue Inventory [MFI-20] [22], Mental Fatigue Scale [MFS] [23], and Modified Fatigue Impact Scale [M-FIS] [24]. In

addition, we chose to include Profile of Mood States-Fatigue [POMS] [25], Neurological Fatigue Index-Stroke [NFI-Stroke] [26], and Dutch Multifactorial Fatigue Scale [DMFS] [27], as these PROMs were partly developed or tested in a stroke population [27–29]. This resulted in eleven PROMs included in the final analysis.

### 2.2. Qualitative data analysis

The eleven PROMs and affiliated items were subject to a qualitative content analysis [30,31], consisting of two stages: data preparation and organizing of data [32].

#### 2.2.1. Preparation of data

The eleven PROMs contained a total of 156 items (i.e. questions), ranging from one to 38 items in each of the separate PROMs (Supplemental Table III). We analyzed the items, but not the scoring method (Supplemental Table III), as this was beyond the scope of this study. NVivo (Version 11.4.1.1064) was used to keep track of the analysis [33].

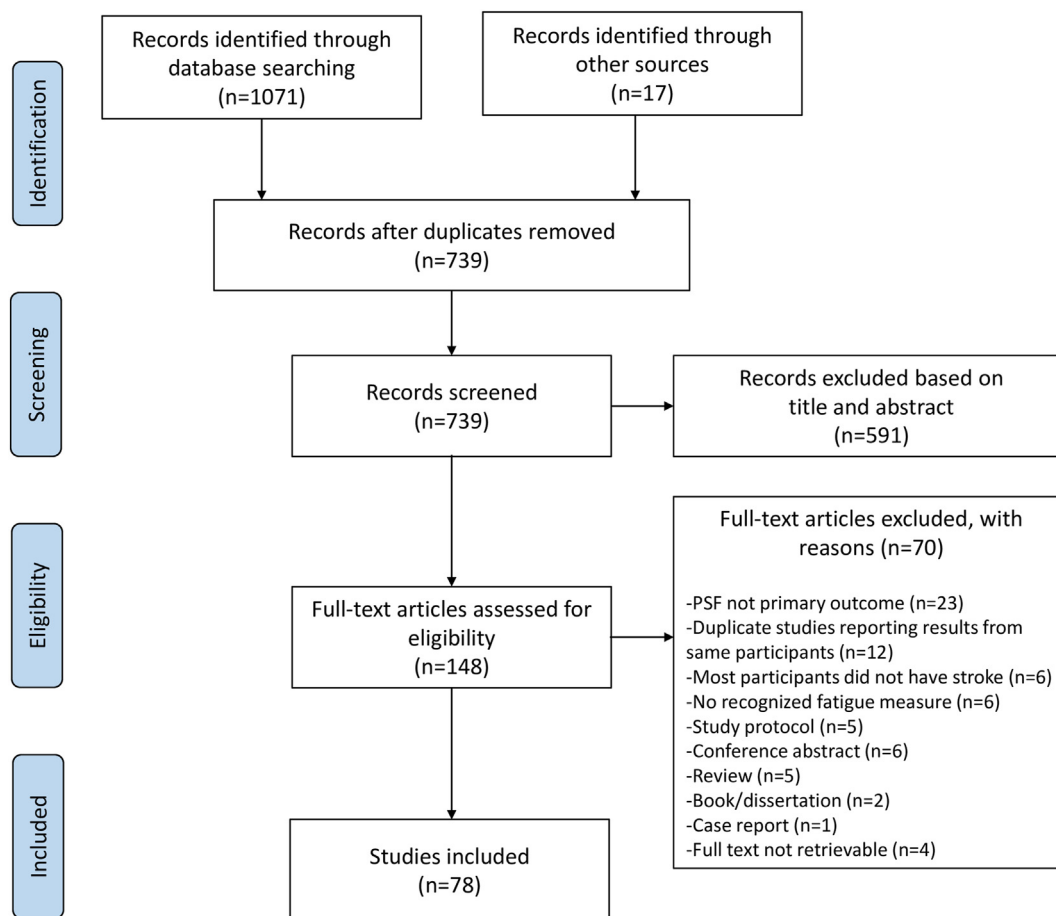
#### 2.2.2. Organizing of data

This second stage consisted of analyzing and comparing items, organizing items into sub-dimensions, and identifying overarching dimensions. The first step in this inductive analytical process included reading and re-reading all PROMs separately to get a sense of each instrument in its entirety. Then all 156 items were compared to each other and classified in two different ways; as *single items* (meaning no other items were similar to that specific item) or *similar items* (this included items that were either identical in wording or had substantially the same meaning). E.g. “*Physically, I feel in a good shape*” and “*Physically, I feel I am in excellent condition*” were interpreted as similar items having the same substantial meaning, whereas the following borderline case “*My muscles have felt weak*” and “*Sometimes I lose my body strength*” were interpreted as different items because they lack a common concept. Studies have shown that even small differences in the items used, i.e. asking individuals to state if they feel “*fatigued*” vs. “*lack of energy*” may lead to considerable heterogeneity in prevalence rates [34]. Thus, in order to group items as similar, they had to feature very comparable content. Despite some analytical challenges, most single items had substantial differences such as “*I can follow conversations without getting tired*” and “*I have had difficulties making decisions*”.

