Chapter 6

Reliability and structural validity of the Multidimensional Fatigue Inventory (MFI) in patients with idiopathic Parkinson's disease

Roy G. Elbers, Erwin E.H. van Wegen, John Verhoef, Gert Kwakkel

Parkinsonism and Related Disorders 2012; 18: 532-536

Abstract

Introduction The Multidimensional Fatigue Inventory (MFI) is commonly used in patients with Parkinson's disease (PD). However, most measurement properties have not been investigated in this population. The aim of this study was to investigate internal consistency, test-retest reliability, measurement error, structural validity, and floor and ceiling effects of the MFI in PD.

Methods Patients with PD (N = 153) completed the MFI at baseline and week 3 in a randomized clinical trial. Cronbach's α, intraclass correlation coefficient (ICC) and the smallest detectable change (SDC) were calculated. Bland and Altman analysis was performed. Principal Component Analysis (PCA) was used to explore structural validity. Floor and ceiling effects were investigated.

Results Cronbach's α for the MFI-total and subscales ranged from 0.74 (reduced motivation) to 0.92 (MFI-total). ICC's ranged from 0.65 (mental fatigue) to 0.81 (physical fatigue), SDC ranged from 6 points (physical fatigue and reduced motivation) to 24 points (MFI-total). Bland and Altman analysis showed no systematic differences between assessments. A floor effect was found for mental fatigue and ceiling effects for physical fatigue and reduced activity. A four-factor model was extracted, combining general fatigue and physical fatigue as one factor.

Conclusions The MFI is reliable and valid to assess fatigue in patients with PD. Clinicians and researchers interested in assessing specific aspects of fatigue should consider interpreting general fatigue and physical fatigue as one subscale measuring physical aspects of fatigue. To establish whether the MFI can detect meaningful changes, studies on anchor-based responsiveness and the minimal important change are needed in PD.

Introduction

Fatigue is common in patients with idiopathic Parkinson's disease (PD) and has a negative impact on health-related quality of life [1-3]. Prevalence rates reported in the literature range from 32% to 50% [4, 5]. One of the challenges in assessing fatigue is the lack of a widely accepted definition [6] and with that, differentiating its many dimensions [7, 8]. Fatigue usually refers to the difficulty initiating or sustaining voluntary activities [9]. Its multidimensionality is believed to result from a complex interplay between the underlying disease process, peripheral control systems (i.e. muscle fatigability), central control systems (i.e. subjective sense of fatigue) and environmental factors [9]. This complexity may be reflected in the large number of self-report questionnaires that are currently available to measure fatigue as either a multidimensional or a unidimensional assessment in patients with PD.

A commonly used instrument to assess the multidimensional aspects of fatigue in patients with PD is the Multidimensional Fatigue Inventory (MFI) [8]. The MFI is a self-report questionnaire that assesses five dimensions of fatigue (i.e. general fatigue, physical fatigue, reduced activity, mental fatigue and reduced motivation). Recently, the Movement Disorders Society Task Force on Rating Scales for Parkinson's disease suggested the MFI as a screening instrument for fatigue and recommended the MFI for the assessment of fatigue severity in patients with PD [10]. Unfortunately, these recommendations were largely based on evidence derived from studies in non-PD samples. In addition, measurement properties such as internal consistency, test-retest reliability, measurement error, structural validity and responsiveness have not been investigated in patients with PD [11].

The aim of the present study was therefore to investigate the internal consistency, test-retest reliability, measurement error, structural validity and floor and ceiling effects of the MFI in patients with PD.

Methods

Population and design

This study was part of a randomized clinical trial (the 'Rescue' trial (Rehabilitation in Parkinson's Disease: Strategies for Cueing) QLK6-CT-2001-00120) about the effects of cueing training on gait and gait-related activity in patients with PD [12]. In this study, 153 patients with PD were recruited from three European centers: Northumbria University, Newcastle upon Tyne (UK); Katholieke Universiteit Leuven, Leuven

(Belgium) and the VU University Medical Center, Amsterdam (The Netherlands). The study was approved by the ethics committee of each centre. All patients gave written informed consent. Patients were randomly allocated to an early or late intervention group by an independent person, not involved in the study. Further details about design and outcomes of the study have been published previously [12].

