Fatigue in Parkinson's disease: Italian validation of the Parkinson Fatigue Scale and the Fatigue Severity Scale using a Rasch analysis approach

M. Siciliano[a,b], C. Chiorri[c], R. De Micco[a], A. Russo[a] , G. Tedeschi[a], L. Trojano[b,*], A. Tessitore[a]

[a]Department of Advanced Medical and Surgical Sciences - MRI Research Center SUN-FISM, University of Campania "Luigi Vanvitelli", Piazza Miraglia 2, 80138, Naples (IT);

[b]Department of Psychology, University of Campania "Luigi Vanvitelli", Viale Ellittico 31, 81100, Caserta (IT);

[c]Department of Educational Sciences, University of Genova, Genova, Italy.

*Corresponding author: Prof. Luigi Trojano, Department of Psychology, University of Campania "Luigi Vanvitelli", Viale Ellittico 31, 81100, Caserta (IT); e-mail: luigi.trojano@unicampania.it; Tel +39 0823274784; Fax: +39 0823274774

Keywords: fatigue; Rasch analysis; Fatigue Severity Scale; Parkinson Fatigue Scale; Parkinson's disease, Italian version.

Abstract

Introduction: The Fatigue Severity Scale (FSS-9) and the Parkinson Fatigue Scale (PFS-16) are commonly used for assessing fatigue in Parkinson's disease (PD). Here we validated the Italian version of these scales, assessed their psychometric properties by Rasch analysis, and computed their optimal cut-off scores using clinical diagnosis of PD-related fatigue as the gold standard.

Methods: PD patients (n= 167) completed the Italian versions of FSS-9 and PFS-16. Each item of PFS-16 was scored both on a 5-point (PFS-16polytomous) and on a 2-point scale (PFS-16dichotomous).

Results: All scales showed an adequate overall Rasch model fit, high reliability, and good discriminant, convergent, and concurrent validity, but were less accurate in measuring very high and very low fatigue levels. No evidence of differential item functioning with respect to age, sex, and severity of parkinsonian symptoms was found. Some items of FSS-9 (item 1), PFS-16polytomous (items 1 and 13), and PFS-16dichotomous (items 1, 8, and 13) showed misfit, possibly due to their content concerning sleep and motivation disorders. When FSS-9 and PFS-16polytomous' responses were rescored on a 3-point scale, the discriminability across response categories improved. The optimal cut-off score in detecting clinically-diagnosed fatigue (observed in 20% of the sample) was 3.09 for PFS-16polytomous, 8.00 for PFS-16dichotomous, and 4.67 for FSS-9.

Conclusions: The Italian version of PFS-16 and FSS-9 showed sound psychometric properties and can be confidently used to quantify fatigue symptoms in PD, although clinical diagnosis of fatigue should rely on validated criteria. The PFS-16polytomous exhibited advantages with respect to PFS-16dichotomous.

Introduction

Fatigue has been recently defined as a significantly diminished energy level or an increased perception of effort that are disproportionate to attempted activities [1,2].

In Parkinson's disease (PD), fatigue is one of the most common and bothersome non-motor symptoms [1,3,4], which may manifest even during premotor stages of the disease [5], and negatively impacts patients' quality of life [6,7].

Specific diagnostic criteria for defining PD-related fatigue have been recently proposed to facilitate fatigue-related patient disability claims and medication coverage, and guide participant selection for clinical trials [3]. Prevalence estimates of fatigue in PD range from 15 to 78%, and this variability is mostly due to the different instruments used to measure it [8]. For this reason, the assessment of fatigue severity in clinical and research contexts should be performed by means of sound standardized tools. The Movement Disorders Society Task Force on Rating Scales [9] "recommended" the Fatigue Severity Scale (FSS-9; [10]), and the Parkinson Fatigue Scale (PFS-16; [11]).

The FSS-9 and PFS-16 have been validated in many languages [12-18], but not in Italian. Two studies [19,20] used modern test theory (Rasch analysis and item response models, as opposed to classical test theory, that mainly relies on factor analytic approaches), which allows accurate and distribution-free analysis of the psychometric properties of the scales, to compare PFS-16 and FSS-9 with other scales, but no published report systematically compared PFS-16 and FSS-9 with each other [19-21]. Likewise, no previous studies considered the clinical diagnosis of PD-related fatigue [3] as "gold standard" to estimate the optimal cut-off scores for PFS-16 and FSS-9 for screening fatigued from nonfatigued patients.

Herein, we recruited a large sample of patients with PD with or without fatigue identified according to standard diagnostic clinical criteria [3]. After having investigated the clinical features of this sample, we aimed to: i) validate in Italian and head-to-head compare the psychometric properties of

PFS-16 and FSS-9 using a Rasch analysis approach; ii) provide the optimal cut-off scores for PFS-16 and FSS-9 using the clinical diagnosis of PD-related fatigue as the gold standard.

Methods

Patients

We screened consecutive patients at the Movement Disorders Outpatient Clinic of the First Division of Neurology of the University of Campania "Luigi Vanvitelli" (Naples, Italy) from April 2015 to February 2018.

