


RESEARCH

Open Access



Validation of a short version of the Lee fatigue scale in adults living in Norway: a cross-sectional population survey

Anners Lerdal^{1,2*} , Caryl Gay^{2,3}, Tore Bonsaksen^{4,5}, Øivind Ekeberg⁶, Trine Grimholt^{7,8}, Trond Heir^{9,10}, Anders Kottorp¹¹, Kathryn A. Lee³, Laila Skogstad¹² and Inger Schou-Bredal¹³

Abstract

Background Due to the nature of fatigue, a brief reliable measure of fatigue severity is needed. Thus, the aim of our study was to evaluate a short version of the Lee Fatigue Scale (LFS) in the Norwegian general population.

Methods This cross-sectional survey consists of a representative sample from the Norwegian population drawn by The National Population Register in Norway. The study is part of a larger study (NORPOP) aimed at collecting normative data from several questionnaires focused on health in adults living in Norway. Registered citizens between 18 and 94 years of age were randomly selected stratified by age, sex and geographic region. Of the 4971 respondents eligible for the study, 1792 (36%) responded to the survey. In addition to age and sex, we collected responses on a 5-item version of the LFS measuring current fatigue severity. The psychometric properties focusing on internal structure and precision of the LFS items were analyzed by a Rasch rating scale model.

Results Complete LFS scores for analyses were available for 1767 adults. Women had higher LFS-scores than men, and adults < 55 years old had higher scores than older respondents. Our analysis of the LFS showed that the average category on each item advanced monotonically. Two of the five items demonstrated misfit, while the three other items demonstrated goodness-of-fit to the model and uni-dimensionality. Items #1 and #4 (tired and fatigue respectively) showed differential item functioning (DIF) by sex, but no items showed DIFs in relation to age. The separation index of the LFS 3-item scale showed that the sample could be separated into three different groups according to the respondents' fatigue levels. The LFS-3 raw scores correlated strongly with the Rasch measure from the three items. The core dimensions in these individual items were very similarly expressed in the Norwegian language version and this may be a threat to the cultural-related or language validity of a short version of the LFS using these particular items.

Conclusions The study provides validation of a short LFS 3-item version for estimating fatigue in the general population.

Keywords Fatigue, Lee Fatigue Scale, Low energy, Psychometrics, Rasch analysis, Population survey

*Correspondence:

Anners Lerdal

anners.lerdal@medisin.uio.no

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

The Lee Fatigue Scale (LFS) [1] is frequently used to measure fatigue severity in a variety of patient populations. It has mostly been used to measure fatigue severity in patients with cancer [2–4], human immunodeficiency virus (HIV) [5], osteoarthritis [6, 7], stroke [7], obstructive sleep apnea [8], patients undergoing dialysis [9], and patients treated in intensive care units [10] as well as in pregnant women [11] and parents of preterm infants [12]. Several studies have shown that the LFS is sensitive in measuring diurnal patterns of fatigue, distinguishing between morning and evening levels of fatigue [3, 5, 13, 14].

When the psychometric properties of the original 13-item version of fatigue items on the LFS were evaluated with a Rasch analysis approach that compared fatigue scores obtained in the morning and evening [15], nine of the items showed satisfactory acceptable goodness-of-fit for the morning and 10 items for the evening measures. These items also demonstrated an acceptable level of uni-dimensionality and were able to differentiate the patients into four levels of fatigue severity.

Because of the nature of fatigue, it is important to reduce participant burden when conducting research on fatigue. Reducing the burden on participants is also relevant in the context of public health surveys, where participants are asked to respond to a broad range of questionnaires. A previous study found that most of the frequently used instruments to assess fatigue consisted of 10–20 items [16], which indicates that the development and validation of short fatigue instruments may be particularly useful. In a study that aimed to develop a short version of the LFS [17], a sub-set of five of the original 13 items was found to be sufficient to measure fatigue severity and satisfy criteria for internal scale validity, uni-dimensionality and separation of patients' fatigue into three distinct levels, which is sufficient for many clinical and research purposes. A short version in Norwegian of the LFS has been used in clinical studies of Norwegian patients, but the psychometric properties has not been tested in the Norwegian general population. This is need in order to know if normative data from the general population can be used as reference values in relation to LSF scores in clinical studies. Furthermore, the LFS, in contrast to other fatigue scales, has been used for two daily measurements over several days to describe morning and evening fatigue [3, 5, 18]. Since the patients have to fill in the LFS so frequently, it is of particular interest to have a valid and reliable short version of the instrument.