After comparison and classification of the total 156 items, the analysis resulted in 83 unique items, each representing a distinct attribute of fatigue. The 83 items were organized into sub-dimensions, and then we identified four central dimensions (Supplemental Fig. 1). In order to display all items in Fig. 2, the items were given shortened names (e.g., “*I feel rested*” was called “*rested*”). Despite presentation of the analysis as a linear process, the analysis was iterative and we went back and forth in the analytical process. The analysis was conducted independently by IJS and MK and was continuously discussed in the research group until agreement was reached.

### 2.3. Statistics

Results from the qualitative analysis of items, sub-dimensions and dimensions were quantified and presented with descriptive statistics. Based on the analysis of items as either single or similar (0 or 1), we estimated statistical content overlap between PROMs using the Jaccard Index (i.e. Jaccard similarity coefficient). This is a method for comparing the similarity and diversity of sample sets and is defined as the size of the intersection divided by the size of the union of sample sets (i.e. the number of shared items divided by the total number of items in any two PROMs) [35].



Note : Flow chart reported based on the PRISMA-guidelines (2009).

Fig. 1. Flow chart of literature search and inclusion of studies [50].

### 3. Results

Analysis of the eleven PROMs included in this study revealed 83 unique items. We found four item dimensions and each dimension had sub-dimensions (Fig. 2). The dimension *characteristics* appeared early in the analysis, and involves typical attributes related to fatigue awareness and perception. Two sub-dimensions were identified: *quality* and *diurnal variations*. Quality was further organized into three groups: “general” items measuring the patient’s overall feelings/condition (e.g. fatigued, tiredness and level of energy), “physical” items related to the state of the body (e.g. weak, fit, body strength), and “mental” items measuring cognitive functioning (e.g. concentration, thinking, coordination). Several of the “physical” and “mental” items are not directly related to fatigue (e.g. physically, I feel I am in excellent condition, and I have been forgetful). Diurnal variations include items measuring timing of fatigue (e.g. I get fatigued in the afternoon), and are considered a part of the *characteristics* dimension as it may describe individual fatigue patterns.