To be enrolled in the study, patients had to fulfil the Movement Disorder Society (MDS) clinical diagnostic criteria for PD [e-1]. Exclusion criteria were: 1) dementia associated with PD [e-2]; 2) history of cerebrovascular or major unstable medical diseases; 3) lifetime or current psychotic disorders and/or previous treatment with antipsychotic drugs; 4) current or recurrent major depressive episode. The last two criteria were ascertained using the clinician-rated format of the Mini International Neuropsychiatric Inventory [e-3].

Patients were assessed in the "ON" state. A small proportion of them (30%) was re-assessed after 3 months in order to investigate the test-retest reliability of the scores provided by the PFS-16 and the FSS-9.

All procedures were approved and supervised by the local Ethical Committee, in accordance with the Declaration of Helsinki. Informed consent was obtained from all participants.

Clinical diagnosis of fatigue and fatigue scales

For the clinical diagnosis of fatigue, two neurologists (RDM and AT) used a structured clinical interview based on diagnostic criteria for PD-related fatigue [3].

The PFS-16 [11] and FSS-9 [9,10,19] were described in Table 1.

Translation and cultural adaptation

As no Italian standardized version of the two scales, specifically validated for PD patients, was available, before starting the study, we translated and adapted into Italian both the PFS-16 (Supplementary Material 1) and the FSS-9 (Supplementary Material 2) following international guidelines [e-4].

Other measures

The severity of motor symptoms was rated using the motor section of the Unified Parkinson's Disease Rating Scale (UPDRS; [e-5]) and the Hoehn and Yahr staging system (HY; [e-6]). The total amount of dopaminergic medication was expressed as the levodopa equivalent daily dosage (LEDD), determined by previously reported methods [e-7]. To profile the study sample, and gather data for testing convergent, discriminant, and concurrent validity, PD-related non-motor symptoms were evaluated using the following tests: the Montreal Cognitive Assessment (MoCA) for global cognitive status [e-8], the Beck Depression Inventory (BDI; [e- 9]) for depressive symptoms, the Parkinson Anxiety Scale for anxiety (PAS; [e-10,e-11]), the Apathy Evaluation Scale for apathy (AES; [e-12,e-13]); the Parkinson's Disease Sleep Scale for sleep disorders [e-14], and the Epworth Sleepiness Scale (ESS; [e-15]) for daytime sleepiness.

Statistical analysis

Motor and non-motor features of clinically diagnosed fatigued compared to nonfatigued PD patients were explored.

Data from the fatigue scales were fitted to a Rasch-Andrich Rating Scale Model (RA-RSM, [e-16,e-17]) using the Winsteps 4.1.0 [e-18]. According to this model, the probability of a given respondent of endorsing a specific response to an item is a logistic function of the relative distance between the item location and the respondent location on a latent trait [e-19]. In order to test how well the observed data fitted the values predicted by the RA-RSM, and to evaluate the accuracy of fatigue scales in classifying clinically diagnosed fatigued patients, we assessed: 1) *unidimensionality*, 2)

local item independence, 3) *Rasch model fit*, 4) *differential item functioning*, 5) *person/item separation and test-retest reliability*, 6) *targeting*, 7) *rating scale functioning*, 8) *convergent* and *discriminant validity*, and 9) *concurrent validity*. Moreover, 10) the *optimal cut-off scores* that best differentiate clinically-diagnosed fatigued from nonfatigued patients were provided for PFS-16p, PFS-16d, and FSS-9.

A detailed explanation of statistical analyses was reported in Supplementary Material 3.

For Rasch analysis, a sample of at least 150 individuals was needed for obtaining stable item calibration or person measures within a ± 0.5 logit interval, with a confidence level of 99% [e-18].

Results

Motor and non-motor features of clinically-diagnosed fatigued PD patients

One hundred sixty-seven patients with PD were enrolled in the study (see Table 2 for descriptive statistics). Thirty-three (20%) were affected by clinically relevant fatigue according to standard criteria. Demographic characteristics, clinical features, and cognitive scores did not significantly differ between clinically-diagnosed fatigued and nonfatigued patients. In contrast, fatigued patients had more severe depressive (BDI), anxious (PAS), apathetic (AES), sleep disorder (PDSS), and sleepiness (ESS) symptoms in comparison to nonfatigued patients (Table 2).

Validation of the PFS-16 and FSS-9

1) *Unidimensionality*

A Parallel Analysis procedure, carried out using simulated polychoric correlation matrices, suggested that the optimal number of factors for PFS-16p, PFS-16d, and FSS-9 was 1. Moreover, in all scales the first-to-second eigenvalue ratio was larger than 3, the first factor accounted for more than 65% of variance, and the eigenvalues began to flatten out from the second factor (Supplementay Material 4). Last, the number of statistically significant *t*-tests (outside the ± 1.96

range) between subsets of items were less than 5% for all fatigue scales (i.e., 3% for FSS-9, 0.6% for PFS-16p, and 2% for PFS-16d). All these findings supported the unidimensionality of the scales.