Although fatigue is often studied in chronic illness populations, several studies show that fatigue is also experienced in the general population, including people without current diseases. Depending on the definition

and cut-off value for fatigue cases, the prevalence of current fatigue in the general population has been reported to vary between 5 and 30% [19–21], and the prevalence of chronic fatigue (lasting more than six months) has ranged from 6 to 30% [19, 20, 22]. Given the large variation in prevalence between these studies the validity of cut-off scores designating a positive case of fatigue is of concern. Due to the nature of fatigue, a short fatigue instrument that requires little energy for respondents to complete and with satisfactory psychometric properties is warranted to survey fatigue in the general population. It is important that fatigue measures be validated for use not only in various patient populations, but in the general population as well. No previous validation studies of the LFS in the general population exist. Thus, the aim of this study was to examine the psychometric properties focusing on internal structure and precision [23] of a short version of the LFS in the general population in Norway.

Methods

A representative sample of the Norwegian population was surveyed in a cross-sectional study in order to establish normative data for a number of different instruments measuring different symptoms, health behaviors and attitudes. The National Population Register in Norway drew a representative sample of the Norwegian population. All registered citizens in Norway between 18 and 94 years of age were eligible to participate, and the sample was stratified according to age, sex, and geographic region. A total of 5500 citizens were invited to participate, including citizens from all of the country's 19 counties. A more detailed description of the recruitment process has been published elsewhere [24, 25].

Procedures

The sample was mailed the questionnaires in 2015 with information about the study and a pre-paid return envelope. Each individual was mailed two reminders. Of the 4,971 survey recipients, 1,792 (36%) returned their survey. Between responders and non-responders, there were no significant differences in mean age, gender proportions or proportions living in rural versus urban areas [25]. The proportion of the sample in active work was 66%, compared to 67% in the general population [26]. Among responders and non-responders alike, 17% lived alone. Among responders, 1.3% were without work and 53% had higher education, compared to 4.4% and 41.0% in the general population, respectively [24]. In view of these comparisons, the sample was deemed fairly representative of the general Norwegian population.

Measurements

Fatigue severity was measured with a 5-item version of the LFS [1]. Each of the items has two anchor statements, and participants responded on an 11-point numeric rating scale, with responses ranging from 0 to 10. The five items anchor statements were: *not at all tired – extremely tired* (item #1), *not at all fatigued – extremely fatigued* (item #4), *not at all worn out – extremely worn out* (item #5), *not at all bushed – extremely bushed* (item #11), and *not at all exhausted – extremely exhausted* (item #12). The item numbers refer to the original LFS version [1]. The mean score of the 5 item scores constitutes each participant's fatigue severity score. Demographic information on age (in years) and sex (male, female) was also collected.

Ethics

The Regional Committee for Medical and Health Research Ethics South East was consulted prior to distributing the survey. Because this was an anonymous survey conducted by mail, the committee did not require formal review of the study. Returned surveys signified consent to participate.

Statistical analysis

Descriptive statistics were used to summarize sample demographics and fatigue scores.

The psychometric analysis of the LFS was guided by a Rasch rating scale model [27]. The transformed 11-category raw scores (0 to 10) from the five LFS items were analyzed using the WINSTEPS Rasch computer software program, version 3.91.0.0 [28]. The analyses were performed using a systematic stepwise approach similar to that used in previous studies [5].

In *Step 1*, an evaluation of the psychometric properties of the fatigue rating scale was conducted to determine whether the average measures for each category on each item advanced monotonically, i.e. whether the Outfit Mean Square (*MnSq*) values were < 2.0 for each of the step calibrations [29].

Step 2 aimed to evaluate the fit of the item responses [27]. Any item that did not show acceptable goodness-of-fit to the model was removed, and the psychometric properties of the remaining items were re-analyzed until all remaining items demonstrated acceptable goodness-of-fit. For this study, acceptable goodness-of-fit was defined as *Infit MnSq* values between 0.7 and 1.3, which is stricter than the suggested guidelines for surveys using rating scales [30]. We choose to focus on *infit* statistics in this study as they are considered the most informative measure of goodness-of-fit given that they focus on the degree of fit in the most typical observations in the data

[31]. We also focus on *MnSq* values rather than standardized *z*-values, as *z*-values are highly influenced by sample size [32]. In *Step 2*, we also evaluated the level of unidimensionality in the generated LFS measure by a principal component analysis (PCA) of the residuals, with the criterion that the first latent dimension should explain at least 50% of total variance [33]. We also monitored the standardized item residual correlations between items using the Winsteps output and considered a correlation coefficient between items of 0.5 or higher (equal to a shared variance of 25% or more) to be a threat to local independence.