Subjects

Patients were recruited according to the following criteria: 1) age 18-80; 2) diagnosis of PD, defined by the UK Brain Bank Criteria [13]; 3) Hoehn and Yahr (H&Y) stage II-IV [14]; 4) stable drug usage and 5) mild to severe gait disturbance (score > 1 on the Unified Parkinson's Disease Rating Scale (UPDRS) item 29) [15]. Patients were excluded if they had: 1) undergone deep brain stimulation or other stereotactic neurosurgery; 2) cognitive impairment (Mini Mental State Examination (MMSE) < 24) [16]; 3) disorders interfering with participation in cueing training, including neurological (stroke, multiple sclerosis, brain tumor), cardiopulmonary (chronic obstructive disorders, angina pectoris) and orthopedic (osteoarthritis, rheumatoid arthritis and back pain) conditions; 4) unpredictable and long lasting off periods (score 1 on item 37 and score > 2 on item 39 of the UPDRS) [15] or 5) had participated in a physiotherapy program two months before starting the trial.

The Multidimensional Fatigue Inventory

The Multidimensional Fatigue Inventory (MFI) was originally developed and validated in the Dutch language in patients with cancer and patients with chronic fatigue syndrome [8] and was translated and validated in English in patients with cancer [17]. The MFI is a self-report questionnaire that assesses the impact of fatigue and comprises five dimensions (general fatigue, physical fatigue, reduced activity, mental fatigue and reduced motivation). Each subscale contains four items, with two items formulated in a positive (e.g. 'I feel fit') and two formulated in a negative direction (e.g. 'I feel fatigued'). The addressed recall period is 'lately'. All items are scored on a five-point Likert scale ranging from 1 ('yes, that is true') to 5 ('no, that is not true'). The negative formulated items must be recoded before adding up scores. The obtainable score within each subscale ranges from 4 (absence of fatigue) to 20 (maximum fatigue).

Procedure

Patients completed the MFI at baseline (t1) and week 3 (t2), during visits from a trained observer blinded to treatment allocation and not involved in data analysis.

Both assessments were performed in the patients' homes at the same time of the day in the on phase, approximately 1 hour after medication intake.

Statistical analysis

All data were analyzed with PASW statistical package (PASW Statistics version 18.0, IBM Corp., New York, USA). The mean scores of the MFI-total and subscales were investigated for statistically significant differences between Dutch and English speaking patients. Dependent on distribution by visual plot, parametric or non-parametric analyses were applied. A two-tailed significance level of 0.05 was used for all tests.

Reliability

Internal consistency is the degree of the interrelatedness among items, assuming the questionnaire to be unidimensional [18]. Cronbach's α was calculated for the total scale and for all subscales separately at t1 (N = 153) and considered adequate if it ranged from 0.70 to 0.95 [19].

Reliability was defined as the proportion of the total variance in the measurements due to 'true' differences between patients [18]. To ensure that patients were stable in the period between two assessments (i.e. no intervention was applied), only data from the late intervention group (N = 77) at t1 and t2 were used for analyses. For test-retest reliability, the intraclass correlation coefficient (ICC) was calculated. A two-way mixed effects model with an absolute agreement definition was used, assuming that included patients are a random selection of the population and the raters (i.e. items) are fixed. The ICC was considered adequate if ≥ 0.70 [19].

Measurement error, defined as the systematic and random error of a score that is not attributed to true changes in the construct to be measured [18], was determined with the Bland and Altman method [20]. The limits of agreement were calculated as the mean difference between two consecutive assessments ± 1.96 x standard deviation (SD) of this difference. In addition, the smallest detectable change (SDC) was calculated, based on the standard error of measurement (SEM). The SDC was calculated by $1.96 \times \sqrt{2} \times SEM$, where the SEM was computed by $SD \times \sqrt{(1\text{-}ICC)}$ [21].

Structural validity

Structural validity, defined as the degree to which scores of a questionnaire are an adequate reflection of the dimensionality of the construct to be measured [18], was investigated by means of exploratory factor analysis. A Principal Component Analysis (PCA) was used to extract factors. As correlations between factors were expected, the

obtained factors were rotated oblique using the direct oblimin procedure. A minimum eigenvalue of 1 was specified as extraction criterion and the criterion for factor loading was set at ≥ 0.40. The existing names of the MFI subscales [8] were used to label the extracted factors. Data from t1 were used for analysis (N = 153).