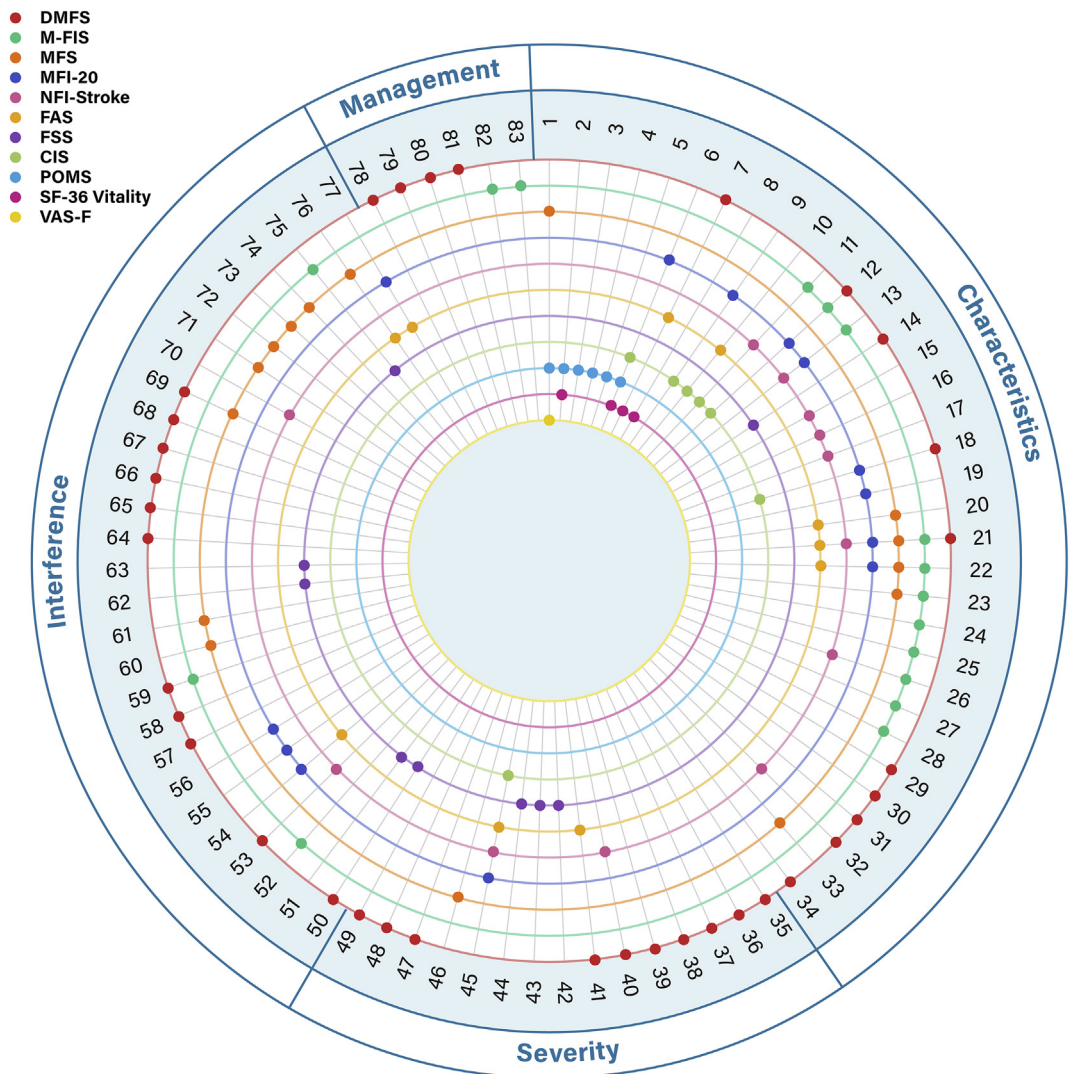
The *severity* dimension involves items related to the degree to which fatigue is bothering a person and to the intensity of fatigue. This dimension has three sub-dimensions. *Impact* measures how much suffering fatigue causes for the person (e.g. serious complaint, severe fatigue, bothered by fatigue). *Onset speed* includes items measuring how easily tired and fatigued a patient is and is related to severity because it describes the degree of vulnerability to reduced capacity. Onset speed is not considered to be a defining attribute of fatigue, and thus does not fit within the *characteristics* dimension. *Recovery time* includes items measuring the possibility and time needed to regain restoration. This

relates to *severity* because a long recovery time might reflect how seriously one’s condition is affected.

The *interference* dimension has three sub-dimensions. *Behavioral interference* items measure general activities that are affected when fatigued (e.g. hinders duties, do little, increased sleep). *Physical interference* items relate to how fatigue affects one’s physical condition (e.g., fatigue prevents sustained physical functioning, body aches when fatigued). These items explicitly measure physical reactions during fatigue, and differ from the physical items under the *quality* sub-dimension of *characteristics* since those items reflect physical functioning as a part of the overall feeling or characteristics of fatigue (e.g. “fatigue interferes with physical function” vs. “physically, I feel I am in a good condition”). *Mental interference* items measure mental function when fatigued (e.g. cannot think when fatigued, make mistakes when fatigued). Items measuring sensitivity to stress, light and sound, as well as lack of motivation are considered part of *mental interference*.

The fourth dimension contains items related to *management* of fatigue and has one sub-dimension, *coping*, which measures how the patient deals with their fatigue (e.g. by avoiding overtiredness, planning to rest, and limiting physical activities).

Of the 83 items, 75% were categorized as either characteristics or interference, and 66% appear only in one PROM (Table 1). The most-used PROMs (FSS, VAS-F and SF-36 Vitality) cover only a few dimensions. Most items in FSS measure *interference*, and VAS-F and SF-36 Vitality only measure *quality* (Table 2). The majority of all items in this study are classified in the *quality* sub-dimension or the *interference* dimension (Table 1). MFS and DMFS are the only PROMs that include an item about diurnal variations of fatigue (Table 2). None of the



**Characteristics**

**Quality**

*General Items*

- 1. Fatigued
- 2. Worn out
- 3. Sluggish
- 4. Weary
- 5. Sleepy
- 6. Tired
- 7. Energy
- 8. Full of pep
- 9. Rested

*Physical items*

- 10. Exhausted
- 11. Feel weak
- 12. Physical condition
- 13. Physical capacity
- 14. Exercise causes fatigue
- 15. Body strength
- 16. Heavy limbs
- 17. Body unable to keep up
- 18. Feel fit
- 19. Feel active

*Mental items*

- 20. Mentally exhausted
- 21. Concentration

- 22. Thinking
- 23. Memory
- 24. Less alert
- 25. Organizing thoughts
- 26. Coordination
- 27. Attention
- 28. Making decisions
- 29. Conversations
- 30. Tired by emotions
- 31. Noise-induced fatigue
- 32. Thinking-induced fatigue