Step 3 evaluated aspects of person response validity. Person goodness-of-fit was defined as *Infit MnSq* values less than 1.4 logits or associated with a *z*-value < 2 , accepting that 5% of the sample may by chance fail to demonstrate acceptable goodness-of-fit without threatening evidence of person response validity [34–36].

In *Step 4*, differential item functioning (DIF) analyses were performed in order to evaluate the stability of the LFS response patterns in relation to age and sex using the Mantel–Haenszel statistics for polytomous scales using log-odds estimators [37, 38] as reported from the WINSTEPS program ($p < 0.01$ with Bonferroni correction). If DIF was detected for an item, a supplementary analysis of the impact of the item's DIF was then calculated, based on a standardized *z*-comparison of the individual Rasch measures produced by the generic vs the sample-specific item hierarchies. Item DIF was considered to have minimal impact if 5% or less of the sample had a change of more than ± 1.96 *z*-value in their Rasch measures.

Step 5 assessed several aspects of the fatigue scale's reliability. The unidimensional scale's ability to separate participants into distinct groups was estimated using the person separation index. The Rasch-equivalent person reliability coefficient, as well as the Cronbach's alpha reliability coefficient based on the LFS raw item scores, were also reported for the final unidimensional scale. Finally, a Pearson's correlation coefficient was used to evaluate the relationship between the LFS mean raw score (calculated as the mean of the item raw scores) and the Rasch-generated measures.

Results

Fatigue in the general population

Of the 1792 survey respondents, 1767 had complete LFS scores for analysis. The mean age of respondents was 53.2 years (± 16.6 SD), with a range 18–94 years. Fewer men (46.9%) responded compared to women (53.1%). As shown in Table 1, women had a higher mean LFS score than men ($p < 0.001$), and adults < 55 years old had a higher mean LFS score than adults 55 years and older

Table 1 Lee Fatigue Scale (LFS) mean scores for men and women ($n = 1767$)

Characteristic	% (n)	LFS score mean (SD)	p-value
Total sample	100% (1767) ^a	3.33 (2.54)	
Sex	99% (1755) ^b		< .001
Men	47% (822)	2.99 (2.38)	
Women	53% (933)	3.65 (2.64)	
Age	99% (1748) ^c		< .001
18–54 years	50% (874)	3.60 (2.60)	
≥ 55 years	50% (874)	3.07 (2.46)	

^a 25 of the 1792 respondents answered < 4 LFS items, so a mean score was not computed

^b 13 of the 1767 respondents with LFS scores had missing data for self-identified gender

^c 19 of the 1767 respondents with LFS scores had missing data for age

Table 2 Descriptive statistics for five LFS items ($n = 1767^a$)

Item	Missing(n)	Mean (SD)	% with minimum score of 0	% with maximum score of 10
1. Tired	0	4.51 (2.69)	8.1%	3.1%
4. Fatigued	0	3.25 (2.84)	22.9%	2.2%
5. Worn out	4	3.30 (2.82)	21.7%	2.0%
11. Bushed	1	2.95 (2.73)	26.4%	1.8%
12. Exhausted	3	2.66 (2.74)	31.7%	1.8%

^a 25 participants were excluded from the analysis because they were missing data on > 1 item

($p < 0.001$). Descriptive statistics for each of the LFS items are shown in Table 2.

Rasch analysis of LFS psychometric properties

Table 3 summarizes the findings of the Rasch analysis. In Step 1, the rating scale of the LFS demonstrated acceptable outcomes in relation to the set criteria. When analyzing the infit mean square statistics for the five items in Step 2, two items (items #1 tired and #5 worn out) demonstrated unacceptable fit statistics (see Table 3). We excluded these items and re-ran the analysis with the remaining three items. In the second iteration, all three remaining items demonstrated acceptable goodness-of-fit to the model. The uni-dimensionality of the 3-item LFS scale was also acceptable (83.6%).

In Step 3, a proportion of the sample close to the set criteria demonstrated misfit to the Rasch model in the 3-item LFS scale (6.4%). In Step 4, none of the three items demonstrated DIF in relation to Gender or Age (using median splits), so no supplementary analysis of the impact of item DIF was performed.

In Step 5, the separation index of the LFS scale decreased (from 3.45 to 2.49) after deleting the two items demonstrating misfit in Step 2, but still exceeded our set criterion. There was a strong correlation between the LFS raw score and the Rasch measure, which also remained stable (from 0.96 to 0.97) after deleting the two misfitting items. Person reliability for the Rasch measure and Cronbach's alpha for the raw score were also acceptable but decreased after deleting the two misfitting items. Given the proportions of minimum and maximum scores, there was a minimal ceiling effect, but some evidence of a floor effect, which worsened after deleting the two misfitting items (from 7.3% to 18.0% having minimum scores).

