Multiple Sclerosis and Related Disorders 65 (2022) 104012

Contents lists available at ScienceDirect

Multiple Sclerosis and Related Disorders

journal homepage: www.elsevier.com/locate/msard

Multimethodological validation of the modified fatigue impact scale in a Danish population of people with Multiple Sclerosis

Morten Riemenschneider [a,b,*], Philipp Trénel [a,c], Michael Nørgaard [a], Finn Boesen [a]

[a] The Danish MS Hospitals, Ry & Haslev, Denmark
[b] Exercise Biology, Department of Public Health, Aarhus University, Denmark
[c] Danish Technological Institute, Aarhus, Denmark

A B S T R A C T

Objective: To provide a multimethodological validation of MFIS in compliance with the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) guidelines using a Danish population of patients with Multiple Sclerosis (pwMS).
Methods: Factor analytic and multidimensional Rasch analytic methods were applied to investigate the structural validity of MFIS. We employed measurement and bifactor models to address divergence and convergence. McDonalds ω was used to assess reliability. Cross-cultural and nomological validity was assessed in relation to relevant external populations and constructs, respectively. Responsiveness was evaluated following one month of multi-disciplinary rehabilitation (MDR).
Results: Based on data from 424 Danish MS patients, the MFIS showed an acceptable fit to both factor analytic and multidimensional Rasch models. MFIS was found to be dominated by a general factor but with considerable substructure present. Sufficient divergence was found between cognitive and physical subdomains, but not the psychosocial subscale. Mc Donald's ω_t of 0.96 indicated good reliability of the scale, however, with low specific reliability of the psychosocial subscale. Homogeneity across Danish and European populations of pwMS supported cross-cultural validity. Explorative factor analysis of the nomological network of MFIS revealed expected convergence with cognitive and physical aspects. MFIS showed good responsiveness as indicated by moderate to large effect sizes following MDR.
Conclusion: The Danish version of the MFIS showed good reliability, a good structural, cross-cultural, and nomological validity an acceptable fit to a multidimensional Rasch analysis, and a good responsiveness. The psychosocial subscale of the MFIS however should be interpreted with caution.

1. Introduction

Multiple sclerosis (MS) is a chronic, progressive inflammatory and degenerative disease affecting the central nervous system, characterized by a complex pathogenesis and heterogenous symptoms across physical, cognitive, and neuropsychological domains (Filippi et al., 2018; Thrue et al., 2020). Fatigue is the most common symptom reported by patients with multiple sclerosis (MS) (Rommer et al., 2019), and it is known to have a detrimental impact on patients perception of health and quality of life (Green et al., 2017; Amato et al., 2001). The most commonly used assessment of subjective perceptions of fatigue is based on measurements through self-report questionnaires, i.e. health-related patient-reported outcomes (HR-PROs). For this purpose, the Multiple Sclerosis Council for Clinical Practice Guidelines derived the 21-item Modified Fatigue Impact Scale (MFIS), from the original 40-item Fatigue Impact Scale (FIS), and recommended it for clinical use and research (MSCfCP, 1998; Fisk et al., 1994). Specifically, the MFIS is a three-factor composite scale evaluating and quantifying the trait level of experienced fatigue impact on physical (9 items), cognitive (10 items) and psychosocial functioning (2 items). The MFIS is an acknowledged outcome measure in clinical trials, and despite being widely used in the Danish MS community (Taul-Madsen et al., 2020; Johansson et al., 2021; Oervik et al., 2017; Riemenschneider et al., 2021; Callesen et al., 2020), the Danish translation of the questionnaire has not yet been validated. Moreover, previous validation studies of other translations of the MFIS scale have reached different conclusions regarding the validity of the MFIS. Kos et al. found MFIS to be a valid scale in a multinational European study, albeit suggesting that the psychosocial subscale should be interpreted with caution as it showed weak homogeneity and added limited extra information (Kos et al., 2005). Another study, using Rasch analysis, concluded that the MFIS is not valid in its current form as it did not fit the unidimensional Rasch model used (Mills et al., 2010). The inherent multidimensionality of the MFIS scale however warrants a multidimensional Rasch analysis (Whitehead, 2009).

Proper validation of HR-PROs is vital as healthcare systems often base important decisions on diagnosis and treatment options on measurement of health status measurements. Generally, a measurement scale is considered methodologically valid, if it measures what it intends to measure with negligible bias, sufficient precision, and sufficient consistency across varying measurement conditions within a pre-

* Corresponding author at: The Danish MS Hospitals, Ry & Haslev, Denmark.
E-mail address: mori@ph.au.dk (M. Riemenschneider).

https://doi.org/10.1016/j.msard.2022.104012
Received 28 April 2022; Received in revised form 3 June 2022; Accepted 1 July 2022
Available online 2 July 2022

specified range (Squara et al., 2021a; 2021b). This also applies to psychometric scales. However, validation of psychometric scales is a conceptually and methodologically complex task, and as evidenced by the conflicting results on the validity of the MFIS, the choice of validation methodology is crucial for the precision and conclusion of the validation (Mokkink et al., 2010; Boateng et al., 2018). On this basis, a comprehensive standardized checklist has been developed to evaluate measurement properties of health status measurement instruments: the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) (Mokkink et al., 2010). According to this checklist, the most important measurement properties when evaluating a HR-PRO is the construct validity (including structural validity, hypotheses testing, and cross-cultural validity), the criterion validity, and the content validity. Moreover, the reliability and the responsiveness of the HR-PRO is also considered important for the quality of a measure (Mokkink et al., 2010).

On this basis, the aims of the present study were to (1) validate the Danish translation of the MFIS, by (2) making a comprehensive validation of the MFIS according to the COSMIN guidelines and thereby addressing the limitations of the previous conflicting conclusions on the validity of the MFIS. The latter will be done by (a) addressing the dimensionality of the construct (structural validation) as well as the convergent and discriminant validity of the three subscales within the MFIS construct, (b) investigating the reliability of the construct and its subscales, (c) investigating test- and item information using multidimensional Rasch analysis, (d) addressing group-invariance and performing a cross-cultural validation by comparison with published MFIS results for European populations (Kos et al., 2005), (e) addressing the nomological validity of MFIS by testing hypothesized relationships to relevant constructs and external criteria (criterion validity), and (f) investigating the responsiveness of the scale in patients following one month of multidisciplinary rehabilitation (MDR).

2. Methods

2.1. Study population

The study population for this validation study consisted of a total of 424 Danish MS patients. All patients were part of The Danish MS Hospitals Rehabilitation Study, and a detailed description of inclusion criteria, study design, data collection, and data management procedures can be found in the paper by Boesen et al. (2018) The study population consisted of two geographical subpopulations corresponding to the two Danish MS hospitals positioned in Haslev, Zealand ($n = 192$), and Ry, Jutland ($n = 232$). Data were collected at baseline, i.e., before any treatment received, with the exception of follow-up data (collected after one-month of MDR in a subset of the study population ($n = 182$)). Patients of this subpopulation had been assigned to the Fatigue main focus area of the MDR program of the Danish MS Hospitals based on an initial consultation prior to hospital admission indicating patient challenges and concerns in relation to fatigue.