Floor and ceiling effects

A floor or ceiling effect was present if more than 15% of patients achieved the lowest or highest possible score on a questionnaire [19]. Data from t1 were used for analysis (N = 153).

Results

Table 6.1 presents the characteristics of the total sample (N = 153) and of the late intervention group subsample (N = 77). One hundred and five patients completed the

Table 6.1 Patient characteristics at baseline

	Total sample (N = 153)[a]	**Late intervention (N = 77)[a]**
Demography		
Male/female[b]	88/65	38/39
Age (years)	67.06 (7.54)	67.38 (8.11)
Language Dutch/English[b]	105/48	53/24
PD characteristics		
Disease duration (years)	8.25 (5.09)	8.63 (5.55)
H&Y (on)	2.78 (0.60)	2.82 (0.64)
H&Y II/III/IV (on)[b]	71/64/18	33/33/11
Clinical data		
UPDRS-total (on)	56.03 (16.01)	55.29 (15.71)
UPDRS I (on)	3.30 (1.72)	3.08 (1.71)
UPDRS II (on)	16.42 (6.03)	16.36 (5.95)
UPDRS III (on)	33.05 (11.28)	32.81 (11.06)
UPDRS IV (on)	3.34 (3.26)	3.22 (3.39)
MFI-total	62.74 (17.94)	62.69 (19.23)
MFI general fatigue	13.83 (4.30)	14.19 (4.34)
MFI physical fatigue	13.93 (4.51)	14.23 (4.80)
MFI reduced activity	13.45 (4.98)	13.00 (5.36)
MFI mental fatigue	10.36 (4.68)	10.38 (4.76)
MFI reduced motivation	11.16 (4.30)	10.88 (4.69)
HADS anxiety	6.90 (3.91)	6.82 (4.00)
HADS depression	7.20 (3.50)	7.09 (3.78)

[a]Expressed as mean (SD); [b]Expressed as number of patients

Dutch language version and 48 patients completed the English language version of the MFI. No statistically significant differences in fatigue scores were found between Dutch and English speaking patients. The mean level of fatigue (MFI-total) was 62.74 (SD = 17.94). Most patients had mild-to-moderate disease severity as 46% (N = 71) of patients were in H&Y stage II, 42% (N = 64) in stage III, and 12% (N = 18) in stage IV.

There were no item responses missing and all data were normally distributed by visual plot.

Reliability

Table 6.2 presents the results for internal consistency, test-retest reliability and measurement error. The MFI-total and all subscales showed adequate internal consistency reflected by a Cronbach's α ranging from 0.74 (reduced motivation) to 0.92 (MFI-total).

Test-retest reliability was adequate for the MFI-total (ICC = 0.80, 95% confidence interval (CI) = 0.70 to 0.87) and most subscales (ICC ranged from 0.73 (95% CI = 0.61 to 0.82) (general fatigue) to 0.81 (95% CI = 0.71 to 0.87) (physical fatigue). Test-retest reliability was not adequate for the mental fatigue dimension (ICC = 0.65, 95% CI = 0.50 to 0.76).

Measurement error, expressed by the SDC, was 24 points for the MFI-total and ranged from 6 (physical fatigue and reduced motivation) to 8 points (mental fatigue and

Table 6.2 Reliability and measurement error

	Reliability		**Measurement error**		
	Internal consistency[a] Cronbach's α	**Test-retest reliability[b] ICC (95% CI)**	**LOA[b]**	**SDC[b]**	**SEM[b]**
MFI-total	0.92	0.80 (0.70 to 0.87)	-22.21 to 24.34	23.24	8.38
MFI general fatigue	0.79	0.73 (0.61 to 0.82)	-6.13 to 6.23	6.20	2.24
MFI physical fatigue	0.83	0.81 (0.71 to 0.87)	-5.87 to 5.71	5.73	2.07
MFI reduced activity	0.88	0.74 (0.62 to 0.83)	-6.79 to 7.96	7.42	2.68
MFI mental fatigue	0.86	0.65 (0.50 to 0.76)	-6.95 to 7.91	7.41	2.67
MFI reduced motivation	0.74	0.79 (0.69 to 0.86)	-5.63 to 5.68	5.61	2.03