Discussion

In this population-based study, a 3-item version of the Lee Fatigue Scale demonstrated acceptable psychometric properties for assessing fatigue in the general population. These findings provide further evidence for the psychometric properties of the LFS and support prior studies where we have shown that short versions of the LFS have satisfactory internal scale validity, uni-dimensionality and sensitivity to separate individuals with fatigue into three different groups (low, medium and high levels of fatigue) [17, 39]. Thus, a shorter measure of fatigue severity may be as psychometrically sound as longer measures. This is particularly important for measures of fatigue where the level of fatigue may impact on the respondents' ability and capacity to participate in fatigue studies. While previous studies have assessed the psychometric properties of the LFS-VAS with scores converted from the 100 mm VAS to numeric scores from 0–10, the present study showed that the average measures for each response category advanced monotonically. The floor effect of the LFS-3 item was relatively high and higher than in previous studies using other versions of the LFS. However, this can be explained by the fact that the current sample represents the general population where fatigue is expected to be considerably lower compared to samples of clinical patients with acute or chronic illness. Thus, the floor effect in this study may not be considered a psychometric weakness, as it is expected that not everyone in the population would experience even a low perceived level of fatigue. While a previous psychometric assessment of the LFS was performed in a sample of women, this current study included men and demonstrated that two of the items from LFS measure showed DIFs biased by sex, i.e. women endorsed "fatigue" more easily than men, while men endorsed "tired" more easily than women.

The core dimension of fatigue severity used in the 5-item English language LFS is reflected in the similar but unique wording of items (tired, fatigued, worn out, bushed and exhausted). The same core dimension may

Table 3 Overview of the Statistical Approach, Criteria, and Results of the Rasch Analysis of the LFS short form used in the general population (n = 1767)

Aspect of validity evidence measured	Statistical approach and criteria	Initial LFS short form scale (5 items)	Final LFS short form scale (3 items)	Comparison LFS short form scale (Bragstad et al. 2020) [7] (3 items)
Step 1: Internal scale validity	<p>Rating scale functioning The average measures for each response category should advance monotonically associated with outfit <i>MnSq</i> values < 2.0 for each of the step calibrations</p> <p>Item goodness-of-fit statistics A sample-size adjusted criterion for item goodness-of-fit requiring outfit <i>MnSq</i> values between 0.7 and 1.3</p>	All criteria met	All criteria met	All criteria met
Step 2: Internal scale validity	<p>Principal component analysis of the residuals The criterion was set for at least 50% of the total variance to be explained by the first latent variable</p> <p>Local independence among items The criterion was set as all correlation coefficients between item residuals should be 0.5 or less (equal to a shared variance of 25% or less)</p>	<p>#12 exhausted: <i>MnSq</i> = 1.29</p> <p>#11 bushed: <i>MnSq</i> = 0.99</p> <p>#5 worn out: <i>MnSq</i> = 0.64</p> <p>#1 tired: <i>MnSq</i> = 1.33</p> <p>#4 fatigued: <i>MnSq</i> = 0.77</p> <p>83.3%</p>	<p>#12 exhausted: <i>MnSq</i> 1.04</p> <p>#11 bushed: <i>MnSq</i> 0.86</p> <p>#4 fatigued: <i>MnSq</i> 1.05</p> <p>83.6%</p>	<p>#5 worn out: <i>MnSq</i> 0.87</p> <p>#1 tired: <i>MnSq</i> 1.16</p> <p>#4 fatigued: <i>MnSq</i> 0.87</p> <p>88.1%</p>
Step 3: Person response validity	<p>Person goodness-of-fit statistics A criterion of < 5% of the sample has unacceptable person goodness-of-fit, as indicated by outfit <i>MnSq</i> values lower than 1.4 associated with a standardized z value < 2.0</p>	<p>The largest standardized residual correlation coefficient was 0.02 between the items #4 Bushed and #5 Exhausted</p>	<p>All standardized item residual correlation coefficients were negative</p>	<p>Not reported</p>
Step 4: Item stability	<p>Differential item functioning (DIF) No item DIF should exceed a criterion of $p = .01$</p>	-	<p>113 (6.4%)</p> <p>Gender: None</p> <p>Age: None</p>	<p>86 (4.9%)</p> <p>Gender: Item #4 Fatigued: less challenging to endorse for men (mean of 54.62 logits) than for women (mean of 56.38 logits, $p < .01$)</p> <p>Item #1 Tired: Less challenging to endorse for women (mean of 38.57 logits) than for men (mean of 40.83 logits, $p < .01$)</p> <p>Age: None</p>