2.2. Item translation

The 21 original MFIS items in English were translated to Danish, as well as back-translated to English by Sanofi-Aventis. Back-translation was used to evaluate the adequacy of the translation. Table 1 lists the 21 MFIS items and the Danish translation of each item.

2.3. Statistical analyses

All of the following validation methods were performed in R version 3.6.3 (2020–02–29) (R Core Team, 2020), especially using R-packages psych, (21) lavaan, (Rosseel, 2012) and mirt, (Chalmers, 2012) see supplementary material for more details.

Table 1
The 21 MFIS items, their subscale and Danish translation.

Nr	Subscale	Abbreviation	English	Danish
01	Cogn.	Alertness	01 I have been less alert.	01 Jeg har været mindre oplagt
02	Cogn.	Attention	02 I have had difficulty paying attention for long periods of time.	02 Jeg har haft problemer med at koncentrere mig i længere tid ad gangen
03	Cogn.	Clear.thinking	03 I have been unable to think clearly.	03 Jeg har ikke været i stand til at tænke klart
04	Phys.	Coordination	04 I have been clumsy and uncoordinated.	04 Jeg har været klodset og haft besvær med at koordinere mine bevægelser
05	Cogn.	Forgetfulness	05 I have been forgetful.	05 Jeg har været glemsom
06	Phys.	Pace.phys.act	06 I have had to pace myself in my physical activities.	06 Jeg har været nødt til at gøre en ekstra anstrengelse for at udføre fysiske aktiviteter
07	Phys.	Motiv.phys.effort	07 I have been less motivated to do anything that requires physical effort.	07 Jeg har haft mindre lyst til at foretage mig noget, der kræver fysisk anstrengelse
08	PsSoc.	Motiv.social.act	08 I have been less motivated to participate in social activities.	08 Jeg har haft mindre lyst til at deltage i sociale aktiviteter
09	PsSoc.	Outdoor.act	09 I have been limited in my ability to do things away from home.	09 Mine muligheder for at deltage i aktiviteter uden for hjemmet har været begrænsede
10	Phys.	Maintain.phys.effort	10 I have trouble maintaining physical effort for long periods.	10 Jeg har haft svært ved at udføre fysisk krævende aktiviteter i længere tid ad gangen
11	Cogn.	Making.decisions	11 I have had difficulty making decisions.	11 Jeg har haft besvær med at træffe beslutninger
12	Cogn.	Motiv.mental.effort	12 I have been less motivated to do anything that requires thinking.	12 Jeg har haft mindre lyst til at foretage mig noget, der krævede tankevirksomhed
13	Phys.	Muscle.weakness	13 My muscles have felt weak.	13 Mine muskler har været svage
14	Phys.	Phys.uncomfortable	14 I have been physically uncomfortable.	14 Jeg har haft fysisk ubehag
15	Cogn.	Finish.tasks	15 I have had trouble finishing tasks that require thinking.	15 Jeg har haft besvær med at fuldføre opgaver, der krævede tankevirksomhed
16	Cogn.	Organize.thoughts	16 I have had difficulty organizing my thoughts when doing things at home or at work.	16 Jeg har haft besvær med at samle tankerne, når jeg foretog mig noget hjemme eller på arbejde
17	Phys.	Complete.phys.tasks	17 I have been less able to complete tasks that require physical effort.	17 Jeg har haft besvær med at fuldføre opgaver, der krævede fysisk anstrengelse
18	Cogn.	Slow.thinking	18 My thinking has been slowed down.	18 Min tankevirksomhed er blevet mere langsom
19	Cogn.	Concentration	19 I have had trouble concentrating.	19 Jeg har haft problemer med at koncentrerer mig
20	Phys.	Limit.phys.act	20 I have limited my physical activities.	20 Jeg har begrænset mine fysiske aktiviteter
21	Phys.	Have.rest	21 I have needed to rest more often or for longer periods.	21 Jeg har haft behov for at hvile hyppigere og/eller i længere tid ad gangen

2.4. Data quality and scale acceptability

We assessed the quality and scale acceptability of the data by analyzing missingness of data, scale range coverage, and flooring and ceiling effects (see supplementary material).

2.5. Local independence, monotonicity, multivariate normality and tau-equivalence

The validity of a psychometric scale as well as its assessment depends on a set of assumptions and the degree of deviation from those assumptions. We used non-parametric Mokken scale analysis in order to address the assumptions of local independence and monotonicity of the latent scale (Dima, 2018). Local independence states that items are independent contingent on the latent scale, while monotonicity states that the latent construct is a monotone increasing scale. Both factor analytic and Rasch analytic models assume local independence and monotonicity. Factor analytic models are furthermore based on a normal-distributional assumption underlying the variance-covariance matrix of item responses, but estimators relaxing this assumption are available, for example the diagonally weighted least squares (DWLS) estimator. We used Mardia's test to test for multivariate normality among items on the total MFIS scale, as well as on the Cognitive and Physical subscales, but not the Psychosocial subscale due to its limited number of items. Reliability of a psychometric scale is often reported using Cronbach's α, but this reliability measure requires the assumption of tau-equivalence to be met. In factor analytic models, tau-equivalence poses a constraint on the variance-covariance matrix by restricting all covariances to be equal. Rasch models share the equivalent property of constraining all items to share the same slope on the latent construct. We evaluated the plausibility of tau-equivalence using confirmatory factor analysis (CFA) by comparing models with increasingly relaxed variance-covariance constraints: (1) all items share a common variance and all inter-item covariances are constrained to be equal (parallel items model), (2) unrestricted item variances but equal inter-item covariances (tau-equivalence model), and (3) unconstrained variance-covariance matrix (congeneric model).

2.6. Structural validity and reliability

We addressed the adequacy of the three-subscale model of MFIS consisting of the Cognitive, Physical, and Psychosocial subscales both in a factor analytic and multidimensional Rasch analytic context using the following two models: (1) a bi-factorial model and (2) a measurement model. The bi-factorial model hypothesizes that each item depends on three sources of variance: (1) a common or general factor, (2) a specific or subconstruct factor and (3) random measurement error. A bi-factor analysis (BFA) therefore reveals the degree of convergence of items towards a single underlying dimension. Furthermore, BFA allows quantification of the degree of convergence of items onto the general factor versus onto their specific subconstruct factor. In contrast, the multi-dimensional measurement model provides a different parametrization, with items depending only on its specific subscale factor and random measurement error, whereas items across subscale factors are related only through parameters describing between-factor correlations. The multi-dimensional measurement model is therefore well-suited to address the sufficiency of convergence and divergence of factors. Sufficient item-on-subscale convergence was evaluated as standardized factor loadings > 0.4. Sufficient divergence among subscales was evaluated as inter-factor correlations < 0.8 (Gana and Broc, 2019).

For comparison with previous studies, we also investigated uni-dimensional (1D) models, ignoring the three-subscale structure of MFIS (Whitehead, 2009), and two-dimensional (2D) models. In the 2D models, the Psychosocial subscale was combined with the Physical subscale into a single subcomponent following the findings of Kos et al. (2005). Finally, in the factor analytic context, we also investigated a two-layered structural equation model (3D2L), where the Psychosocial subscale is modelled to be secondary to (i.e. depending on) the Physical subscale, as the latter is more likely to be a primary source of fatigue and as physical and psychosocial fatigue factors have been shown to be associated (Kluger et al., 2013; Trojan et al., 2007).