2) *Local item independence*

For all fatigue scales, Yen's Q_3 statistic [e-20] between pairs of items was always ≤0.30, suggesting that the items were locally independent when the level on the latent trait was held constant, and that there was no additional latent trait in the measurement.

3) *Rasch model fit*

PFS-16p [log-likelihood $\chi^2(5155)= 5127.14$, $p= 0.62$], PFS-16d [log-likelihood $\chi^2(1768)= 1682.11$, $p= 0.94$], and FSS-9 [log-likelihood $\chi^2(3491)= 3439.43$, $p= 0.73$] had an adequate overall fit to the Rasch model.

At the item-level, the items that showed underfit at a statistically significant level were Item 1 of PFS-16p and PFS-16d ("I have to rest during the day"), Item 8 of PFS-16d ("I have a feeling of 'heaviness'"), and Item 1 of FSS-9 ("My motivation is lower when I am fatigued"). Item 13 of PFS-16p and PFS-16d ("Fatigue makes it difficult for me to cope with everyday activities") demonstrated overfit at a statistically significant level (Table 3). After removing the items with the least acceptable goodness of fit, the overall fit of the scales did not improve (Supplementary Material 5), hence we decided to retain all the items.

Item hierarchy for PFS-16p, PFS-16d, and FSS-9 was shown in Table 3.

Person misfit occurred in 12 patients (7.18%) for PFS-16p, 3 patients (1.79%) for PFS-16d, and 14 patients (8.38%) for FSS-9. Their removal, however, did not produce noticeable changes in the fit to the Rasch model at the scale level, therefore no patient was removed from the overall sample.

4) *Differential Item Functioning (DIF)*

All scales were free from DIF for age, sex, and severity of parkinsonian symptoms (Supplementary Material 6 and 7).

5) *Person/item separation and test-retest reliability*

All scales showed good person- and item-separation reliability and good temporal stability of fatigue scores (test-rest reliability) (Supplementary Material 8).

6) *Targeting*

The mean sample location of PFS-16p and FSS-9 well approximated the mean item location (near to 0), but the mean sample location of PFS-16d was far (1 logit or more) from the mean item location (Supplementary Material 9).

The person-item histograms revealed that all scales covered a relatively narrow range of the latent trait, as none of their items could identify patients expressing extreme levels of fatigue (i.e., very high or low). In other terms, no items were agreed or disagreed with exclusively by patients with extreme levels of fatigue (Supplementary Material 10).

7) *Rating scale functioning*

The rating scales of all scales showed (Supplementary Material 11): i) at least 10 observations for each response category; ii) an irregular distribution of the observations among response categories (except PFS-16d); iii) an average rating scale category measures advancing monotonically; iv) OUTFIT MnSq for each response category less than 2.00 logit; iv) disordered transition threshold points between categories (except PFS-16d) (Supplementary Material 12). As for this last issue, the overall sample demonstrated a general inability to discriminate the mid response category area (response options ranging from 2: 'Disagree' up to 4: 'Agree' for PFS-16p; response options ranging from 2: 'Mainly Disagree' up to 6: 'Mainly Agree' for FSS-9). Based on this evidence and considering the criteria reported above, we systematically rescored rating scales as follows: from 5

to 3 levels for PFS-16p by combining 2 with 3 and 4 with 5 (rescored rating scale:12223); from 7 to 3 levels for FSS-9 by combining 1 with 2, 3 with 4 and 5, and 6 with 7 (rescored rating scale: 1122233). This revised scoring system improved discriminability among response categories (Supplementary Material 12).

8) *Convergent* and *discriminant validity*

All scales were very strongly correlated among them, negligibly or weakly correlated with BDI, PAS, AES, PDSS and ESS (Supplementary Material 13).

9) *Concurrent validity*

Twenty percent of the sample (33 of 167 PD patients) met the criteria for PD-related fatigue [3]. A moderate degree of agreement in differentiating fatigued from nonfatigued patients was found between diagnostic criteria proposed by Kluger et al. [3] and cut-off scores available for PFS-16 [11] and FSS-9 [10,18] (Supplementary Table 14).

10) *Optimal cut-off scores*

For each fatigue scale, the optimal cut-off scores were shown in Table 4, while the details of the ROC analyses were reported in Supplementary Material 15.

Discussion

The present study provided the first Italian validation and Rasch-based head-to-head comparison of PFS-16 and FSS-9, and determined the optimal cut-off scores for these scales for screening PD-related fatigue [3].

Overall, our study supported the psychometric robustness of PFS-16 and FSS-9 in PD, although no convincing empirical support for the adequacy of the dichotomous scoring method applied to PFS-16 (i.e., PFS-16d) was found.

The unidimensional structure of the latent trait underlying all scales was supported by the results (e.g., Parallel Analysis, local item independency), replicating previous studies [19,20].