[a]N = 153; [b]N = 77

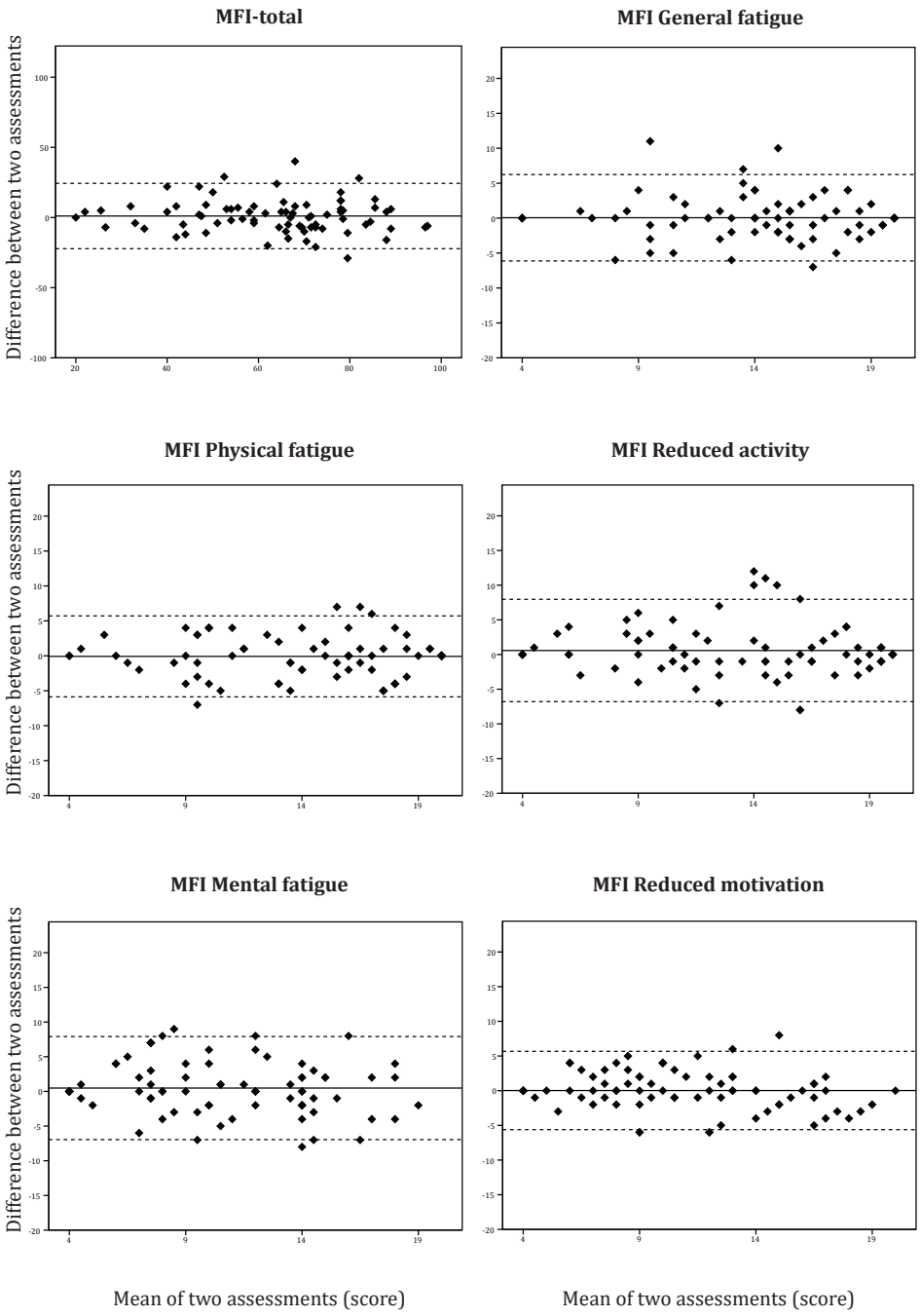

Figure 6.1 Graphic representation according to the Bland and Altman method for the MFI-total and all subscales. Bold lines represent the mean difference score, dotted lines represent the limits of agreement (N = 77).

reduced activity) for all subscales. Figure 6.1 presents the Bland and Altman plots for the MFI-total and all subscales. No systematic differences were found in the late intervention group between the first (t1) and second (t2) administration.