In all factor analytic analyses in the current study, we employed the DWLS estimator and constrained the two Psychosocial items (items 8 and 9) to share the same variance to ensure identifiability. Due to the polytomous nature of the MFIS scale, Rasch analyses were carried out using the Rating Scale Model (RSM) and the Partial Credit Model (PCM). The PCM is a generalization of the RSM by including specific item-category parameters for each item, thereby allowing items to have differing category metrics, i.e. each item and item category has its own difficulty parameter.

From the BFA output we calculated Mc Donald's ω_h, the Explained Common Variance (ECV), and Mc Donald's ω_t. Mc Donald's ω_h gives the proportion of variance in observed scores accounted for by a single general factor and, hence, is indicative of unidimensionality if high (Mair, 2018). The ECV is the ratio of the general factor eigen value to the sum of all factor eigen values and hence indicates the degree of uni-dimensionality versus substructure. Mc Donald's ω_t, on the other hand, is the proportion of variance in scores due to all factors and, hence, expresses the expected reliability of the scale.

2.7. Fit indices

The following global fit indices were used in all analyses with abbreviations and adequacy thresholds in parentheses: Comparative Fit Index (CFI, good fit: $0.90 \leq \text{CFI} \leq 0.95$, very good fit: ≥ 0.95), Tucker-Lewis Index (TFI, good fit: $0.90 \leq \text{CFI} \leq 0.95$, very good fit: ≥ 0.95), Standardized Root Mean Squared Residuals (SRMSR, ≤ 0.08), and the Root Mean Squared Error of Approximation (RMSEA, ≤ 0.05, and with RMSEA 90% upper confidence limits 90%-UCL ≤ 0.1) (Gana and Broc, 2019). Adequacy thresholds for validation followed (Gana and Broc, 2019).

Group invariance and cross-cultural validity – A psychometric scale should be consistent across varying measurement conditions as for example gender, geography and cultural background. Here, we addressed group invariance among gender and among the two Danish MS hospitals located in Ry (Jutland) and Haslev (Sealand) using Andersen's Likelihood Ratio test of group invariance on the uni-dimensional PCM models for the Cognitive and Physical MFIS subscales. Furthermore, we addressed cross-cultural validity using the Tucker index of factor congruence comparing PCA loadings reported by Kos et al. (2005) for a sample across four different European countries (Belgium, Italy, Slovenia, Spain) with PCA loadings for our Danish sample population, with PC loadings calculated according to Kos et al. (2005). We also report the RV coefficient of overall matrix correlation.

Nomological validity – We investigated the nomological net of MFIS by performing an Exploratory Factor Analysis (EFA) including the six subscale scores of the Functional Assessment in Multiple Sclerosis scale (FAMS, ref.), Beck's Depression Inventory II (BDI) and physical performance as measured by the 6 min walking test (6MWT) along with the MFIS Cognitive, Physical and Psychosocial subscale scores.

2.8. Responsiveness

For a psychometric scale to be valuable, it must be able to capture hypothesized effects of relevant size in the population (Mokkink et al., 2010; Hobart et al., 2004). Here, we investigated the responsiveness of the MFIS scale and its subscales as Cohen's effect size d for changes from baseline to discharge for patients in the thinking/fatigue-targeted patient subset ($n = 181$) following one month of fatigue-targeted MDR. When assessing the responsiveness, the intervention must be assumed to be effective, why this fatigue-specific subgroup of patients was chosen for this analysis. See Boesen et al. for more details on the MDR intervention (Boesen et al., 2018). Effect sizes will be interpreted according to the recent empirically based guidelines for rehabilitation research, with effect sizes of 0.14–0.30 being considered small, effect sizes of 0.31–0.60 moderate, and effect sizes $\geq$ 0.61 being large (Kinney et al., 2020).

3. Results

Characteristics of the Danish study population are shown in Table 2. The thinking/fatigue-targeted patient subset used for the responsiveness evaluation differed from the main study population by higher MFIS at baseline, slightly higher proportion of younger patients, shorter disease duration, higher proportion of relapsing remitting MS and patients with EDSS scores $<$ 3.5, and longer 6MWT distances, see Table 2.

3.1. Data quality and scale acceptability

Ten out of the 424 patients (2.4%) did not return the MFIS questionnaire at baseline. Scale coverages for MFIS and its subscales were high, ranging from 88.1% (total MFIS) to 100%, and floor and ceiling effects were found to be small (<8% of the sample). See Table S1 in the supplementary material for full results on data quality and scale acceptability.

Table 2
Characteristics of the study population.

	All (n = 424)	Thinking/fatigue-targeted subpopulation[a] (n = 182)
	n (%)	
Missing data	10 (2%)	1 (1%)
Gender: women	290 (68%)	132 (73%)
Age group: 40 years or less	66 (16%)	37 (20%)
41–50 years	137 (32%)	55 (30%)
51–60 years	170 (40%)	73 (40%)
over 60 years	51 (12%)	17 (9%)
Disease duration: 5 years or less	62 (15%)	29 (16%)
6–10 years	86 (20%)	45 (25%)
11–20 years	148 (35%)	62 (34%)
over 20 years	128 (30%)	46 (25%)
Type of MS: relapsing remitting	174 (41%)	93 (51%)
secondary progressive	180 (42%)	67 (37%)
primary progressive	70 (17%)	22 (12%)
EDSS: EDSS $<$ 3.5	124 (29%)	70 (38%)
EDSS 3.5–5.5	140 (33%)	61 (34%)
EDSS 5.0–6.5	145 (34%)	48 (26%)
EDSS $>$ 6.5	15 (4%)	3 (2%)
Work situation full-time	15 (4%)	7 (4%)
part-time	80 (19%)	40 (22%)
unemployed	51 (12%)	30 (17%)
retired/early retirement	261 (63%)	101 (56%)
otherwise	5 (1%)	3 (2%)
Current or former smoker Yes	286 (69%)	129 (71%)
Marital status married	288 (70%)	133 (73%)
	median (Q1, Q3)	
MFIS	41 (30–50)	43 (33–50)
BMI	24 (22–27)	24 (23–28)
BDI	11 (6–17)	11 (7–16)
FAMS	114 (96–133)	114 (98–130)
6MWT	239 (107–356)	270 (119–315)

[a] This subpopulation was used to investigate the responsiveness of MFIS as mean marginal changes in MFIS from baseline to discharge after one month of multidisciplinary rehabilitation.