Although, at the overall level, the scales did not depart from Rasch model expectations, some misfitting items were detected. From a substantive point of view, the underfit items may suggest that daytime sleepiness (Item 1 of PFS-16p and PFS-16d; "I have to rest during the day"), and fatigue-related lack of motivation (Item 1 of FSS-9; "My motivation is lower when I am fatigued") did not fully conform to the latent trait expressed by the remaining items of the respective scales (roughly oriented towards the impact of fatigue on activities of daily living). This result is consistent with previous evidence showing that although sleep disorders, lack of motivation (apathy), and fatigue are frequent and co-occur in PD [8,22-26], they should be considered as distinct symptoms [19,24]. Item 8 of PFS-16 ("I have a feeling of 'heaviness'") also showed a tendency to underfit when the dichotomous scoring method was used. Conversely, Item 13 of PFS-16p and PFS-16d ("Fatigue makes it difficult for me to cope with everyday activities") showed overfitting problems, i.e. did not add information to the scale, fostering the possibility of future item reduction [12]. However, when the misfitting items were removed, the overall fit of PFS-16p, PFS-16d, and FSS-9 did not improve, suggesting that these items were not degrading, although being poorly productive for the overall scale fit [e-18]. For this reason, we chose a conservative approach and retained the original version of the fatigue scales for all statistical analyses.

The item hierarchy of the fatigue scales showed that Item 7 of PFS-16p and PFS-16d ("Because of fatigue it takes me longer to get things done") and Item 4 of FSS-9 ("Fatigue interferes with my physical functioning") were those most easily agreed with, while Item 6 of PFS-16p and PFS-16d ("Fatigue makes me reluctant to socialise") and Item 7 of FSS-9 ("Fatigue interferes with carrying out certain duties and responsibilities") were the most difficult to agree with. Although this issue has not been addressed before in PD, similar results have been reported for FSS-9 in Multiple Sclerosis [27,28]. This evidence supports the idea that the hierarchical structure at least of the FSS-

9 holds across diverse neurological diseases, and can be used to compare distressing fatigue across neurological populations.

The absence of DIF supports the Hagell et al. [19] and Nilsson et al. 's results [20], sustaining that PFS-16 and FSS-9 work in the same way, independent of sex, age, and disease severity in PD. However, further studies are necessary to evaluate measurement invariance across other subgroups, such as language/cultural groups [19,28].

All scales showed high and comparable values of reliability (separation indices), meaning high reproducibility of item and person measure location [19,20]. In other words, these measures of fatigue are highly reliable. In the same vein, and similarly to previous findings [11,14,29], we found a good stability over time of the fatigue measures provided by PFS-16 or FSS-9, which support their usefulness for clinical and research longitudinal studies.

Rasch analysis also suggested that PFS-16p and FSS-9 were well-targeted fatigue scales in PD, although no item of these scales was able to identify patients affected by extreme fatigue levels (i.e., very high or low fatigue levels). This result indicates that the scales might not be enough sensitive to extreme expressions of the latent trait and is consistent with the failure of either scale to reach high levels of sensitivity and specificity (i.e., >90) in ROC analyses. In turn, this suggests that the fulfillment of diagnostic criteria (as those proposed by[3]) remains necessary when using either scale for avoiding misdiagnosis of fatigue.

The mean negative sample location associated to PFS-16d (compared with the mean value of 0 logits routinely assigned for items) suggested that this version is at risk of underestimating fatigue in PD. This is not surprising as the dichotomization of response categories could lead to loss of information and poor ability to detect individual differences [20].

The analysis of rating scale functioning suggested that PD patients did not discern among the mid response categories of PFS-16p and FSS-9. On this basis, the collapsing procedure suggested a three-point scoring method as the optimal solution for classifying original responses categories. This issue has arisen also in other neurological populations [28], supporting the hypothesis that, in

general, patients may have difficulty to discriminate among more than three response categories. Future studies should consider this issue when developing revised versions of these scales.

PFS-16 and FSS-9 were strongly correlated with each other (convergent validity), negligibly to weakly correlated with measures of other non-motor symptoms (discriminant validity), and moderately associated with diagnostic criteria (concurrent validity), indicating that they operationalize the same underlying construct and are valid measures of fatigue [12].

It is worth noting that the present study showed that patients with PD-related fatigue as defined by recent diagnostic criteria [3] did not show peculiar motor symptoms with respect to patients not affected by fatigue, but they did show higher severity of several non-motor symptoms (depression, anxiety, apathy, sleep disorders, and sleepiness). These findings are consistent with the hypothesis that fatigue together with other non-motor symptoms in PD may result from disruption of nondopaminergic pathways [30] and reflect serotonergic dysfunction in basal ganglia and limbic circuits [8].

Some limitations of this study need to be pointed out. These include the monocentric design and the relatively small sample size. Moreover, the lack of a control group of healthy participants did not allow the comparison of the fatigue severity between patients with PD and the general population. Finally, we did not include patients who underwent Deep Brain Stimulation or continuous intrajejunal infusion of L-dopa-carbidopa limiting the generalizability of our results to advanced stages of disease.