Structural validity

The results of the PCA are presented in Table 6.3. Four factors were extracted. The first factor was interpreted as a combination of the general fatigue and physical fatigue dimensions and the other three factors as the mental fatigue, reduced motivation and reduced activity dimensions.

All 20 items had an unique loading of ≥ 0.40 on one of four factors in the pattern matrix. Three items loaded on other factors compared to the original MFI subscales. Item 3 ('I feel very active') loaded on factor 3 (reduced motivation) instead of the reduced activity subscale. Item 8 ('Physically I can take on a lot') and item 9 ('I dread having to do things') loaded on factor 4 (reduced activity) instead of the physical fatigue and reduced motivation subscale respectively. The structure matrix shows a more complex model as 13 items loaded ≥ 0.40 across different factors. The maximum loading for most items was consistent between the pattern- and the structure matrix. Item 3 ('I feel very active') loaded in the pattern matrix on factor 3 (reduced motivation) and had its maximum loading in the structure matrix on factor 4 (reduced activity).

The factor correlation matrix shows moderate correlations between factor 1 (general fatigue/physical fatigue) and factor 2 (mental fatigue) (r = 0.35), between factor 1 (general fatigue/physical fatigue) and factor 4 (reduced activity) (r = 0.44), and between factor 3 (reduced motivation) and factor 4 (reduced activity) (r = 0.38).

Floor and ceiling effects

No floor or ceiling effects were found for the MFI-total and the general fatigue and reduced motivation subscales. The mental fatigue subscale showed a floor effect (18.30% of patients achieved the lowest possible score). Ceiling effects were found for the physical fatigue and reduced activities subscales as respectively 16.30% and 15.70% of patients achieved the highest possible score.

Table 6.3 Principal component analysis, loadings in pattern- and structure matrix[a,b] and factor correlation matrix (N = 153)

	Factor 1 General/Physical fatigue	Factor 2 Mental fatigue	Factor 3 Reduced motivation	Factor 4 Reduced activity
MFI general fatigue				
I feel fit (item 1)	**0.78 (0.82)**	-0.11 (0.23)	0.28 **(0.42)**	0.07 **(0.48)**
I feel tired (item 5)	**0.68 (0.73)**	0.14 (0.35)	-0.32 (-0.13)	0.12 (0.35)
I feel rested (item 12)	**0.71 (0.67)**	0.20 (0.36)	-0.18 (-0.08)	-0.17 (0.13)
I tired easily (item 16)	**0.73 (0.79)**	0.03 (0.30)	-0.12 (0.07)	0.14 **(0.43)**
MFI physical fatigue				
Physically I feel I am in an excellent condition (item 20)	**0.71 (0.76)**	0.04 (0.33)	0.28 **(0.40)**	-0.03 **(0.40)**
Physically I feel I am in a bad condition (item 14)	**0.73 (0.78)**	0.00 (0.30)	0.10 (0.27)	0.08 **(0.45)**
Physically I can take on a lot (item 8)	0.35 **(0.56)**	-0.11 (0.19)	0.27 **(0.48)**	**0.45 (0.67)**
Physically I feel only able to do a little (item 2)	**0.53 (0.71)**	-0.02 (0.30)	0.17 **(0.40)**	0.35 **(0.64)**
MFI reduced activity				
I think I do very little in a day (item 10)	0.01 **(0.42)**	0.07 (0.32)	-0.16 (0.20)	**0.94 (0.90)**
I think I do a lot in a day (item 6)	-0.09 (0.31)	-0.04 (0.21)	0.12 **(0.43)**	**0.87 (0.87)**
I get little done (item 17)	0.07 **(0.48)**	0.17 **(0.43)**	-0.01 (0.33)	**0.80 (0.88)**
I feel very active (item 3)	0.37 **(0.58)**	-0.07 (0.24)	**0.42 (0.61)**	0.36 **(0.66)**