3.2. Local independence, monotonicity, multivariate normality and tau-equivalence

All MFIS items passed the Mokken scale monotonicity and local independence checks (see supplementary Table S2). The assumption of multivariate normality was rejected for both the total MFIS scale and the Cognitive and Physical subscales, as indicated by Mardia tests for skewness and kurtosis (all, $p < 0.001$) (see supplementary Table S1). Hence, the DWLS estimator was applied in all factor-analytic analyses. CFA models addressing tau-equivalence rejected both the parallel items model (LRT, $p < 0.001$; ECVI = 2.98) and the tau-equivalence model (LRT, $p < 0.001$; ECVI = 2.87), when compared to the congeneric model (ECVI = 1.44). Model fit indices of the congeneric model indicated that this model is an adequate representation of the data (CFI = 0.98, TLI = 0.97, RMSEA = 0.06, RMSEA 90%-UCL = 0.07, SRMR = 0.08) and is superior to the parallel item or tau-equivalence version of the model, see Table 3. This indicates that split-half reliability (requiring the parallel condition) and Cronbach's alpha (requiring tau-equivalence) are inadequate estimators of global test reliability for the MFIS scale, and, hence, McDonald's omega (ω) was calculated in the present study.

3.3. Structural validity and reliability

MFIS in its original form did show poor fit to one-dimensional models that ignore the subscale structure of MFIS regardless of the model framework (Table 3). However, when acknowledging the subscale structure of MFIS, model fit indices improved to acceptable fits. This was the case for both bi-factor models and 3-dimensional measurement models in the factor-analytic and Rasch-analytic context (Table 3 and Table 4). Rasch-analytic models showed generally poorer fits in line with their higher scale standards according to theory. Standardized loadings are shown in Table 5 and Fig. 1. The bi-factor models supported uni-dimensionality of MFIS as indicated by Mc Donald's ω_h = 0.72, however, with considerable substructure present (ECV = 0.47), as also indicated by the eigenvalues of the general and specific factors Cognitive, Physical and Psychosocial were 6.48, 4.81, 2.24, and 0.22, respectively. Reliability of the MFIS scale was found to be high as indicated by a Mc Donald's ω_t of 0.96. All items showed loadings onto the general factor greater than 0.30, but not all items showed loadings onto their specific factor exceeding the 0.30 threshold (items 1, 8, 14, and 21; see Table 5). The Psychosocial subscale was poorly supported in the bi-factor model, with its two sole items (8 and 9) showing considerable larger loadings onto the general factor than onto their specific factor (Table 5). Item 8 was even found to be uncorrelated with the Psychosocial subscale conditional on its loading onto the general factor.

Fit indices for the 3-dimensional measurement model were comparable to the bi-factor model (Table 3). All items showed standardized loadings onto their specific factor that were larger than 0.4, indicating sufficient convergence of items onto their factors (Fig. 1, Table 5). This was also found for the two items 8 and 9 of the Psychosocial subscale. Inter-subscale correlations and their 95%-CI were estimated to be 0.44 (0.42,0.47), 0.81 (0.72,0.9), and 0.57 (0.5,0.64) for Physical-Cognitive, Physical-Psychosocial, and Cognitive-Psychosocial correlations, respectively. This indicates sufficient discrimination between the Physical and Cognitive as well as between the Cognitive and Psychosocial subscales, but not between the Physical and Psychosocial subscales. AVE estimates for Cognitive, Physical, and Psychosocial subscales were found to be 0.63, 0.47, and 0.52, respectively, with AVE = 0.55 for the total scale. This indicates that items in the Physical subscale on average are marginally insufficient representatives of the Physical subconstruct, when taking measurement error into account. This is due to items 4, 13, and 14, that all showed $R^2 < 0.40$. Mc Donald's ω_t reliability estimates were 0.94, 0.89, and 0.68 for the Cognitive, Physical, and Psychosocial subscales, respectively, compared to a ω_t = 0.95 for the total scale.

The 3D two-layered structural model showed fit indices comparable to the 3D bi-factor and congeneric 3D one-layered measurement models

Table 3
Model fit statistics. Model fit sufficiency is indicated with a good fit shown in bold and a very good fit shown in bold italics.

		Model	TLI	CFI	RMSEA	Upper 95%- CL	SRMR
	CFA models						
		1D1L	0.86	0.87	0.15	0.15	0.15
		2D1L	***0.97***	***0.97***	**0.07**	**0.08**	**0.08**
		3D1L, parallel	**0.93**	**0.93**	0.10	0.11	0.12
		3D1L, tau-equivalent	**0.93**	**0.93**	0.10	0.11	0.12
		3D1L, congeneric	***0.97***	***0.98***	**0.06**	**0.07**	**0.08**
		Bi-factor 3D	***0.99***	***0.99***	***0.03***	**0.04**	**0.05**
		3D2L	***0.97***	***0.97***	**0.07**	**0.08**	**0.08**
IRT models							
	Rasch models						
		1D RSM	0.69	0.61	0.23	0.24	n.a.
		1D RSM_{Cogn}	***0.96***	**0.93**	0.13	0.14	n.a.
		1D RSM_{Phys}	**0.94**	**0.91**	0.11	0.12	n.a.
		1D RSM_{Cogn5} [a]	***0.98***	***0.95***	0.11	0.13	n.a.
		1D RSM_{Phys8} [a]	***0.96***	**0.93**	0.10	0.11	n.a.
		1D PCM	0.62	0.62	0.26	0.26	0.18
		2D PCM	**0.93**	**0.93**	0.11	0.12	0.11
		3D PCM	**0.93**	**0.93**	0.11	0.12	0.10
		Bi-factor 3D PCM	**0.92**	**0.92**	0.11	0.12	0.11
		3D GRM	***0.96***	***0.96***	**0.08**	**0.09**	0.10

CFA: confirmatory factor analysis.
1D: one-dimensional model of MFIS.
2D: two-dimensional measurement model of MFIS with the Psychosocial subscale combined with the Physical subscale.
3D: three-dimensional measurement model of MFIS with Cognitive, Physical and Psychosocial subscales.
1L: Single-layered measurement model, with no hierarchical structure among sub-components.
2L: Two-layered structural model, with the Psychosocial sub-component secondary to the Physical subcomponent.
IRT: Item response theory.
RSM: Rating Scale Model; PCM: Partial Credit Model; GRM: Graded Response Model.
TLI: Tucker-Lewis Index (TFI, good fit: $0.90 \leq CFI \leq 0.95$, very good fit: ≥ 0.95).
CFI: Comparative Fit Index (CFI, good fit: $0.90 \leq CFI \leq 0.95$, very good fit: ≥ 0.95).
RMSEA: Root Mean Squared Error of Approximation (RMSEA, ≤ 0.05, and with RMSEA 90% upper confidence limits 90%-UCL ≤ 0.1).
SRMR: Standardized Root Mean Squared Residuals (SRMSR, ≤ 0.08).
[a] : Scales suggested by Mills et al.[14].

Table 4
Infit fit statistics for multidimensional and unidimensional Rasch analyses. Infit and outfit mean squared (MS) values between 0.5 and 1.5 are considered satisfactory for measurement (Linacre, 2002). Deviating items are shown in bold.