Table 3 (continued)

Aspect of validity evidence measured	Statistical approach and criteria	Initial LFS short form scale (5 items)	Final LFS short form scale (3 items)	Comparison LFS short form scale (Bragstad et al. 2020) [7] (3 items)
Step 5: Test reliability	Person-separation reliability	Separation index = 3.45 Correlation between raw score and Rasch measure = .96 Person reliability (Rasch) = .92 Cr- α (raw score) = .95 Maximum score = 20 (1.1%) Minimum score = 129 (7.3%)	Separation index = 2.59 Correlation between raw score and Rasch measure = .97 Person reliability (Rasch) = .87 Cr- α (raw score) = .93 Maximum score = 20 (1.1%) Minimum score = 319 (18.0%)	Separation index = 3.55 Correlation between raw score and Rasch measure = .97 Person reliability (Rasch) = .93 Cr- α (raw score) = .94 Maximum score = 29 (1.6%) Minimum score = 135 (7.6%)
	Person reliability (Rasch) Cronbach's alpha coefficient for raw scores			

MnSq Mean square

not be reflected when using Norwegian words with very similar meanings. Language may indeed impact the cultural validity of a short Norwegian version of LFS with the particular combination of items used in our study. Conceptual translations for cross-cultural comparisons are complex and would require additional studies to ensure that items or concepts in different languages address the same construct.

Slightly higher levels of fatigue than expected were found in persons demonstrating misfit to the Rasch model. More data may be needed here to assess whether any systematic pattern among the persons demonstrating misfit can be detected and allow for specific subgroup analysis. Earlier studies using the LFS have shown item calibration differences between diagnostic groups [7], which may be hidden in general adult population samples, such as the sample used in this study.

Some of the items had a relatively high proportion of minimum scores. However, since our study surveyed fatigue in the general population, we expected that a large proportion did not experience fatigue.

A significant challenge for building cumulative knowledge of fatigue in different groups of patients and populations is the absence of consensus among researchers about which fatigue instrument should be used. There are several promising initiatives to develop international recommendations for measuring fatigue and psychometrically sound instruments that are available in many languages, such as the 13-item PROMIS SF v1.0 Fatigue Scale [40]. While such measures may have some advantages, the psychometrically valid 3-item LFS may be more suitable when a shorter scale is needed.

Strengths

A relatively large sample drawn from the general population provides evidence of validity of a short three-item version of the LFS that is also sensitive enough to differentiate the sample into three distinct groups by level of fatigue.

Limitations

Although our findings in a Norwegian version of the LFS are similar to those in the English version, caution in interpreting the findings is warranted. Translating fatigue and similar symptoms from the original English version is difficult and can be problematic when attempting to generalize our findings to other non-English speaking cultures. Even American English could have subtle differences in meaning compared to British English wording for items in the LFS. The Norwegian LFS items evaluated in this study originate from a translation done for a prior study with a diagnosis-specific female sample [41].

In addition, the translation process was not performed according to the current standards described in the COSMIN framework [42]. Furthermore, since the survey did not include the full original Lee Fatigue Scale we only had the opportunity to validate this short version based on data generated using the pre-selected items. Thus, the generalization and applicability of the LFS scale in general to a wider population can be questioned. With a limited number of fatigue items in this study, there is also a risk that some other LFS items, now not included, could have demonstrated acceptable goodness-of-fit to the Rasch model. This could also impact on evidence based on test content, potentially missing additional essential aspects of fatigue. Various validity studies of the LFS scale have also suggested different item combinations to measure the level of fatigue in a valid manner [17, 39]. This can be viewed as a limitation especially when comparing outcomes from different studies using different item combinations. A solution to this challenge could be to apply a Rasch analysis model to generate a stable item bank or item hierarchy across large samples with diversity in diagnoses and languages that demonstrate acceptable validity, evidence of stability and internal construct validity. Providing the generic weights of each item calibration can then be used to select and use subsets of items for specific studies and samples, and still generate comparable measures. Such systematic approach should also be better grounded in the COSMIN guidelines regarding the full validation process, including translation.

Conclusions

We surveyed Norwegian adults across the age spectrum to evaluate the psychometric properties of a Norwegian version of the LFS. Our results provide evidence for the validity of a short three-item version of the tool. Therefore, this version of the instrument can be readily applied to measure levels of fatigue in the general Norwegian population.