MFI mental fatigue				
When I am doing something, I can keep my thoughts on it (item 7)	-0.01 (0.30)	**0.70 (0.74)**	-0.03 (0.16)	0.17 (0.36)
I can concentrate well (item 11)	0.01 (0.34)	**0.84 (0.87)**	0.08 (0.25)	0.04 (0.32)
My thoughts easily wander (item 19)	0.01 (0.30)	**0.81 (0.82)**	0.10 (0.24)	-0.02 (0.26)
It takes a lot of effort to concentrate on things (item 13)	0.08 (0.34)	**0.86 (0.86)**	0.02 (0.15)	-0.10 (0.20)
MFI reduced motivation				
I have a lot of plans (item 15)	-0.08 (0.17)	0.21 (0.35)	**0.71 (0.78)**	0.12 **(0.41)**
I feel like doing all sorts of nice things (item 4)	0.08 (0.23)	0.09 (0.24)	**0.83 (0.83)**	-0.08 (0.30)
I dread having to do things (item 9)	0.24 **(0.48)**	0.11 (0.33)	0.02 (0.25)	**0.45 (0.60)**
I don't feel like doing anything (item 18)	-0.02 (0.31)	0.36 **(0.52)**	**0.45 (0.62)**	0.29 **(0.55)**
Eigenvalue	5.90	4.40	3.44	5.80
Kaiser-Meyer-Olkin (Barlett's Test of Sphericity)	0.88 ($p < 0.001$)			
Factor correlation matrix[c]				
Factor 1 GF / PF	-	0.35	0.18	0.44
Factor 2 MF	0.35	-	0.18	0.29
Factor 3 RM	0.18	0.18	-	0.38
Factor 4 RA	0.44	0.29	0.38	-

[a]Loadings ≥ 0.40 are presented in bold; [b]The figures in parentheses indicate loadings in the structure matrix; [c]Expressed as Pearson r

Discussion

To our knowledge, this study is the first that investigated reliability, structural validity, and floor and ceiling effects of the MFI in patients with PD. The present study shows that the MFI-total and the general fatigue, physical fatigue, reduced activity and reduced motivation subscales have adequate internal consistency and test-retest reliability. However, it is unclear whether measurement error was adequate. Measurement error can be considered adequate if the SDC is smaller than the minimal important change (MIC) or if the MIC is outside the limits of agreement [19]. Unfortunately, no results are known on anchor-based values for the MIC of the MFI in patients with PD. Two studies [22, 23] that investigated the MIC in patients with systemic lupus erythematosus and rheumatoid arthritis found MIC values ranging from 12 [22] to 14 points [23] for the MFI-total. One study [24] investigated the MIC for all MFI subscales in patients with cancer and proposed a MIC of two points for all subscales [24]. This suggests that the measurement error found in our study may not be adequate and that the MFI may not be responsive to detect changes considered important by patients. However, studies that use anchor-based methods to investigate the MIC of the MFI in patients with PD are needed to confirm whether measurement error is adequate or not.

The SEM and SDC have been suggested as distribution-based MIC values. However, distribution-based methods are most applicable when the estimation of a clinically meaningful change does not rely on the estimate needing to be minimal [25]. The SDC provides supportive information whether change scores on the MFI exceed measurement error.

Principal component analysis failed to fully replicate the original five-factor model. However, the found four-factor model, combining most items of the general fatigue and physical fatigue dimensions in one factor, is in line with two other studies that investigated structural validity of the MFI in patients with cancer [26, 27]. In addition, in the original paper Smets and colleagues [8] found a four-factor model with a combined general fatigue and physical fatigue subscale that was equally acceptable as the postulated five-factor model. The developers decided to retain the five-factor model and concluded that if future research turns out that using both the general fatigue and physical fatigue subscale does not provide additional information these two subscales may be combined [8].