**Diurnal variations**

- 33. 24-hour variations
- 34. Fatigued in afternoon

**Severity**

**Fatigue impact**

- 35. Affects whole life
- 36. Serious complaint
- 37. Complaints get worse
- 38. Suffer terribly
- 39. Severe fatigue
- 40. Tired every day
- 41. Bothered by fatigue
- 42. Causes problems
- 43. Most disabling symptom

**Onset speed**

- 44. Easily fatigued
- 45. Easily tired

**Recovery time**

- 46. Long mental recovery
- 47. Recover easily
- 48. Fatigued next day
- 49. Wake up rested

**Interference**

**Behavioral**

- 50. Duties
- 51. Work, family, social life
- 52. Limited ability
- 53. Unable to go further
- 54. Think I do little
- 55. Get little done
- 56. Few plans
- 57. Say things I regret
- 58. Other people notice
- 59. Need to rest
- 60. Decreased sleep
- 61. Increased sleep

**Physical**

- 62. Interferes with physical function

- 63. Prevents sustained function
- 64. Induces headache
- 65. Body aches

**Mental**

- 66. Letting thoughts go
- 67. Unable to think
- 68. Makes mistakes
- 69. Emotional tendency
- 70. Speech problems
- 71. Stress sensitivity
- 72. Light sensitivity
- 73. Noise sensitivity
- 74. Irritability
- 75. Low motivation
- 76. Problems starting things
- 77. Feel no desire

**Management**

**Coping**

- 78. Let myself get tired out
- 79. Avoid overtiredness
- 80. Plan to rest
- 81. Finish tasks
- 82. Limit physical activity
- 83. Pace physical activity

**Note.** Each colored circle represents a PROM and each dot represents one item. We have analyzed all the items and categorized them in dimensions (characteristics, severity, interference and management) and sub-dimensions.

**Fig. 2.** Distribution of 83 different items in eleven fatigue PROMs used to measure fatigue in stroke survivors. Organized in relation to four fatigue dimensions.

**Table 1**  
Number and percentage of items (n = 83) appearing in each dimension and sub-dimension and in one or more PROMs.

Dimension	Number of items	% of items
Sub-dimension		
Characteristics	34	41%
Quality	32	39%
General	9	11%
Physical	10	12%
Mental	13	16%
Diurnal variations	2	2%
Severity	15	18%
Fatigue impact	9	11%
Onset speed	2	2%
Recovery time	4	5%
Interference	28	34%
Behavioral	12	14%
Physical	4	5%
Mental	12	14%
Management	6	7%
Coping	6	7%
Number of PROMs containing each item		
1	55	66%
2	18	22%
3	5	6%
4	4	5%
6	1	1%

Note. Percentages calculated based on the 83 unique fatigue items identified through the qualitative analysis as displayed in Fig. 3.

PROMs in this study contain items on pre-stroke fatigue or a comprehensive assessment of diurnal variations, fatigue trajectory or aggravating factors (Table 2), which may have implications for management of PSF [2]. In total, this indicates a lack of multidimensionality among the fatigue PROMs in this study.

There is considerable lack of content overlap between PROMs, and the three most commonly used PROMs (FSS, VAS-F and SF-36 Vitality) do not overlap at all with each other (Table 3). FSS does not overlap with other PROMs except for the DMFS (38 items). MFI-20 has the highest degree of overlap with other PROMs, with 33% overlap with CIS and 27% overlap with FAS (Table 3).

#### 4. Discussion

This study shows that the most commonly used PSF PROMs do not address potentially relevant aspects specific to PSF, such as diurnal variations and pre-stroke fatigue, do not account for the multidimensional nature of PSF and lack content overlap. Our analysis

**Table 2**  
Distribution of items in each PROM organized into fatigue dimensions and sub-dimensions, % (n).

PROM	Characteristics		Severity			Interference			Management	Total number of items, N
	Quality	Diurnal variations	Fatigue impact	Onset speed	Recovery time	Behavioral	Physical	Mental	Coping	
FSS	11% (1)	0% (0)	22% (2)	11% (1)	0% (0)	22% (2)	22% (2)	11% (1)	0% (0)	9
VAS-F	100% (1)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	1
SF-36V	100% (4)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	4
CIS	86% (6)	0% (0)	0% (0)	14% (1)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	7 <sup>a</sup>
FAS	50% (5)	0% (0)	10% (1)	10% (1)	0% (0)	10% (1)	0% (0)	20% (2)	0% (0)	10
MFI-20	61% (8)	0% (0)	0% (0)	8% (1)	0% (0)	23% (3)	0% (0)	8% (1)	0% (0)	13 <sup>a</sup>
MFS	33% (5)	7% (1)	0% (0)	0% (0)	7% (1)	13% (2)	0% (0)	40% (6)	0% (0)	15
M-FIS	68% (11)	0% (0)	0% (0)	0% (0)	0% (0)	13% (2)	0% (0)	6% (1)	13% (2)	16 <sup>a</sup>
POMS	100% (6)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	6
NFI-Stroke	67% (8)	0% (0)	8% (1)	8% (1)	0% (0)	8% (1)	0% (0)	8% (1)	0% (0)	12
DMFS	26% (9)	3% (1)	20% (7)	0% (0)	9% (3)	14% (5)	6% (2)	11% (4)	11% (4)	35 <sup>a</sup>