Abbreviations

DIF	Differential Item Functioning
HIV	Human Immunodeficiency Virus
LFS	Lee Fatigue Scale
NORPOP	The Norwegian Population Study
MnSq	Mean Square
VAS	Visual Analogue Scale

Acknowledgements

Not applicable

Authors' contributions

ISB designed the study, TB, AL, ØE, TG, TH, LS and ISB contributed to the collection of the data. AK performed the statistical analysis, AL, AK, and CG drafted the manuscript. AL, AK, CG and KL participated in interpretation of the data. All authors read, gave input to and approved the final manuscript.

Funding

Open access funding provided by University of Oslo (incl Oslo University Hospital) The authors received no financial support for the research, authorship, and/or publication of this article.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

No person-identifying information was collected. Those who consented to participate did so by returning their completed questionnaires in a sealed envelope. The requirement for informed consent and approval was waived by the Ethics Committee of Health South-East because the anonymous data collected. The study was performed in accordance with the Declarations of Helsinki.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Research Department, Lovisenberg Diaconal Hospital, Oslo, Norway. ²Department of Interdisciplinary Health Sciences, Faculty of Medicine, Institute of Health and Society, University of Oslo, Oslo, Norway. ³Department of Family Health Care Nursing, University of California, San Francisco, USA. ⁴Department of Health and Nursing, Faculty of Social and Health Sciences, Inland Norway University of Applied Sciences, Elverum, Norway. ⁵Department of Health, Faculty of Health Studies, VID Specialized University, Stavanger, Norway. ⁶Psychosomatic and CL Psychiatry, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway. ⁷Department of Health, Faculty of Health Studies, VID Specialized University, Oslo, Norway. ⁸Department of Acute Medicine, Oslo University Hospital, Oslo, Norway. ⁹Norwegian Center for Violence and Traumatic Stress Studies, Oslo, Norway. ¹⁰Institute of Clinical Medicine, University of Oslo, Oslo, Norway. ¹¹Faculty of Health and Society, Malmö University, Malmö, Sweden. ¹²Faculty of Health Sciences, Department of Health and Care Sciences, UiT The Arctic University of Norway, Tromsø, Norway. ¹³Department of Public Health Science, Faculty of Medicine, Institute of Health and Society, University of Oslo, Oslo, Norway.