Combining the general fatigue and physical fatigue subscale poses a problem in calculating a global score for fatigue. The general fatigue subscale has been proposed as a short assessment for fatigue [8] and may be considered as a global score for

fatigue instead of the total summed score (MFI-total) of all 20 items [3, 28]. However, acknowledging the four-factor model, this global score (general fatigue) may reflect mainly physical aspects of fatigue and may therefore not be valid to represent all dimensions of fatigue measured with the MFI. Although PCA showed four distinct factors, the interrelatedness of these factors may allow the use of a total summed score of all 20 items to obtain a more valid global indication of fatigue.

The factor correlation matrix showed moderate correlations between most extracted factors, confirming the complex interrelatedness between different aspects of fatigue. Research, focused on physiological and clinical aspects contributing to peripheral and central fatigue [9] may be helpful to define concepts and dimensions of fatigue more clearly. As both fatigue and most clinical aspects contributing to fatigue fluctuate over time and show circadian rhythms, associations between these factors may be more accurately studied by using longitudinal studies with intensive, repeated measures in time [29].

Three items that loaded in the pattern matrix on other factors compared to the original MFI subscales suggested a misfit of items. However, taking the structure matrix and factor correlations between factor 1 (general fatigue/physical fatigue), factor 3 (reduced motivation) and factor 4 (reduced activity) into account, it is difficult to assign these items uniquely to one factor. With that, we decided not to consider these differences in factor-loadings as misfits.

The found floor effect for the mental fatigue subscale and the ceiling effects for the physical fatigue and reduced activity subscales should be considered when evaluating one of these aspects of fatigue. Furthermore, these results suggest that mental fatigue and physical fatigue are two different aspects of fatigue and further confirm previous findings that mental fatigue and physical fatigue are independent symptoms in PD [29, 30].

There are some study limitations that should be acknowledged. First, we combined results from the Dutch and English language version of the MFI. This may have resulted in cross-cultural differences within our sample. However, albeit in other patient populations, both language versions were previously validated [8, 17] and we found no statistically significant differences in fatigue scores between Dutch and English speaking patients in our sample. Second, we used PCA to investigate structural validity in a relatively small sample. Although the Kaiser-Meyer-Olkin statistic indicated reliable factors, this method provides preliminary data of the factorial structure of the MFI and these results should be confirmed in future studies using more robust statistical analyses such as Item Response Theory methods.

In conclusion, the present study shows that the MFI is a reliable and valid instrument to assess the multidimensional aspects of fatigue in patients with PD. Our results indicate that the found four-factor model, combining the general fatigue and physical fatigue dimensions into one subscale, more validly measures the different aspects of fatigue compared to the originally proposed five dimensions of the MFI. However this model as well as the construct of underlying dimensions of fatigue has to be confirmed in future studies. We recommend the use of the original version of the MFI to obtain a global indication of fatigue by calculating a total summed score of all 20 items. Clinicians and researchers interested in assessing specific aspects of fatigue should consider interpreting the general fatigue and physical fatigue dimensions as one subscale measuring physical aspects of fatigue. Furthermore, a comprehensive evaluation of fatigue should be accompanied by the assessment of clinically related factors such as mood and sleep. To establish whether the MFI can detect meaningful changes in clinical practice and research, studies on anchor-based responsiveness and the MIC are needed in patients with PD.

Acknowledgments

This research project was supported by a grant from the European Commission (QLK6-CT-2001-00120; Rehabilitation in Parkinson's Disease: Strategies for Cueing).

References

1. Havlikova E, Rosenberger J, Nagyova I, et al.: Impact of fatigue on quality of life in patients with Parkinson's disease. Eur J Neurol 2008; 15: 475-480.

2. Herlofson K, Larsen J: The influence of fatigue on health-related quality of life in patients with Parkinson's disease. Acta Neurol Scand 2003; 107: 1-6.

3. Martínez-Martín P, Catalan M, Benito-León J, et al.: Impact of fatigue in Parkinson's disease: the fatigue impact scale for daily use (D-FIS). Qual Life Res 2006; 15: 597-606.

4. Alves G, Wentzel-Larsen T, Larsen J: Is fatigue an independent and persistent symptom in patients with Parkinson disease? Neurology 2004; 63: 1908-1911.

5. Herlofson K, Larsen J: Measuring fatigue in patients with Parkinson's disease - the fatigue severity scale. Eur J Neurol 2002; 9: 595-600.