Note. Table showing the percentage and number (n) of items in each PROM measuring the different fatigue dimensions and sub-dimensions.

<sup>a</sup> This PROM contains items that were determined in the qualitative analysis to be similar, and thus, the total number of unique items in each instrument in this table is less than in the original version of the instrument.

suggests that these PSF PROMs have important limitations, which might impair further progress in PSF research and patient care.

The analysis revealed 83 unique fatigue-related items in eleven PROMs used in stroke research. We classified the items in four main dimensions covering different fatigue-related aspects. Despite the consensus that fatigue is a multidimensional phenomenon [36–38], the three most widely used PROMs in PSF research (FSS, VAS-F, and SF-36 Vitality) mainly assess fatigue characteristics or interference. One previous study on PSF PROMs considered FSS to have insufficient face validity [28], and it is unclear why FSS remains the preferred PROM in stroke research.

One of the core characteristics of the experience of PSF is the unpredictability of when and why fatigue occurs [39]. Fatigue is not a constant symptom, but fluctuates throughout the day [2,40]. However, only two PROMs in this study had included this dimension, having one item each measuring diurnal variations. These two items lack specific details needed to investigate possible fluctuations, such as percent of fatigue during a typical day or week, as suggested in the Lynch case definition of post-stroke fatigue [41].

Most PROMs in this study, such as the FSS, were developed based on expert opinion and theory, and designed for conditions other than stroke. Many items in these PROMs refer to factors that are often directly limited by the stroke. For example, most items in the group “physical” (Fig. 2) under the quality sub-dimension are examples of normal physical stroke sequela (e.g. feel weak and low physical capacity), and these items are indistinguishable from the general consequences of the stroke and might bias the results [10,28].

The considerable lack of item overlap between the PROMs of post-stroke fatigue showed in this study, may affect the replicability and generalizability of PSF research. Current PSF research yields inconsistent findings regarding prevalence rates and associated factors, such as depression, anxiety, cognitive functioning, pain, sleep problems, lesion location, and co-morbidities [1,7,34,42]. Falconer et al. found significant variation in the prevalence of PSF depending on the PROMs and items used for detecting it, which could partly be explained by differences in terminology and descriptors used to assess fatigue [34].

Knowledge on the etiology of PSF is limited [38], but some studies report associations between PSF and biological as well as immunological factors [43]. Given the hypothesis that PSF is caused by cerebrovascular pathology, it would be of interest to discern pre-stroke fatigue conditions from PSF in order to elaborate our etiological understanding of PSF [2,44]. However, the PROMs in this study have no items measuring pre-stroke fatigue. Only a few previous studies have included pre-stroke fatigue assessment [45,46], and found that that 27–31% of stroke survivors reported having fatigue before their stroke [45,47–48]. In further studies it could be of relevance to distinguish

**Table 3**  
Overlap of items as measured by the Jaccard index for eleven PROMs.

PROM	FSS	VAS-F	SF-36 V	CIS <sup>a</sup>	FAS	MFI-20 <sup>a</sup>	MFS	M-FIS <sup>a</sup>	POMS	NFI-Stroke	DMFS <sup>a</sup>
FSS	1.000	.000	.000	.000	.000	.000	.000	.042	.000	.000	.048
VAS-F	.000	1.000	.000	.000	.000	.000	.067	.000	.167	.000	.000
SF-36V	.000	.000	1.000	.100	.077	.063	.000	.000	.250	.000	.026
CIS <sup>a</sup>	.000	.000	.100	1.000	.133	.333	.000	.095	.083	.118	.050
FAS	.000	.000	.077	.133	1.000	.278	.190	.083	.000	.100	.071
MFI-20 <sup>a</sup>	.000	.000	.063	.333	.278	1.000	.077	.160	.056	.136	.067
MFS	.000	.067	.000	.000	.190	.077	1.000	.107	.050	.038	.042
M-FIS <sup>a</sup>	.042	.000	.000	.095	.083	.160	.107	1.000	.000	.167	.063
POMS	.000	.167	.250	.083	.000	.056	.050	.000	1.000	.000	.000
NFI-Stroke	.000	.000	.000	.118	.100	.136	.038	.167	.000	1.000	.093
DMFS <sup>a</sup>	.048	.000	.026	.050	.071	.067	.042	.063	.000	.093	1.000

Note: The Jaccard index measures similarity between two sets of data with a range from 0 (no overlap) to 1 (total overlap), and is calculated as the number of shared items divided by the total number of items in any two PROMs. The PROMs are arranged by the most used to the least used instrument.