			Multidimensional Rasch model						Unidimensional Rasch model		
			3D RSM		3D PCM		Bi-factor 3D PCM		1D PCM_{Cogn}		
			Infit MS	Outfit MS	Infit MS	Outfit MS	Infit MS	Outfit MS	Infit MS	Outfit MS	Discrimination
Cogn.	01	Alertness	0.71	0.69	**1.61**	**1.63**	1.50	**1.50**	**1.62**	**1.65**	0.47
	02	Attention	0.90	0.91	0.82	0.83	0.78	0.80	0.75	0.77	0.82
	03	Clear.thinking	0.83	0.81	0.96	0.97	0.93	0.95	0.93	0.95	0.75
	05	Forgetfulness	0.95	0.94	1.14	1.13	1.15	1.15	0.96	0.96	0.75
	11	Making.decisions	0.81	0.82	0.95	0.96	0.92	0.95	0.79	0.81	0.82
	12	Motiv.mental.effort	0.85	0.88	0.92	0.96	0.91	0.96	0.56	0.56	0.88
	15	Finish.tasks	0.73	0.75	0.67	0.67	0.62	0.62	0.84	0.82	0.69
	16	Organize.thoughts	0.76	0.78	0.7	0.7	0.65	0.65	0.79	0.80	0.73
	18	Slow.thinking	0.97	0.99	0.83	0.84	0.8	0.83	1.39	1.39	0.47
	19	Concentration	0.79	0.80	0.67	0.66	0.61	0.61	0.61	0.60	0.82
									1D PCM_{Phys}		
Phys.	04	Coordination	1.28	1.24	1.18	1.18	1.26	1.24	1.24	1.19	0.69
	06	Pace.phys.act	0.85	0.86	0.74	0.74	0.7	0.71	0.94	0.97	0.76
	07	Motiv.phys.effort	0.76	0.77	0.86	0.85	0.83	0.82	0.64	0.58	0.87
	10	Maintain.phys.effort	1.18	1.19	0.80	0.81	0.83	0.84	0.62	0.62	0.87
	13	Muscle.weakness	1.32	1.28	0.94	0.94	1.01	0.99	1.29	1.24	0.47
	14	Phys.uncomfortable	1.30	1.32	1.26	1.28	1.32	1.34	0.67	0.68	0.77
	17	Complete.phys.tasks	0.89	0.90	0.66	0.66	0.64	0.64	1.00	0.95	0.64
	20	Limit.phys.act	1.05	1.05	0.74	0.75	0.73	0.75	0.71	0.71	0.76
	21	Have.rest	1.02	1.03	0.93	0.94	0.92	0.94	0.97	0.97	0.64
PsSoc.	08	Motiv.social.act	0.83	0.84	0.86	0.84	0.54	0.54	n.a.	n.a.	n.a.
	09	Outdoor.act	1.17	1.15	0.79	0.78	0.57	0.58	n.a.	n.a.	n.a.

3D: three-dimensional MFIS model with Cognitive, Physical and Psychosocial subcomponents.
RSM: Rating Scale Model; PCM: Partial Credit Model.

Table 5
Results of confirmatory factor analyses (CFA) including 3-dimensional (3D) bi-factor analysis, 3D one-layered (3D1L) and 3D two-layered (3D2L) CFA. Item loadings < 0.40 are indicated in bold.

Subscale	Item		Bi-factor 3D			3D1L		3D2L	
			Std. loadings (95%-CI)		R^2	Std. loadings (95%-CI)	R^2	Std. loadings (95%-CI)	R^2
			General factor *g*	Specific factor					
Cogn.	01	Alertness	0.64 (0.56,0.71)	**0.19 (0.12,0.27)**	0.60	0.62 (0.55,0.68)	0.38	0.62 (0.56,0.68)	0.38
	02	Attention	0.42 (0.36,0.48)	0.71 (0.64,0.79)	0.58	0.81 (0.75,0.87)	0.66	0.81 (0.75,0.87)	0.66
	03	Clear.thinking	0.43 (0.37,0.49)	0.63 (0.55,0.70)	0.44	0.76 (0.70,0.82)	0.57	0.76 (0.70,0.82)	0.58
	05	Forgetfulness	**0.36 (0.30,0.42)**	0.60 (0.53,0.68)	0.69	0.69 (0.63,0.74)	0.47	0.69 (0.63,0.74)	0.47
	11	Making.decisions	0.48 (0.42,0.54)	0.60 (0.53,0.67)	0.58	0.78 (0.72,0.84)	0.61	0.78 (0.72,0.84)	0.61
	12	Motiv.mental.effort	0.44 (0.38,0.50)	0.62 (0.55,0.70)	0.49	0.76 (0.71,0.82)	0.58	0.76 (0.71,0.82)	0.58
	15	Finish.tasks	0.47 (0.41,0.53)	0.73 (0.66,0.81)	0.59	0.86 (0.80,0.93)	0.74	0.86 (0.80,0.92)	0.74
	16	Organize.thoughts	0.50 (0.44,0.56)	0.71 (0.63,0.78)	0.58	0.87 (0.81,0.93)	0.76	0.87 (0.81,0.93)	0.76
	18	Slow.thinking	0.42 (0.36,0.48)	0.71 (0.64,0.79)	0.76	0.81 (0.75,0.87)	0.65	0.81 (0.75,0.87)	0.66
	19	Concentration	0.44 (0.38,0.50)	0.79 (0.71,0.87)	0.75	0.87 (0.81,0.93)	0.76	0.87 (0.81,0.93)	0.76
Phys.	04	Coordination	**0.35 (0.29,0.40)**	**0.34 (0.25,0.44)**	0.58	0.48 (0.43,0.53)	0.23	0.47 (0.42,0.52)	0.22
	06	Pace.phys.act	0.54 (0.47,0.60)	0.58 (0.47,0.69)	0.65	0.75 (0.68,0.81)	0.56	0.73 (0.66,0.79)	0.53
	07	Motiv.phys.effort	0.64 (0.57,0.71)	**0.38 (0.28,0.48)**	0.24	0.76 (0.70,0.83)	0.58	0.75 (0.68,0.81)	0.56
	10	Maintain.phys.effort	0.45 (0.39,0.51)	0.63 (0.52,0.74)	0.62	0.68 (0.62,0.74)	0.46	0.66 (0.60,0.71)	0.43
	13	Muscle.weakness	**0.36 (0.30,0.42)**	0.55 (0.44,0.65)	0.55	0.57 (0.52,0.63)	0.33	0.55 (0.50,0.61)	0.31
	14	Phys.uncomfortable	0.56 (0.50,0.62)	**0.12 (0.03,0.21)**	0.60	0.57 (0.52,0.62)	0.32	0.56 (0.51,0.61)	0.32
	17	Complete.phys.tasks	0.55 (0.49,0.62)	0.65 (0.53,0.76)	0.43	0.78 (0.72,0.85)	0.61	0.76 (0.70,0.82)	0.58
	20	Limit.phys.act	0.55 (0.48,0.61)	0.54 (0.43,0.65)	0.33	0.74 (0.68,0.81)	0.55	0.72 (0.66,0.79)	0.52
	21	Have.rest	0.74 (0.67,0.81)	**0.19 (0.09,0.29)**	0.72	0.78 (0.71,0.84)	0.60	0.76 (0.70,0.82)	0.58
PsSoc.	08	Motiv.social.act	0.76 (0.70,0.83)	**0.00 (-0.28,0.27)**	0.69	0.76 (0.66,0.85)	0.57	0.71 (0.63,0.79)	0.51
	09	Outdoor.act	0.69 (0.63,0.75)	0.42 (0.24,0.60)	0.81	0.69 (0.61,0.77)	0.48	0.74 (0.67,0.82)	0.55