Nonetheless, the present study supported the reliability and validity of PFS-16 and FSS-9 (here in their Italian version) in PD, thus promoting their use in cross-sectional and longitudinal clinical studies and fostering cross-cultural studies for a deeper understanding of this distressing, common and underestimated non-motor symptom. Importantly, and in agreement with earlier observation [19,20], when comparing the two versions of PFS-16, the polytomous one exhibited higher measurement precision than the dichotomous version, but a 3-point scoring system might improve

discriminability among response categories for both scales and ease the completion of the scale by patients.

Conflicts of interest: none

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' role:

Mattia Siciliano contributed to the conception and design of the study, the acquisition, analysis and interpretation of data, drafted and revised critically the article for important intellectual content.

Carlo Chiorri contributed to the conception and design of the study, the acquisition, analysis and interpretation of data, drafted and revised critically the article for important intellectual content.

Rosa De Micco contributed to the design of the study, the acquisition, analysis and interpretation of data, drafted and revised critically the article for important intellectual content.

Antonio Russo, Antonio De Mase and Federica Garramone contributed to the conception and design of the study, the acquisition and interpretation of data.

Gioacchino Tedeschi contributed to the design of the study, the acquisition, analysis and interpretation of data, drafted and revised critically the article for important intellectual content.

Luigi Trojano contributed to the conception and design of the study, the acquisition, analysis and interpretation of data, drafted and revised critically the article for important intellectual content.

Alessandro Tessitore contributed to the conception and design of the study, the acquisition, analysis and interpretation of data, drafted and revised critically the article for important intellectual content.

All authors approved the final version of the article.

References

[1] B.M. Kluger, L.B. Krupp, R.M. Enoka, Fatigue and fatigability in neurologic illnesses: proposal for a unified taxonomy, Neurology 80 (2013) 409-416.

[2] J.H. Friedman, J.C. Beck, K.L. Chou, G. Clark, C.P. Fagundes, C.G. Goetz, K. Herlofson, B. Kluger, L.B. Krupp, A.E. Lang, J.S. Lou, L. Marsh, A. Newbould, D. Weintraub, Fatigue in Parkinson's disease: report from a mutidisciplinary symposium, N.P.J. Parkinsons Dis. (2016) 15025.

[3] B.M. Kluger, K. Herlofson, K.L. Chou, J.S. Lou, C.G. Goetz, A.E. Lang, D. Weintraub, J. Friedman, Parkinson's disease-related fatigue: A case definition and recommendations for clinical research, Mov. Disord. 31 (2016) 625-631.

[4] K. Herlofson, B.M. Kluger, Fatigue in Parkinson's disease, J. Neurol. Sci. 15 (2017) 38-41.

[5] C. Pont-Sunyer, A. Hotter, C. Gaig, K. Seppi, Y. Compta, R. Katzenschlager, N. Mas, D. Hofeneder, T. Brücke, A. Bayés, K. Wenzel, J. Infante, H. Zach, W. Pirker, I.J. Posada, R. . Álvarez, L. Ispierto, O. De Fàbregues, A. Callén, A. Palasí, M. Aguilar, M.J. Martí, F. Valldeoriola, M. Salamero, W. Poewe, E. Tolosa, The onset of nonmotor symptoms in Parkinson's disease (the ONSET PD study), Mov. Disord. 30 (2015) 229-237.

[6] P. Barone, A. Antonini, C. Colosimo, R. Marconi, L. Morgante, T.P. Avarello, E. Bottacchi, A. Cannas, G. Ceravolo, R. Ceravolo, G. Cicarelli, R.M. Gaglio, R.M. Giglia, F. Iemolo, M. Manfredi, G. Meco, A. Nicoletti, M. Pederzoli, A. Petrone, A. Pisani, F.E. Pontieri, R. Quatrale, S. Ramat, R. Scala, G. Volpe, S. Zappulla, A.R. Bentivoglio, F. Stocchi, G. Trianni, P.D. Dotto, PRIAMO study group, The PRIAMO study: A multicenter assessment of nonmotor symptoms and their impact on quality of life in Parkinson's disease, Mov. Disord. 24 (2009) 1641-1649

[7] F. Stocchi, G. Abbruzzese, R. Ceravolo, P. Cortelli, M. D'Amelio, M.F. De Pandis, G. Fabbrini, C. Pacchetti, G. Pezzoli, A. Tessitore, M. Canesi, C. Iannacone, M. Zappia, FORTE Study Group, Prevalence of fatigue in Parkinson disease and its clinical correlates, Neurology 15 (2014) 215-220.

[8] M. Siciliano, L. Trojano, G. Santangelo, R. De Micco, G. Tedeschi, A. Tessitore, Fatigue in Parkinson's disease: A systematic review and meta-analysis, Mov. Disord. 33 (2018) 1712-1723.

[9] J.H. Friedman, G. Alves, P. Hagell, J. Marinus, L. Marsh, P. Martinez-Martin, C.G. Goetz, W. Poewe, O. Rascol, C. Sampaio, G. Stebbins, A. Schrag, Fatigue rating scales critique and recommendations by the Movement Disorders Society task force on rating scales for Parkinson's disease, Mov. Disord. 25 (2010) 805-822.

[10] 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.

[11] R.G. Brown, A. Dittner, L. Findley, S.C. Wessely, The Parkinson fatigue scale, Parkinsonism Relat. Disord. 11 (2005) 49-55.