6. Lou J: Physical and mental fatigue in Parkinson's disease: epidemiology, pathophysiology and treatment. Drugs Aging 2009; 26: 195-208.

7. Friedman J, Brown R, Comella C, et al.: Fatigue in Parkinson's disease: a review. Mov Disord 2007; 22: 297-308.

8. Smets E, Garssen B, Bonke B, et al.: The multidimensional fatigue inventory (MFI) psychometric qualities of an instrument to assess fatigue. J Psychosom Res 1995; 39: 315-325.

9. Chaudhuri A, Behan P: Fatigue in neurological disorders. Lancet 2004; 363: 978-988.

10. Friedman J, Alves G, Hagell P, et al.: Fatigue rating scales critique and recommendations by the Movement Disorders Society task force on rating scales for Parkinson's disease. Mov Disord 2010; 25: 805-822.

11. Elbers R, Rietberg M, van Wegen E, et al.: Self-report fatigue questionnaires in multiple sclerosis, Parkinson's disease and stroke: a systematic review of measurement properties. Qual Life Res 2012; 21: 925-944.

12. Nieuwboer A, Kwakkel G, Rochester L, et al.: Cueing training in the home improves gait-related mobility in Parkinson's disease: the RESCUE trial. J Neurol Neurosurg Psychiatry 2007; 78: 134-140.

13. Hughes A, Daniel S, Kilford L, et al.: Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. J Neurol Neurosurg Psychiatry 1992; 55: 181-184.

14. Hoehn M, Yahr M: Parkinsonism: onset, progression and mortality. Neurology 1967; 17: 427-442.

15. Fahn S, Elton E: The unified Parkinson's disease rating scale, in Recent developments in Parkinson's disease, edited by Calne D. New Jersey, Macmillan Healthcare Information, 1987, pp. 153-163.

16. Folstein M, Folstein S, McHugh P: 'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975; 12: 189-198.

17. Smets E, Garssen B, Cull A, et al.: Application of the multidimensional fatigue inventory (MFI-20) in cancer patients receiving radiotherapy. Br J Cancer 1996; 73: 241-245.

18. Mokkink L, Terwee C, Patrick D, et al.: The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010; 63: 737-745.

19. Terwee C, Bot S, de Boer M, et al.: Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007; 60: 34-42.

20. Bland J, Altman D: Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1: 307-310.

21. De Vet H, Terwee C, Knol D, et al.: When to use agreement versus reliability measures. J CLin Epidemiol 2006; 59: 1033-1039.

22. Goligher E, Pouchot J, Brant R, et al.: Minimal clinically important difference for 7 measures of fatigue in patients with systemic lupus erythematosus. J Rheumatol 2008; 35: 635-642.

23. Pouchot J, Kherani R, Brant R, et al.: Determination of the minimal clinically important difference for seven fatigue measures in rheumatoid arthritis. J Clin Epidemiol 2008; 61: 705-713.

24. Purcell A, Fleming J, Bennett S, et al.: Determining the minimal clinically important difference criteria for the multidimensional fatigue inventory in a radiotherapy population. Support Care Cancer 2010; 18: 307-315.

25. Revicki D, Hays R, Cella D, et al.: Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol 2008; 61: 102-109.

26. Fillion L, Gélinas C, Simard S, et al.: Validation evidence for the French Canadian adaptation of the multidimensional fatigue inventory as a measure of cancer-related fatigue. Cancer Nurs 2003; 26: 143-154.

27. Gentile S, Delarozière J, Favre F, et al.: Validation of the French 'multidimensional fatigue inventory' (MFI 20). Eur J Cancer Care 2003; 12: 58-64.

28. Benito-León J, Martínez-Martín P, Frades B, et al.: Impact of fatigue in multiple sclerosis: the fatigue impact scale for daily use (D-FIS). Mult Scler 2007; 13: 645-651.

29. Elbers R, van Wegen E, Rochester L, et al.: Is impact of fatigue an independent factor associated with physical activity in patients with idiopathic Parkinson's disease? Mov Disord 2009; 24: 1512-1518.

30. Lou J, Kearns G, Oken B, et al.: Exacerbated physical fatigue and mental fatigue in Parkinson's disease. Mov Disord 2001; 16: 190-196.