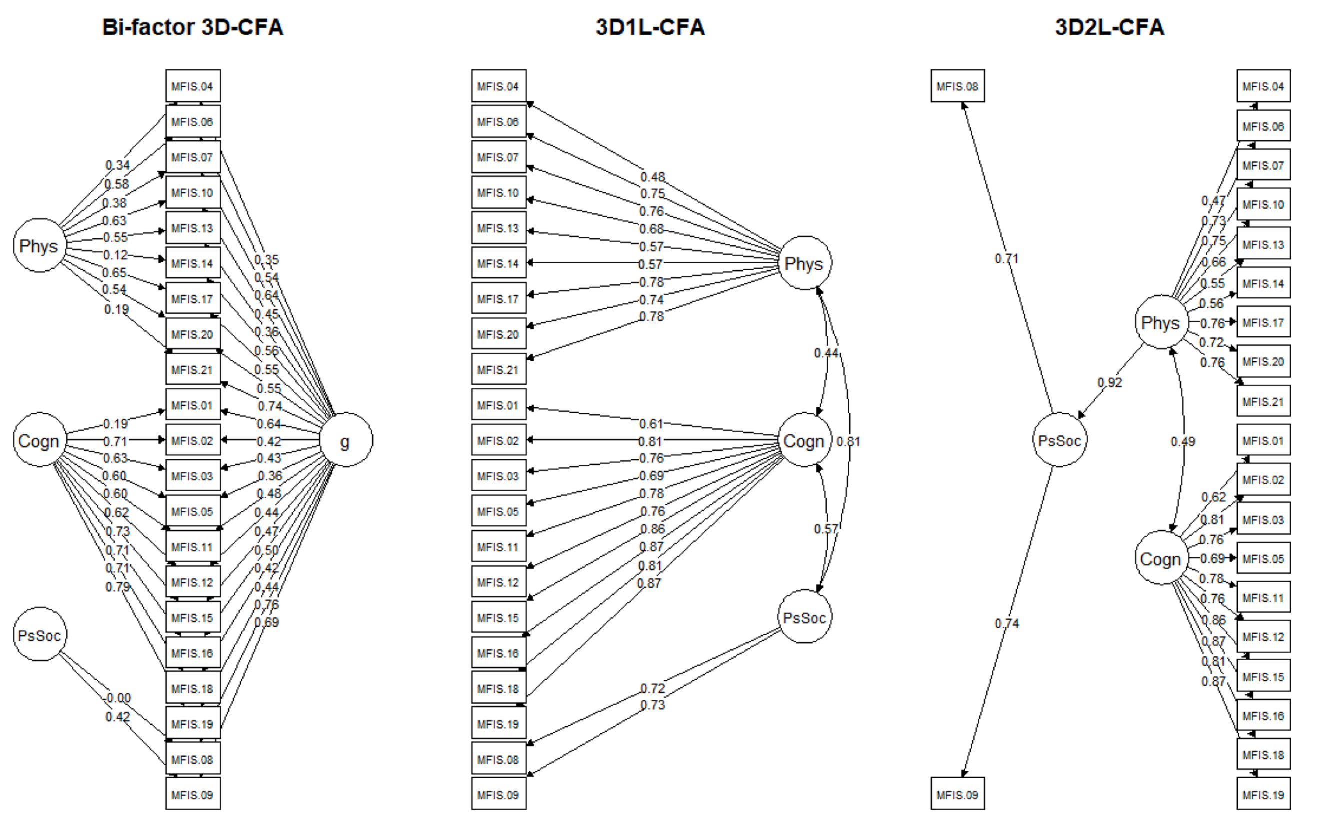

Fig 1. The bi-factor 3-dimensional model (3D), the 3D one-layered measurement model (3D1L), and the 3D two-layered (3D2L) structural equation model of MFIS. All estimates are standardized estimates (correlations). Manifest standard deviations are not shown.

(Table 3), and standard loadings comparable to the congeneric 3D one-layered model (Table 5).

3.4. Group invariance and cross-cultural validity

Results on group invariance within the current study population revealed no differences in group comparisons between hospitalization locality (Ry vs. Haslev) in 1D PCM models of the Cognitive and Physical subdimensions but did reveal a difference between males and females for the 1D PCM model of the Physical subdimension (see supplementary Table S3). The Tucker index of factor congruence between PCA loadings reported by Kos et al. (2005) in a multinational European cohort and PCA loadings from the current Danish study sample revealed a high degree of factor congruence between the cross-cultural loadings in the

Cognitive and Physical subdomains, respectively (Tucker index: Cognitive, 0.97, and Physical, 0.93), but very limited congruence was observed for the Psychosocial subdomain (Tucker index: 0.25). The overall RV coefficient of matrix correlations was estimated to be 0.81 ($p < 0.001$).

3.5. Nomological and criterion validity

Fatigue, measured by the MFIS total score, was found to be unrelated to age, time since diagnosis, and EDSS score, see supplementary Table S4. Exploratory factor analysis revealed an apparently larger nomological network for cognitive fatigue than the physical subdomain of fatigue, as both FAMS Thinking/Fatigue, BDI, and 6MWT loads into the latent factor associated with MFIS Cognitive whereas only the FAMS Mobility and 6MWT performance loads into the latent factor associated with MFIS Physical (Table 6).

3.6. Responsiveness

The MFIS showed overall good responsiveness, as large effect sizes were found on the total score and in the physical subdomain and moderate effect sizes were found for the cognitive and psychosocial subdomains following one month of MDR (Table 7).

4. Discussion

Previous validations of the MFIS have applied different validation methodologies and have reached different conclusions on the validity of the scale (Kos et al., 2005; Mills et al., 2010; Amtmann et al., 2012; Bakalidou et al., 2014). The present study provide the first ever validation of the MFIS following the COSMIN checklist, ensuring a standardized, modern, and comprehensive multimethodological evaluation of the measurement properties of the scale. As this validation is based on the Danish translation of the MFIS, the present study also serves as the first validation of the Danish version of the MFIS.

The present study found that both 3D one-layered measurement, two-layered structural, and bi-factor CFA models showed acceptable fits of the current MFIS data. The fact that the two-layered structural modeling of the data (psychosocial domain modelled as a secondary structure to the cognitive and physical domains) shows an equally good fit of the data as the one-layered measurement model underline the general finding of a limited benefit of the psychosocial subscale. However, the question whether the perception of psychosocial fatigue is secondary to cognitive and physical fatigue remains unanswered. Moreover, eigenvalues of general and specific factors indicated a structure of a dominant general factor, with underlying specific factors.

Table 6
Exploratory Factor Analysis loadings of MFIS subscales and its nomological net including Health Related Quality of Life (HRQoL) measured by FAMS subscales, Beck's Depression Inventory II (BDI) and physical performance as measured by the 6 min walking test (6MWT). Factor loadings < 0.20 are not shown.