<sup>a</sup> This PROM contains items that were determined in the qualitative analysis to be similar, and thus, the total number of unique items in each instrument in this table is less than in the original version of the instrument (i.e. CIS 7 items, MFI-20 13 items, M-FIS 16 items and DMFS 35 items).

pre-stroke fatigue from PSF. However, potential recall bias and response shift might influence the validity of PROM answers if the recall interval is lengthy, thus collecting pre-stroke data is only advisable in the acute phase after stroke [49].

There are some limitations to this study. First, we analyzed a selected group of fatigue PROMs. Inclusion of other PROMs could have yielded a different result. However, we believe that the inclusion criteria cover the most relevant PROMs, as we only included the most commonly used fatigue PROMs, or those at least partly developed for a stroke population. Second, during the analysis of items there were some borderline cases, such as whether *single* or *similar* refer to different facets of fatigue and accordingly should be classified as separate items. We carefully discussed how conservative this part of the analysis should be. Generally, we chose to classify items as similar rather than different. Thus, if anything, the results overestimate, rather than underestimate, the homogeneity and overlap between the PROMs.

## 5. Summary

The wide variety of PROMs and dimensions used to assess PSF demonstrates lack of consensus regarding what needs to be assessed when diagnosing and measuring fatigue in stroke populations. Studies on PSF should report why they used a particular fatigue PROM, and what dimension of fatigue they intended to measure (characteristics, severity, interference and/or management). In order to move the research on the etiology of PSF forward, PROMs that capture the relevant experiences of patients with PSF are needed.

## Acknowledgement

We would like to acknowledge specialized research librarians Mia Alexandra Ølnes and Fredrik Solvang Pettersen at Lovisenberg Diaconale University College for their help with the systematic literature search. We would also like to thank our graphic designer Eira Taksdal for designing Fig. 2 and our statistician Milada Småstuen at University of Oslo for cross-checking the Jaccard index.

## Sources of funding

Ingrid Johansen Skogestad is funded by the non-profit Norwegian Health Association doctoral studentship. The funding source had no involvement in conducting or reporting of this study.

## Disclosures

None.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jpsychores.2019.109759>.