[12] J. Grace, A. Mendelsohn, J.H. Friedman, A comparison of fatigue measures in Parkinson's disease, Parkinsonism Relat. Disord. 13 (2007) 443-445.

[13] A. Kummer, P. Scalzo, F. Cardoso, A.L. Teixeira, Evaluation of fatigue in Parkinson's disease using the Brazilian version of Parkinson's Fatigue Scale, Acta Neurol. Scand. 123 (2011) 130-136.

[14] P. Hagell, T. Rosblom, S. Palhagen, A Swedish version of the 16-item Parkinson fatigue scale (PFS-16), Acta Neurol. Scand. 125 (2012) 288-292.

[15] R. Fu, S.S. Cui, J.J. Du, P. Huang, Y.C. He, C. Gao, X.G. Luo, S.D. Chen, Validation of the Parkinson Fatigue Scale in Chinese Parkinson's disease patients, Brain Behav. 7 (2017) e00712.

[16] E.A. Ozturk, B.G. Kocer, E. Umay, A. Cakci, Cross-cultural adaptation and psychometric evaluations of the Turkish version of Parkinson Fatigue Scale, Qual. Life Res. (2018). https://doi: 10.1007/s11136-018-1907-3.

[17] S.M. Fereshtehnejad, H. Hadizadeh, F. Farhadi, G.A. Shahidi, A. Delbari, J. Lökk, Reliability and validity of the persian version of the fatigue severity scale in idiopathic Parkinson's disease patients, Parkinsons Dis. (2013) 935429.

[18] A. Lerdal, A. Wahl, T. Rustøen, B.R. Hanestad, T. Moum, 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 33 (2005) 123-130.

[19] P. Hagell, A. Höglund, J. Reimer, B. Eriksson, I. Knutsson, H. Widner, D. Cella, Measuring fatigue in Parkinson's disease: a psychometric study of two brief generic fatigue questionnaires, J. Pain Symptom Manage. 32 (2006) 420-432.

[20] M.H. Nilsson, S. Bladh, P. Hagell, Fatigue in Parkinson's disease: measurement properties of a generic and a condition-specific rating scale, J. Pain Symptom. Manage. 46 (2013) 737-746.

[21] J. Hobart, S. Cano, Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods, Health. Technol. Assess. 13 (2009) 1-177.

[22] K. Herlofson, J.P. Larsen, Measuring fatigue in patients with Parkinson's disease-the Fatigue Severity Scale, Eur. J. Neurol. 9 (2002) 595-600.

[23] N. Sáez-Francàs, J. Hernández-Vara, M. Corominas Roso, J. Alegre Martín, M. Casas Brugué, The association of apathy with central fatigue perception in patients with Parkinson's disease, Behav. Neurosci. 127 (2013) 237-244.

[24] M. Skorvanek, I. Nagyova, J. Rosenberger, M. Krokavcova, R. Ghorbani Saeedian, J.W. Groothoff, Z. Gdovinova, J.P. van Dijk, Clinical determinants of primary and secondary fatigue in patients with Parkinson's disease, J. Neurol. 260 (2013) 1554-1561.

[25] M. Skorvanek, Z. Gdovinova, J. Rosenberger, R.G. Saeedian, I. Nagyova, J.W. Groothoff, J.P. van Dijk, The associations between fatigue, apathy, and depression in Parkinson's disease, Acta. Neurol. Scand. 131 (2015) 80-87.

[26] G.D. Cochrane, S. Rizvi, A.M. Abrantes, B. Crabtree, J. Cahill, J.H. Friedman, The association between fatigue and apathy in patients with either Parkinson's disease or multiple sclerosis, Parkinsonism Relat. Disord. 21 (2015) 1093-1095.

[27] R. Mills, C. Young, R. Nicholas, J. Pallant, A. Tennant, Rasch analysis of the Fatigue Severity Scale in multiple sclerosis, Mult. Scler. 15 (2009) 81-87.

[28] M. Ottonello, L. Pellicciari, A. Giordano, C. Foti, Rasch analysis of the Fatigue Severity Scale in Italian subjects with multiple sclerosis, J. Rehabil. Med. 48 (2016) 597-603.

[29] S. Valderramas, A.C. Feres, A. Melo, Reliability and validity study of a Brazilian-Portuguese version of the fatigue severity scale in Parkinson's disease patients, Arq. Neuropsiquiatr. 70 (2012) 497-500.

[30] N. Titova, M.A. Qamar, K.R. Chaudhuri, The Nonmotor Features of Parkinson's Disease, Int. Rev. Neurobiol. 132 (2017) 33-54.

Table 1. Brief description of Parkinson Fatigue Scale (PFS-16) and Fatigue Severity Scale (FSS-9).