	Latent factor		
	F1 Cognitive	F2 Emotional	F3 Physical
$MFIS_{Cognitive}$	0.95		
$MFIS_{Physical}$	0.42		−0.56
$MFIS_{Psychosocial}$	0.39		−0.41
$FAMS_{Mobility}$			0.98
$FAMS_{Symptoms}$	−0.39		
$FAMS_{Emotional\ wellbeing}$		0.86	
$FAMS_{General\ contentment}$		0.88	
$FAMS_{Thinking/fatigue}$	−0.84		
$FAMS_{Family/social\ wellbeing}$		0.6	
BDI	0.31	−0.49	
6MWT	0.27		0.52
Proportion of variance explained	0.21	0.2	0.16
Cumulative variance explained	0.21	0.4	0.56

Table 7
Cohen's effect sizes d for the MFIS total scale and the cognitive, physical, and psychosocial subscales after one month of MDR.

Scale	Cohen's d	95% CI		p-value
MFIS	−0.88	−1.02	−0.73	<0.0001
$MFIS_{Cognitive}$	−0.60	−0.72	−0.48	<0.0001
$MFIS_{Physical}$	−0.90	−1.05	−0.74	<0.0001
$MFIS_{Psychosocial}$	−0.48	−0.64	−0.33	<0.0001

Generally, all items loaded satisfactory onto the general factor, but the specific psychosocial subscale was poorly supported as its two items showed larger loading onto the general rather than their specific factor. These findings are well in line with the findings of Kos et al. in the multinational European validation of the MFIS and Bakalidou et al. in the Greek validation (Kos et al., 2005; Bakalidou et al., 2014). Also aligned with previous findings, the reliability of the total score and the cognitive and physical subscales of the MFIS showed good reliability in the current study, whereas the psychosocial subscale did not (Kos et al., 2005; Larson, 2013). The low reliability of the Psychosocial subscales is probably related to only two items constituting the subscale. Overall, the Psychosocial subscale shows low reliability, too high correlation to the Physical subscale, and its items are more related to a general factor than to their own subscale factor. All this indicates that the Psychosocial subscale is poorly defined and cannot be recommended to be used separately. However, the items may still be informative for assessment of the general underlying factor in MFIS as indicated by the BFA.

One previous study evaluated the MFIS in MS using Rasch analysis, and Mills et al. found that the MFIS did not fit the unidimensional Rasch model applied and concluded that the summed score of MFIS is invalid because of multidimensionality (Mills et al., 2010). The authors did find an underlying structure with cognitive and physical subdomains, but was forced to remove items 1, 2, 3, 4, 5, 11, 14, and 17 in order to obtain acceptable model fits. Instead of discarding the MFIS total score based on a unidimensional model we applied multidimensional Rasch analyses, relaxing the assumption of unidimensionality. Using this approach, we found acceptable fits to the models of the total score as well as the cognitive and physical subscales, indicating that these are valid measures of fatigue impact. We also found that correlations between total sum scores and corresponding predicted latent factor values were high for both the CFA models and the more restrictive Rasch models, indicating that the degree of misfit to an ideal Rasch scale influenced MFIS sum scores only to a limited degree in the current data. Cross-cultural PCA loadings revealed a high degree of factor congruence in the cognitive and physical subdomains, but not the psychosocial domain, again adding to the overall validity of the cognitive and physical subscales but also underlining the limited validity of the psychosocial subscale.

In the assessment of the validity of an instrument, it is also important to consider the nomological network of the construct, and to evaluate theoretically sound associations to relevant other constructs (Larson, 2013). The findings of relevant loadings of FAMS subdimensions, BDI, and 6MWT onto the latent factor associated with MFIS Cognitive and FAMS Mobility and 6MWT performance onto the latent factor of MFIS physical are confirmative of the nomological validity of MFIS, and also replicate previous findings (Larson, 2013; Tellez et al., 2005). However, this should also be confirmed by associations to other more fatigue-specific validated constructs (i.e. the Fatigue Severity Scale and/or objective measures of fatigability) in future studies.

The MFIS has been reported to have limited responsiveness to changes in patients with MS (Rietberg et al., 2010), but important for the use in clinical trials, the MFIS showed good responsiveness over time following a 1-month MDR intervention in the current study.

5. Conclusion

The Danish version of the MFIS showed good reliability, a good structural, cross-cultural, and nomological validity with a general and underlying structure including cognitive and physical subscales, an acceptable fit to a multidimensional Rasch analysis, and a good responsiveness. The psychosocial subscale of the MFIS however showed low reliability and low factor loadings and on this basis we do not recommend to interpret the scores of this subscale independently.

CRediT authorship contribution statement

Morten Riemenschneider: Conceptualization, Methodology, Validation, Writing – original draft, Writing – review & editing. **Philipp Trénel:** Conceptualization, Methodology, Data curation, Formal analysis, Software, Validation, Writing – review & editing. **Michael Nørgaard:** Conceptualization, Data curation, Writing – review & editing. **Finn Boesen:** Conceptualization, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors have no conflicts of interests associated with the present study.

Funding

The present study was not financed by any external funds.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.msard.2022.104012.