## References

- [1] T.B. Cumming, A.B. Yeo, J. Marquez, L. Churilov, J.-M. Annoni, U. Badaru, et al., Investigating post-stroke fatigue: an individual participant data meta-analysis, *J. Psychosom. Res.* 113 (2018) 107–112 Oct.
- [2] A. Lerdal, L.N. Bakken, S.E. Kouwenhoven, G. Pedersen, M. Kirkevold, A. Finset, et al., Poststroke fatigue—a review, *J. Pain Symptom Manag.* 38 (2009) 928–949.
- [3] E.L. Glader, B. Stegmayr, K. Asplund, Poststroke fatigue: a 2-year follow-up study of stroke patients in Sweden, *Stroke.* 33 (2002) 1327–1333.
- [4] G. Andersen, D. Christensen, M. Kirkevold, S.P. Johnsen, Post-stroke fatigue and return to work: a 2-year follow-up, *Acta Neurol. Scand.* 125 (2012) 248–253.
- [5] M. Kirkevold, D. Christensen, G. Andersen, S.P. Johansen, I. Harder, Fatigue after stroke: manifestations and strategies, *Disabil. Rehabil.* 34 (2012) 665–670.
- [6] F. Staub, J. Bogousslavsky, Fatigue after stroke: a major but neglected issue, *Cerebrovasc. Dis.* 12 (2001) 75–81.
- [7] S. Wu, M.A. Kutubaev, H.Y. Chun, E. Cowey, A. Pollock, M.R. Macleod, M. Dennis, E. Keane, M. Sharpe, G.E. Mead, Interventions for post-stroke fatigue, *Cochrane Database Syst. Rev.* (7) (2015) CD007030, <https://doi.org/10.1002/14651858.CD007030.pub3>.
- [8] M. Nadarajah, H.T. Goh, Post-stroke fatigue: a review on prevalence, correlates, measurement, and management, *Top. Stroke Rehabil.* 22 (2015) 208–220.
- [9] S. Choi-Kwon, P.H. Mitchell, J.S. Kim, Nursing interventions for poststroke fatigue, *Stroke.* 46 (2015) e224–e227.
- [10] S.F. Tyson, P. Brown, How to measure fatigue in neurological conditions? A systematic review of psychometric properties and clinical utility of measures used so far, *Clin. Rehabil.* 28 (2014) 804–816.
- [11] A.J. Dittner, S.C. Wessely, R.G. Brown, The assessment of fatigue: a practical guide for clinicians and researchers, *J. Psychosom. Res.* 56 (2004) 157–170.
- [12] C.A.C. Prinsen, L.B. Mokkink, L.M. Bouter, J. Alonso, D.L. Patrick, H.C.W. de Vet, et al., COSMIN guideline for systematic reviews of patient-reported outcome measures, *Qual. Life Res.* 27 (2018) 1147–1157.
- [13] C.B. Terwee, C.A.C. Prinsen, A. Chiarotto, M.J. Westerman, D.L. Patrick, J. Alonso, et al., COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study, *Qual. Life Res.* 27 (2018) 1159–1170.
- [14] A. Chiarotto, R.W. Ostelo, M. Boers, C.B. Terwee, A systematic review highlights the need to investigate the content validity of patient-reported outcome measures for physical functioning in patients with low back pain, *J. Clin. Epidemiol.* 95 (2018) 73–93.
- [15] S.P. McKenna, Measuring patient-reported outcomes: moving beyond misplaced common sense to hard science, *BMC Med.* 9 (2011) 86.
- [16] N. Kruithof, M.H. Van Cleef, S.M. Rasquin, T.J. Bovend'Eerd, Screening poststroke fatigue; feasibility and validation of an instrument for the screening of poststroke fatigue throughout the rehabilitation process, *J. Stroke Cerebrovasc. Dis.* 25 (2016) 188–196.
- [17] E.I. Fried, The 52 symptoms of major depression: lack of content overlap among seven common depression scales, *J. Affect. Disord.* 208 (2017) 191–197.
- [18] L.B. Krupp, N.G. LaRocca, J. Muir-Nash, A.D. Steinberg, The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus, *Arch. Neurol.* 46 (1989) 1121–1123.
- [19] J.E. Ware Jr., C.D. Sherbourne, The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection, *Med. Care.* 30 (1992) 473–483.
- [20] J.H. Vercoolen, C.M. Swanink, J.F. Fennis, J.M. Galama, J.W. van der Meer, G. Bleijenberg, Dimensional assessment of chronic fatigue syndrome, *J. Psychosom. Res.* 38 (1994) 383–392.