Received: 4 November 2022 Accepted: 20 October 2023

Published online: 31 October 2023

References

- Lee KA, Hicks G, Nino-Murcia G. Validity and reliability of a scale to assess fatigue. *Psychiatry Res.* 1991;36(3):291–8. [https://doi.org/10.1016/0165-1781\(93\)90022-9](https://doi.org/10.1016/0165-1781(93)90022-9).
- Golan-Vered Y, Pud D. Chemotherapy-induced neuropathic pain and its relation to cluster symptoms in breast cancer patients treated with paclitaxel. *Pain Pract.* 2013;13(1):46–52. <https://doi.org/10.1111/j.1533-2500.2012.00554.x>.
- Lin Y, Bailey DE, Xiao C, Hammer M, Paul SM, Cooper BA, Conley YP, Levine JD, Kober KM, Miaskowski C. Distinct Co-occurring Morning and Evening Fatigue Profiles in Patients With Gastrointestinal Cancers Receiving Chemotherapy. *Cancer Nurs.* 2022. <https://doi.org/10.1097/NCC.0000000000001148>.
- Hugoy T, Lerdal A, Rustoen T, Oksholm T. Predicting postoperative fatigue in surgically treated lung cancer patients in Norway: a longitudinal 5-month follow-up study. *BMJ Open.* 2019;9(9):e028192. <https://doi.org/10.1136/bmjopen-2018-028192>.
- Lerdal A, Gay CL, Aouizerat BE, Portillo CJ, Lee KA. Patterns of morning and evening fatigue among adults with HIV/AIDS. *J Clin Nurs.* 2011;20(15–16):2204–16. <https://doi.org/10.1111/j.1365-2702.2011.03751.x>.
- Lindberg MF, Miaskowski C, Rustoen T, Rosseland LA, Paul SM, Cooper BA, Lerdal A. The Impact of Demographic, Clinical, Symptom and Psychological Characteristics on the Trajectories of Acute Postoperative Pain After Total Knee Arthroplasty. *Pain Med.* 2017;18(1):124–39. <https://doi.org/10.1093/pm/pnw080>.
- Bragstad LK, Lerdal A, Gay CL, Kirkevold M, Lee KA, Lindberg MF, Skogestad IJ, Hjelle EG, Sveen U, Kottorp A. Psychometric properties of a short version of Lee Fatigue Scale used as a generic PROM in persons with stroke or osteoarthritis: assessment using a Rasch analysis approach. *Health Qual Life Outcomes.* 2020;18(1):168. <https://doi.org/10.1186/s12955-020-01419-8>.
- Yang H, Engeland CG, King TS, Sawyer AM. The relationship between diurnal variation of cytokines and symptom expression in mild obstructive sleep apnea. *J Clin Sleep Med.* 2020;16(5):715–23. <https://doi.org/10.5664/jcsm.8332>.
- Chao CT, Huang JW, Chiang CK. group Cs: Functional assessment of chronic illness therapy—the fatigue scale exhibits stronger associations with clinical parameters in chronic dialysis patients compared to other fatigue-assessing instruments. *PeerJ.* 2016;4:e1818. <https://doi.org/10.7717/peerj.1818>.
- Day A, Haj-Bakri S, Lubchansky S, Mehta S. Sleep, anxiety and fatigue in family members of patients admitted to the intensive care unit: a questionnaire study. *Crit Care.* 2013;17(3):R91. <https://doi.org/10.1186/cc12736>.
- Tsai SY, Shun SC, Lai YH, Lee YL, Lee SY. Psychometric evaluation of a Chinese version of the Lee Fatigue Scale-Short Form in women during pregnancy and postpartum. *Int J Nurs Stud.* 2014;51(7):1027–35. <https://doi.org/10.1016/j.ijnurstu.2013.10.023>.
- Nordheim T, Rustoen T, Iversen PO, Nakstad B. Quality of life in parents of preterm infants in a randomized nutritional intervention trial. *Food Nutr Res.* 2016;60:32162. <https://doi.org/10.3402/fnr.v60.32162>.
- Aouizerat BE, Dhruva A, Paul SM, Cooper BA, Kober KM, Miaskowski C: Phenotypic and Molecular Evidence Suggests That Decrements in Morning and Evening Energy Are Distinct but Related Symptoms. *J Pain Symptom Manage* 2015, 50(5):599–614 e593. <https://doi.org/10.1016/j.jpainsymman.2015.05.008>
- Wright F, D'Eramo Melkus G, Hammer M, Schmidt BL, Knobf MT, Paul SM, Cartwright F, Mastick J, Cooper BA, Chen LM, et al. Predictors and Trajectories of Morning Fatigue Are Distinct From Evening Fatigue. *J Pain Symptom Manage.* 2015;50(2):176–89. <https://doi.org/10.1016/j.jpainsymman.2015.02.016>.
- Lerdal A, Kottorp A, Gay C, Aouizerat BE, Lee KA, Miaskowski C. A Rasch Analysis of Assessments of Morning and Evening Fatigue in Oncology Patients Using the Lee Fatigue Scale. *J Pain Symptom Manage.* 2016;51(6):1002–12. <https://doi.org/10.1016/j.jpainsymman.2015.12.331>.
- Hjollund NH, Andersen JH, Bech P. Assessment of fatigue in chronic disease: a bibliographic study of fatigue measurement scales. *Health Qual Life Outcomes.* 2007;5:12. <https://doi.org/10.1186/1477-7525-5-12>.
- Lerdal A, Kottorp A, Gay CL, Lee KA. Development of a short version of the Lee Visual Analogue Fatigue Scale in a sample of women with HIV/AIDS: a Rasch analysis application. *Qual Life Res.* 2013;22(6):1467–72. <https://doi.org/10.1007/s11136-012-0279-3>.
- Kober KM, Harris C, Conley YP, Dhruva A, Dokiparthi V, Hammer MJ, Levine JD, Oppegaard K, Paul S, Shin J, et al. Perturbations in common and distinct inflammatory pathways associated with morning and evening fatigue in outpatients receiving chemotherapy. *Cancer Med.* 2022. <https://doi.org/10.1002/cam4.5435>.
- van't Leven M, Zielhuis GA, van der Meer JW, Verbeek AL, Bleijenberg G: Fatigue and chronic fatigue syndrome-like complaints in the general population. *Eur J Public Health* 2010, 20(3):251–257. <https://doi.org/10.1093/eurpub/ckp113>
- Loge JH, Ekeberg O, Kaasa S: Fatigue in the general Norwegian population: normative data and associations. *J Psychosom Res* 1998, 45(1 Spec No):53–65. [https://doi.org/10.1016/s0022-3999\(97\)00291-2](https://doi.org/10.1016/s0022-3999(97)00291-2)
- Lerdal A, Wahl A, Rustøen T, Hanestad BR, Moum T. Fatigue in the general population: a translation and test of the psychometric properties of the Norwegian version of the Fatigue Severity Scale. *Scand J Public Health.* 2005;33:123–30. <https://doi.org/10.1080/14034940410028406>.