References

Amato, M.P., Ponziani, G., Rossi, F., et al., 2001. Quality of life in multiple sclerosis: the impact of depression, fatigue and disability. Mult. Scler. 7, 340–344.
Amtmann, D., Bamer, A.M., Noonan, V., et al., 2012. Comparison of the psychometric properties of two fatigue scales in multiple sclerosis. Rehabil. Psychol. 57, 159–166. https://doi.org/10.1037/a0027890, 2012/06/13.
Bakalidou, D., Voumvourakis, K., Tsourti, Z., et al., 2014. Validity and reliability of the Greek version of the Modified Fatigue Impact Scale in multiple sclerosis patients. Int. J. Rehabil. Res. 37, 271–276. https://doi.org/10.1097/MRR.0000000000000057, 2014/02/22.
Boateng, G.O., Neilands, T.B., Frongillo, E.A., et al., 2018. Best practices for developing and validating scales for health, social, and behavioral research: a primer. Front. Public Health 6, 149. https://doi.org/10.3389/fpubh.2018.00149, 2018/06/27.
Boesen, F., Norgaard, M., Trenel, P., et al., 2018. Longer term effectiveness of inpatient multidisciplinary rehabilitation on health-related quality of life in MS patients: a pragmatic randomized controlled trial - the Danish MS Hospitals Rehabilitation Study. Mult. Scler. 24, 340–349. https://doi.org/10.1177/1352458517735188, 2017/10/07.
Callesen, J., Cattaneo, D., Brincks, J., et al., 2020. How do resistance training and balance and motor control training affect gait performance and fatigue impact in people with multiple sclerosis? A randomized controlled multi-center study. Mult. Scler. 26, 1420–1432. https://doi.org/10.1177/1352458519865740, 2019/07/25.
Chalmers, R.P., 2012. Mirt: a multidimensional item response theory package for the R environment. J. Stat. Softw. 48, 1–29. https://doi.org/10.18637/jss.v048.i06.
Dima, A.L., 2018. Scale validation in applied health research: tutorial for a 6-step R-based psychometrics protocol. Health Psychol. Behav. Med. 6, 136–161. https://doi.org/10.1080/21642850.2018.1472602, 2018/05/10.
Filippi, M., Bar-Or, A., Piehl, F., et al., 2018. Multiple sclerosis. Nat. Rev. Dis. Prim. 4, 43. https://doi.org/10.1038/s41572-018-0041-4, 2018/11/10.
Fisk, J.D., Ritvo, P.G., Ross, L., et al., 1994. Measuring the functional impact of fatigue: initial validation of the fatigue impact scale. Clin. Infectious Dis. 18 (Suppl 1), S79–S83.
Gana K. and Broc G. Structural equation modeling with lavaan. 2019.
Green, R., Cutter, G., Friendly, M., et al., 2017. Which symptoms contribute the most to patients' perception of health in multiple sclerosis? Mult. Scler. J. Exp. Transl. Clin. 3, 2055217317728301 https://doi.org/10.1177/2055217317728301, 2017/09/15.
Hobart, J.C., Riazi, A., Lamping, D.L., et al., 2004. Improving the evaluation of therapeutic interventions in multiple sclerosis: development of a patient-based measure of outcome. Health Technol. Assess. 8, 1–48. https://doi.org/10.3310/hta8090 iii2004/02/26.
Johansson, S., Skjerbæk, A.G., Nørgaard, M., et al., 2021. Associations between fatigue impact and lifestyle factors in people with multiple sclerosis – the Danish MS hospitals rehabilitation study. Mult. Scler. Relat. Disord. 50, 102799 https://doi.org/10.1016/j.msard.2021.102799.
Kinney, A.R., Eakman, A.M., Graham, J.E., 2020. Novel effect size interpretation guidelines and an evaluation of statistical power in rehabilitation research. Arch. Phys. Med. Rehabil. 101, 2219–2226. https://doi.org/10.1016/j.apmr.2020.02.017, 2020/04/10.
Kluger, B.M., Krupp, L.B., Enoka, R.M., 2013. Fatigue and fatigability in neurologic illnesses: proposal for a unified taxonomy. Neurology 80, 409–416. https://doi.org/10.1212/WNL.0b013e31827f07be [doi].
Kos, D., Kerckhofs, E., Carrea, I., et al., 2005. Evaluation of the Modified Fatigue Impact Scale in four different European countries. Mult. Scler. 11, 76–80. https://doi.org/10.1191/1352458505ms1117oa, 2005/03/01.
Larson, R.D., 2013. Psychometric properties of the modified fatigue impact scale. Int. J. MS Care 15, 15–20. https://doi.org/10.7224/1537-2073.2012-019, 2014/01/24.
Linacre, J.M., 2002. What do infit and outfit, mean-square and standardized mean. Rasch Measur. Trans. 16, 878.
Mair, P., 2018. Modern Psychometrics With R. New York: Springer.
Mills, R.J., Young, C.A., Pallant, J.F., et al., 2010. Rasch analysis of the Modified Fatigue Impact Scale (MFIS) in multiple sclerosis. J. Neurol. Neurosurg. Psychiatry 81, 1049–1051. https://doi.org/10.1136/jnnp.2008.151340, 2010/06/16.
Mokkink, L.B., Terwee, C.B., Patrick, D.L., et al., 2010. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual. Life Res. 19, 539–549. https://doi.org/10.1007/s11136-010-9606-8, 2010/02/20.
Guidelines MSCfCP. *Fatigue and multiple sclerosis: evidence-based management strategies for fatigue in multiple sclerosis*. 1998.
Oervik, M.S., Sejbaek, T., Penner, I.K., et al., 2017. Validation of the fatigue scale for motor and cognitive functions in a danish multiple sclerosis cohort. Mult. Scler. Relat. Disord. 17, 130–134. https://doi.org/10.1016/j.msard.2017.07.017, 2017/10/23.
psych, W.R., 2020. Procedures For Psychological, Psychometric, and Personality Research. Northwestern University, Evanston, Illinois. R package version 2.0.12.
Riemenschneider, M., Hvid, L.G., Ringgaard, S., et al., 2021. Study protocol: randomised controlled trial evaluating exercise therapy as a supplemental treatment strategy in early multiple sclerosis: the early multiple sclerosis exercise study (EMSES). BMJ Open 11, e043699. https://doi.org/10.1136/bmjopen-2020-043699, 2021/01/14.
Rietberg, M.B., Van Wegen, E.E., Kwakkel, G., 2010. Measuring fatigue in patients with multiple sclerosis: reproducibility, responsiveness and concurrent validity of three Dutch self-report questionnaires. Disabil. Rehabil. 32, 1870–1876. https://doi.org/10.3109/09638281003734458, 2010/03/30.
Rommer, P.S., Eichstadt, K., Ellenberger, D., et al., 2019. Symptomatology and symptomatic treatment in multiple sclerosis: results from a nationwide MS registry. Mult. Scler. 25, 1641–1652. https://doi.org/10.1177/1352458518799580, 2018/09/20.
Rosseel, Y., 2012. lavaan: an R Package for Structural Equation Modeling. J. Stat. Softw. 48, 1–36. https://doi.org/10.18637/jss.v048.i02.
Squara, P., Scheeren, T.W.L., Aya, H.D., et al., 2021a. Metrology part 1: definition of quality criteria. J. Clin. Monit. Comput. 35, 17–25. https://doi.org/10.1007/s10877-020-00494-y, 2020/03/19.
Squara, P., Scheeren, T.W.L., Aya, H.D., et al., 2021b. Metrology part 2: procedures for the validation of major measurement quality criteria and measuring instrument properties. J. Clin. Monit. Comput. 35, 27–37. https://doi.org/10.1007/s10877-020-00495-x, 2020/03/19.
Taul-Madsen, L., Dalgas, U., Kjolhede, T., et al., 2020. A head-to-head comparison of an isometric and a concentric fatigability protocol and the association with fatigue and walking in persons with multiple sclerosis. Neurorehabil. Neural Repair 34, 523–532. https://doi.org/10.1177/1545968320920250, 2020/05/13.
Tellez, N., Rio, J., Tintore, M., et al., 2005. Does the Modified Fatigue Impact Scale offer a more comprehensive assessment of fatigue in MS? Mult. Scler. 11, 198–202. https://doi.org/10.1191/1352458505ms1148oa, 2005/03/30.
Thrue, C., Riemenschneider, M., Hvid, L.G., et al., 2020. Time matters: early-phase multiple sclerosis is accompanied by considerable impairments across multiple domains. Mult. Scler. https://doi.org/10.1177/1352458520936231, 13524585209362312020/07/03.
Trojan, D.A., Arnold, D., Collet, J.P., et al., 2007. Fatigue in multiple sclerosis: association with disease-related, behavioural and psychosocial factors. Mult. Scler. 13, 985–995. https://doi.org/10.1177/1352458507077175, 2007/05/01.
Whitehead, L., 2009. The measurement of fatigue in chronic illness: a systematic review of unidimensional and multidimensional fatigue measures. J. Pain Symptom Manag. 37, 107–128. https://doi.org/10.1016/j.jpainsymman.2007.08.019, 2008/12/30.