Fatigue scale	Measure	Scoring Method	Total score	Available cut-off scores	Time frame
PFS-16polytomous (PFS-16p)	Developed for patients with Parkinson's Disease and designed to assess the physical aspects of fatigue and its impact on activities of daily functioning	From 1 (*strongly disagree*) to 5 (*strongly agree*); higher scores indicate more severe fatigue	1 – 5 points obtained by averaging the sum of all item scores	Total score ≥2.95 or ≥3.30 points[a]	Past 2 weeks
PFS-16dichotomous (PFS-16d)	As above	*Agree* and *strongly agree* are scored as 1, all other responses are scored as 0; higher scores indicate more severe fatigue	1 – 16 points obtained by summing all item scores	Total score ≥7 or ≥ 8 points[a]	Past 2 weeks
FSS-9	Developed for patients with Multiple Sclerosis and Systemic Lupus Erythematosus, and designed to assess physical, mental and social aspects of fatigue	From 1 (*strongly disagree*) to 7 (*strongly agree*); higher scores indicate more severe fatigue	1 – 7 points obtained by averaging the sum of all item scores	Total score ≥4[b] or 5[c] points	Past 2 weeks

	and their functional impact of fatigue.

Note. a, Brown et al., 2005 [11]; b, Krupp et al., 1989 [10]; c, Lerdal et al., 2005 [18].

Table 2. Characteristics of patients with and without diagnosis of fatigue, and summary of comparisons performed by one way-Analysis of Variance (ANOVA) for continuous variables and Pearson's chi-squared test (χ^2) for categorical variables.

Variable	Total sample (n= 167)	Fatigued (n= 33)	Nonfatigued (n= 134)	F/χ^2	p-value	Adj-p	Cohen's d/OR
Demographics							
Age	66.53±8.77	67.21±9.63	66.37±8.58	0.24	0.62	1.00	0.09
Education, years	9.90±4.38	10.36±4.80	9.79±4.28	0.45	0.50	1.00	0.12
Sex, male	99 (59.30%)	23 (69.70%)	76 (56.70%)	1.84	0.17	1.00	0.30
Clinical features							
Age at onset	61.93±9.45	62.41±9.94	61.81±9.36	0.09	0.75	1.00	0.06
Disease duration, years	4.44±2.94	4.84±3.67	4.35±2.75	0.68	0.41	1.00	0.15
UPDRS-III	26.68±10.39	30.97±11.90	25.69±9.79	6.28	0.01	0.23	0.48
Hoehn and Yahr stage	2.04±0.48	2.17±0.58	2.00±0.44	3.15	0.07	1.00	0.33
LEDD (mg/day)	415.94±289.13	498.73±356.63	395.55±267.61	3.42	0.06	1.00	0.32
Drug naïve PD patients	22 (13%)	10 (31.20%)	57 (44.20%)	1.76	0.18	1.00	0.17
Cognitive assessment							
MoCA	20.73±3.67	19.90±4.01	20.94±3.57	2.11	0.14	1.00	0.27
Behavioural measures							
BDI	7.25±7.07	12.79±8.32	5.88±6.02	29.57	<0.01	0.01	0.95
PAS	10.22±8.52	15.09±9.08	9.02±7.96	14.51	<0.01	0.01	0.71
AES	32.06±8.63	37.85±8.98	30.63±7.95	20.67	<0.01	0.01	0.85
PDSS	117.65±20.82	103.67±21.72	120.98±19.24	18.66	<0.01	0.01	0.84
ESS	5.90±4.53	8.15±4.79	5.33±4.30	10.80	<0.01	0.01	0.61
Fatigue measures							
PFS-16polytomous	1.77±1.06	3.77±0.95	2.17±1.01	68.59	<0.01	0.01	1.63
PFS-16dichotomous	5.17±4.91	10.45±4.73	3.87±4.00	66.46	<0.01	0.01	1.50
FSS-9	3.18±1.88	5.39±1.53	2.64±1.53	84.85	<0.01	0.01	2.15

Note. Adj-p represents p-value corrected for multiple comparisons using the Bonferroni procedure, OR, Odds Ratio; UPDRS, Unified Parkinson's Disease Rating Scale; LEDD, Levodopa Equivalent Daily Dose; MoCA, Montreal Cognitive Assessment; BDI, Beck Depression

Inventory; PAS, Parkinson Anxiety Scale; AES, Apathy Evaluation Scale; PDSS, Parkinson's disease sleep scale; ESS, Epworth Sleepiness Scale; PFS, Parkinson Fatigue Scale; FSS, Fatigue Severity Scale.

Table 3. Item hierarchy for PFS-16 Polytomous, PFS-16 Dichotomized, and FSS-7; items are sorted in descending order with items that are easier to agree with at the bottom and items that are harder to agree with at the up.