- [21] H.J. Michielsen, J. de Vries, G.L. van Heck, F.J.R. van de Vijver, K. Sijtsma, Examination of the dimensionality of fatigue: the construction of the Fatigue Assessment Scale (FAS), *Eur. J. Psychol. Assess.* 20 (2004) 39–48.
- [22] E.M. Smets, B. Garssen, B. Bonke, J.C. De Haes, The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue, *J. Psychosom. Res.* 39 (1995) 315–325.
- [23] B. Johansson, A. Starmark, P. Berglund, M. Rodholm, L. Ronnback, A self-assessment questionnaire for mental fatigue and related symptoms after neurological disorders and injuries, *Brain Inj.* 24 (2010) 2–12.
- [24] J.D. Fisk, P.G. Ritvo, L. Ross, D.A. Haase, T.J. Marrie, W.F. Schleich, Measuring the functional impact of fatigue: initial validation of the fatigue impact scale, *Clin. Infect. Dis.* 18 (Suppl. 1) (1994) S79–S83.
- [25] D.M. McNair, M. Lorr, An analysis of mood in neurotics, *J. Abnorm. Psychol.* 69 (1964) 620–627.
- [26] R.J. Mills, C.A. Young, J.F. Pallant, A. Tennant, Development of a patient reported outcome scale for fatigue in multiple sclerosis: the neurological fatigue index (NFI-MS), *Health Qual. Life Outcomes* 8 (2010) 22.
- [27] A.C. Visser-Keizer, A. Hogenkamp, H.J. Westerhof-Evers, L.J.L. Egberink, J.M. Spikman, Dutch multifactor fatigue scale: a new scale to measure the different aspects of fatigue after acquired brain injury, *Arch. Phys. Med. Rehabil.* 96 (2015) 1056–1063.
- [28] G. Mead, J. Lynch, C. Greig, A. Young, S. Lewis, M. Sharpe, Evaluation of fatigue scales in stroke patients, *Stroke*. 38 (2007) 2090–2095.
- [29] R.J. Mills, J.F. Pallant, M. Koufali, A. Sharma, S. Day, A. Tennant, et al., Validation of the neurological fatigue index for stroke (NFI-stroke), *Health Qual. Life Outcomes* 10 (2012) 51.
- [30] M. Sandelowski, Triangles and crystals: on the geometry of qualitative research, *Res. Nurs. Health* 18 (1995) 569–574.
- [31] J.D. Comins, M.R. Krogsgaard, J. Brodersen, Ensuring face validity in patient-related outcome scores—a matter of content, *Knee*. 20 (2013) 72–78.
- [32] S. Elo, H. Kyngas, The qualitative content analysis process, *J. Adv. Nurs.* 62 (2008) 107–115.
- [33] C. Houghton, K. Murphy, B. Meehan, J. Thomas, D. Brooker, D. Casey, From screening to synthesis: using nvivo to enhance transparency in qualitative evidence synthesis, *J. Clin. Nurs.* 26 (2017) 873–881.
- [34] M. Falconer, S. Walsh, J.A. Harbison, Estimated prevalence of fatigue following stroke and transient ischemic attack is dependent on terminology used and patient gender, *J. Stroke Cerebrovasc. Dis.* 19 (2010) 431–434.
- [35] M. Levandowsky, D. Winter, Distance between Sets, *Nature*. 234 (1971) 34.
- [36] L. Whitehead, The measurement of fatigue in chronic illness: a systematic review of unidimensional and multidimensional fatigue measures, *J. Pain Symptom Manag.* 37 (2009) 107–128.
- [37] J.L. Hinkle, K.J. Becker, J.S. Kim, S. Choi-Kwon, K.L. Saban, N. McNair, et al., Poststroke fatigue: emerging evidence and approaches to management: a scientific statement for healthcare professionals from the American Heart Association, *Stroke*. 48 (2017) (e159–e70).
- [38] M. Acciarresi, J. Bogousslavsky, M. Paciaroni, Post-stroke fatigue: epidemiology, clinical characteristics and treatment, *Eur. Neurol.* 72 (2014) 255–261.
- [39] G. Eilertsen, H. Ormstad, M. Kirkevold, Experiences of poststroke fatigue: qualitative meta-synthesis, *J. Adv. Nurs.* 69 (2013) 514–525.
- [40] J.H. White, K.R. Gray, P. Magin, J. Attia, J. Sturm, G. Carter, et al., Exploring the experience of post-stroke fatigue in community dwelling stroke survivors: a prospective qualitative study, *Disabil. Rehabil.* 34 (2012) 1376–1384.
- [41] J. Lynch, G. Mead, C. Greig, A. Young, S. Lewis, M. Sharpe, Fatigue after stroke: the development and evaluation of a case definition, *J. Psychosom. Res.* 63 (2007) 539–544.
- [42] S. Wu, A. Barugh, M. Macleod, G. Mead, Psychological associations of poststroke fatigue: a systematic review and meta-analysis, *Stroke*. 45 (2014) 1778–1783.
- [43] R. Aarnes, J. Stubberud, A. Lerdal, A literature review of factors associated with fatigue after stroke and a proposal for a framework for clinical utility, *Neuropsychol. Rehabil.* (2019) 1–28.
- [44] G.E. Mead, C. Graham, P. Dorman, S.K. Bruins, S.C. Lewis, M.S. Dennis, et al., Fatigue after stroke: baseline predictors and influence on survival. Analysis of data from UK patients recruited in the international stroke trial, *PLoS One* 6 (2011) e16988.
- [45] A. Lerdal, L.N. Bakken, E.F. Rasmussen, C. Beiermann, S. Ryen, S. Pynnten, et al., Physical impairment, depressive symptoms and pre-stroke fatigue are related to fatigue in the acute phase after stroke, *Disabil. Rehabil.* 33 (2011) 334–342.
- [46] F. McKechnie, S. Lewis, G. Mead, A pilot observational study of the association between fatigue after stroke and C-reactive protein, *J. R. Coll. Physicians Edinb.* 40 (2010) 9–12.
- [47] T. Egerton, A. Hokstad, T. Askim, J. Bernhardt, B. Indredavik, Prevalence of fatigue in patients 3 months after stroke and association with early motor activity: a prospective study comparing stroke patients with a matched general population cohort, *BMC Neurol.* 15 (2015) 181.
- [48] S.S. Wang, J.J. Wang, P.X. Wang, R. Chen, Determinants of fatigue after first-ever ischemic stroke during acute phase, *PLoS One* 9 (2014) e110037.
- [49] E. Kwong, N. Black, Retrospectively patient-reported pre-event health status showed strong association and agreement with contemporaneous reports, *J. Clin. Epidemiol.* 81 (2017) 22–32.
- [50] D. Liberati Moher, A. Tetzlaff, J.D.G. Altman, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement, *J. Clin. Epidemiol.* 62 (2009) 1006–1012.