22. Martin A, Chalder T, Rief W, Braehler E. The relationship between chronic fatigue and somatization syndrome: a general population survey. *J Psychosom Res.* 2007;63(2):147–56. <https://doi.org/10.1016/j.psychres.2007.05.007>.
23. American Educational Research Association. American Psychological Association, National Council on Measurement in Education: Standards for Educational and Psychological Testing. Washington DC: American Educational Research Association; 2014.
24. Schou-Bredal I, Heir T, Skogstad L, Bonsaksen T, Lerdal A, Grimholt T, Ekeberg O. Population-based norms of the Life Orientation Test-Revised (LOT-R). *Int J Clin Health Psychol.* 2017;17(3):216–24. <https://doi.org/10.1016/j.ijchp.2017.07.005>.
25. Bonsaksen T, Lerdal A, Heir T, Ekeberg O, Skogstad L, Grimholt TK, Schou-Bredal I. General self-efficacy in the Norwegian population: Differences and similarities between sociodemographic groups. *Scand J Public Health.* 2019;47(7):10. <https://doi.org/10.1177/1403494818756701>.
26. Statistics Norway: Labour force survey, Q4. In.; 2016. <https://www.ssb.no/en/arbeid-og-lonn/statistikker/aku>.
27. Bond TG, Fox CM. Applying the Rasch model: fundamental measurement in the human sciences. 2nd ed. Mahwah, N.J.: Lawrence Erlbaum Associates Publishers; 2007.
28. Linacre JM: Winsteps - Rasch Model computer program (Version 3.9.1.0). In. Chicago: www.winsteps.com; 2016.
29. Linacre JM. Optimizing rating scale category effectiveness. *J Appl Meas.* 2002;3(1):85–106.
30. Wright BD, Linacre JM. Reasonable mean-square fit values. *Rasch Measurement Transactions.* 1994;8:2.
31. McNamara TF. Measuring second language performance. London: Longman; 1996.
32. Smith AB, Rush R, Fallowfield LJ, Velikova G, Sharpe M. Rasch fit statistics and sample size considerations for polytomous data. *BMC Med Res Methodol.* 2008;8:33. <https://doi.org/10.1186/1471-2288-8-33>.
33. Linacre JM. A User's Guide to Winsteps. Ministep Rasch-Model Computer Programs. Program Manual 3.73.0. 2011.
34. Kottorp A, Bernspang B, Fisher AG. Validity of a performance assessment of activities of daily living for people with developmental disabilities. *J Intellect Disabil Res.* 2003;47(Pt 8):597–605. <https://doi.org/10.1046/j.1365-2788.2003.00475.x>.
35. Patomella AH, Tham K, Kottorp A. P-drive: assessment of driving performance after stroke. *J Rehabil Med.* 2006;38(5):273–9. <https://doi.org/10.1080/16501970600632594>.
36. Hallgren M, Nygard L, Kottorp A. Technology and everyday functioning in people with intellectual disabilities: a Rasch analysis of the Everyday Technology Use Questionnaire (ETUQ). *J Intellect Disabil Res.* 2011;55(6):610–20. <https://doi.org/10.1111/j.1365-2788.2011.01419.x>.
37. Mantel N. Chi-square tests with one degree of freedom; Extensions of the Mantel-Haenszel procedure. *J Am Stat Assoc.* 1963;58(303):690–700.
38. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst.* 1959;22(4):719–48.
39. Lerdal A, Kottorp A, Gay CL, Lee KA. Lee Fatigue And Energy Scales: exploring aspects of validity in a sample of women with HIV using an application of a Rasch model. *Psychiatry Res.* 2013;205(3):241–6. <https://doi.org/10.1016/j.psychres.2012.08.031>.
40. PROMIS SF v1.0 Fatigue 13a [<https://www.facit.org/measures/PROMIS-SF-v1.0-Fatigue-13a>]
41. Schjolberg TK, Dodd M, Henriksen N, Asplund K, Cvancarova Smastuen M, Rustoen T. Effects of an educational intervention for managing fatigue in women with early stage breast cancer. *European journal of oncology nursing : the official journal of European Oncology Nursing Society.* 2014;18(3):286–94. <https://doi.org/10.1016/j.ejon.2014.01.008>.
42. COSMIN Study Design checklist for Patient-reported outcome measurement instruments [https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf]

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