	Item	Measure	SE	INFIT		OUTFIT	
				MnSq	ZSTD	MnSq	ZSTD
	PFS-16 Polytomous						
6.	Fatigue makes me reluctant to socialise	0.65	0.10	1.25	1.70	0.88	-0.40
12.	I feel totally drained	0.52	0.09	1.13	1.00	1.32	1.30
5.	I feel completely exhausted	0.44	0.08	0.68	-2.90	0.99	0.00
16.	I get so tired I want to lie down wherever I am	0.42	0.08	0.98	-0.10	0.85	-0.60
14.	I feel tired even when I haven't done anything	0.32	0.08	1.07	0.60	1.06	0.30
2.	My life is restricted by fatigue	0.18	0.08	1.15	1.20	0.89	-0.40
4.	Fatigue is one of my three worst symptoms	0.14	0.08	0.98	-0.20	0.85	-0.70
11.	I lack energy for much of the time	0.07	0.08	0.88	-1.00	0.87	-0.60
10.	Everything I do is an effort	0.06	0.08	0.75	-2.40	0.70	-1.60
13.	**Fatigue makes it difficult for me to cope with everyday activities**	**0.03**	**0.07**	**0.55**	**-4.70**	**0.44**	**-3.40**
8.	I have a feeling of 'heaviness'	-0.20	0.08	1.16	1.40	1.30	1.50
3.	I get tired more quickly than other people I know	-0.42	0.08	0.96	-0.30	0.95	-0.20
15.	Because of fatigue I do less in my day than I would like	-0.42	0.08	0.81	-1.60	0.66	-1.90
1.	**I have to rest during the day**	**-0.47**	**0.10**	**2.41**	**8.50**	**3.75**	**8.80**
9.	If I wasn't so tired I could do more things	-0.62	0.08	0.89	-0.80	0.68	-1.80
7.	Because of fatigue it takes me longer to get things done	-0.73	0.08	0.69	-2.60	0.60	-2.50
	PFS-16 Dichotomized						
6.	Fatigue makes me reluctant to socialise	1.63	0.29	1.10	0.70	0.96	0.10
12.	I feel totally drained	1.55	0.28	1.03	0.20	1.10	0.40
5.	I feel completely exhausted	1.40	0.27	0.80	-1.30	0.51	-1.20
14.	I feel tired even when I haven't done anything	1.12	0.27	1.08	0.60	0.87	-0.20
16.	I get so tired I want to lie down wherever I am	0.86	0.25	0.86	-1.10	1.10	0.40
2.	My life is restricted by fatigue	0.39	0.24	1.04	0.40	0.82	-0.60
4.	Fatigue is one of my three worst symptoms	0.39	0.24	0.98	-0.10	0.91	-0.20
11.	I lack energy for much of the time	0.33	0.24	0.95	-0.30	0.82	-0.60
13.	**Fatigue makes it difficult for me to cope with everyday activities**	**0.22**	**0.24**	**0.60**	**-4.00**	**0.41**	**-2.90**

10.	Everything I do is an effort	0.16	0.24	0.81	-1.70	0.59	-1.80
8.	**I have a feeling of 'heaviness'**	**-0.48**	**0.26**	**1.26**	**2.20**	**1.55**	**2.20**
15.	Because of fatigue I do less in my day than I would like	-1.06	0.23	0.75	-2.40	0.57	-1.90
3.	I get tired more quickly than other people I know	-1.22	0.24	1.05	0.50	0.95	-0.10
1.	**I have to rest during the day**	**-1.38**	**0.35**	**2.21**	**7.90**	**3.88**	**6.00**
9.	If I wasn't so tired I could do more things	-1.89	0.24	0.65	-3.10	0.44	1.80
7.	Because of fatigue it takes me longer to get things done	-2.01	0.24	0.69	-2.70	0.49	-1.50
	FSS-9						
7.	Fatigue interferes with carrying out certain duties and responsibilities	0.27	0.06	1.02	0.20	0.88	-0.50
5.	Fatigue causes frequent problems for me	0.25	0.06	1.04	0.40	0.83	-0.80
9.	Fatigue interferes with my work, family, or social life	0.23	0.06	1.10	0.80	1.12	0.60
8.	Fatigue is among my three most disabling symptoms	0.04	0.06	1.09	0.80	0.90	-0.50
1.	**My motivation is lower when I am fatigued**	**-0.08**	**0.07**	**1.45**	**3.30**	**2.00**	**4.8**
3.	I am easily fatigued	-0.11	0.06	0.94	-0.40	0.92	-0.40
6.	My fatigue prevents sustained physical functioning	-0.15	0.06	0.91	-0.70	0.88	-0.70
2.	Exercise brings on my fatigue	-0.15	0.06	0.85	-1.30	0.91	-0.50
4.	Fatigue interferes with my physical functioning	-0.30	0.06	0.71	-2.60	0.65	-2.40

Note. Underfitted and overfitted items are in **bold** and in **highlighted bold**; PFS, Parkinson Fatigue Scale; FSS, Fatigue Severity Scale.

Table 4. Sensitivity, specificity, and Youden's index associated with Parkinson Fatigue Scale and Fatigue Severity Scale optimal cut-off scores.

Fatigue scale	Optimal cut-off	Sensitivity	Specificity	Youden's index
PFS-16p	3.09	0.81	0.78	0.60
PFS-16d	8.00	0.79	0.82	0.62
FSS-9	4.67	0.82	0.87	0.69

Note. PFS-16p, Parkinson Fatigue Scale-16polytomous; PFS-16d, Parkinson Fatigue Scale-16dichotomous; FSS-9, Fatigue Severity Scale-9.