J Rehabil Med 2014; 46: 761–767

ORIGINAL REPORT

MEASURING FATIGUE IN POLIO SURVIVORS: CONTENT COMPARISON AND RELIABILITY OF THE FATIGUE SEVERITY SCALE AND THE CHECKLIST INDIVIDUAL STRENGTH

Fieke S. Koopman, MD[1], Merel A. Brehm, PhD[1], Yvonne F. Heerkens, PhD[2], Frans Nollet, MD, PhD[1] and Anita Beelen, PhD[1]

From the [1]Department of Rehabilitation, Academic Medical Center, University of Amsterdam, Amsterdam and [2]Dutch Institute of Allied Health Professions, Amersfoort, and Department of Occupation & Health, HAN University of Applied Sciences, Nijmegen, The Netherlands

Objectives: **(*i*) To compare the content of the Fatigue Severity Scale and the subscale "subjective experience of fatigue" of the Checklist Individual Strength, and (*ii*) to assess the reliability of both questionnaires in polio survivors.**
Design: **Repeated-measures at a 3-week interval.**
Subjects: **Consecutive series of 61 polio survivors.**
Methods: **Concepts contained in the questionnaire items were linked to the International Classification of Functioning, Disability and Health (ICF), using standardized linking rules. Reliability analyses included tests of internal consistency, test-retest reliability and measurement error.**
Results: **Questionnaires differed in the extent to which they measured other than fatigue-related aspects of functioning (represented ICF components: "Body functions": 50% and 80%, "Activities and Participation": 30% and 0%, for the Fatigue Severity Scale and Checklist Individual Strength, respectively). Internal consistency and test-retest reliability were considered acceptable, while measurement error was large (Cronbach's α: 0.90 and 0.93, intraclass correlation coefficient: 0.80 and 0.85, smallest detectable change: 28.7% and 29.4% for the Fatigue Severity Scale and Checklist Individual Strength, respectively).**
Conclusion: **Considering the acceptable clinimetric properties, we conclude that both the Fatigue Severity Scale and the Checklist Individual Strength can be applied in research on post-poliomyelitis syndrome when measuring fatigue. However, because the 2 questionnaires differ in content they cannot be used interchangeably.**

Key words: fatigue; post-poliomyelitis syndrome; questionnaires.

J Rehabil Med 2014; 46: 761–767

Correspondence address: Fieke S. Koopman, Department of Rehabilitation, Academic Medical Center, University of Amsterdam, PO 22660, 1100 DD Amsterdam, The Netherlands. E-mail: s.koopman@amc.uva.nl

Accepted Mar 18, 2014; Epub ahead of print Jun 17, 2014

INTRODUCTION

Fatigue, defined as an overwhelming sense of tiredness, lack of energy and feeling of exhaustion (1), is one of the most frequently reported complaints in persons with post-poliomyelitis syndrome (PPS) (2, 3). Compared with healthy individuals, persons with PPS experience high levels of fatigue (4) that persist over time (5). Because fatigue in PPS impacts negatively on quality of life (6) it is a relevant problem in this patient group.

Fatigue is commonly quantified by means of self-report questionnaires. Studies have used different questionnaires to investigate the prevalence, severity and natural course of fatigue in PPS (5), as well as to evaluate the effects of treatment to reduce fatigue (7–10). The most frequently used instrument to measure fatigue in PPS is the Fatigue Severity Scale (FSS) (11), which was originally developed and tested in multiple sclerosis and systemic lupus erythematosus (11). More recently, the Checklist Individual Strength (CIS20) has been introduced for assessing fatigue within PPS research (12). The CIS20 was formerly developed for chronic fatigue syndrome (CFS) (13), and has been recommended for research purposes by the International Chronic Fatigue Syndrome Study Group (14). Although the CIS20 consists of 4 sub-scales related to fatigue, the sub-scale subjective experience of fatigue (CIS20-F) is generally used as the primary outcome measure in studies assessing fatigue (15, 16). Over the past decade, this scale has also become increasingly popular for assessing fatigue in neuromuscular disease (17, 18).

The choice of fatigue instrument is dependent on the construct to be measured. Because fatigue is a complex, multidimensional problem (5), care should be taken in comparing the outcomes from different fatigue instruments. That is, instruments may vary greatly with regard to the International Classification of Functioning, Disability and Health (ICF) categories covered (19), indicating that they measure different constructs. At face value, the FSS and CIS20-F appear to measure different constructs (i.e. aspects of fatigue), although detailed information on their content validity is lacking. The current study therefore assessed and compared the conceptual properties of the FSS and the CIS20-F.

In addition to clarifying the conceptual properties of the FSS and CIS20-F, their clinimetric properties on reliability need to be known for valid assessment of fatigue over time, or adequately to distinguish between patients with different levels of fatigue. The reliability of the FSS has been investigated in

 doi: 10.2340/16501977-1838

 ISSN 1650-1977

3 studies in PPS, and has shown acceptable to good internal consistency and test-retest reliability (20–22). The CIS20 has been validated in CFS (23), multiple sclerosis (24) and the working population (25). However, the clinimetric properties on reliability of the CIS20-F in PPS are unknown.

The aims of the current study were: (*i*) to perform a content comparison on the FSS and the CIS20-F questionnaires for measuring fatigue; and (*ii*) to assess the reliability (internal consistency, test-retest reliability and measurement error) of these questionnaires in polio survivors.

METHODS

Participants and procedures

A consecutive series of 80 polio survivors known to the outpatient clinic of Rehabilitation of the Academic Medical Centre in Amsterdam were invited to participate in this study. Written information on the study, a brief survey on demographic variables, 2 fatigue questionnaires (FSS and CIS20-F) and a stamped return envelope were sent out to these patients. To assess test-retest reliability and measurement error, a second package of questionnaires was sent (within 3 weeks) to participants who returned the first questionnaire.

According to Dutch law, studies involving human subjects must undergo medical ethics review if they are subject to the *Medical Research Involving Human Subjects Act* (WMO). A study involving the completion of a questionnaire does not fall within the scope of the Act and, therefore, approval from the medical ethics committee was not required for our study.

Fatigue questionnaires

The FSS (11) is a self-administered questionnaire measuring fatigue. It consists of 9 statements rated on a 7-point scale, ranging from 1 (strong disagreement) to 7 (strong agreement). The total score is calculated as the mean of the responses to the 9 statements. A score of 4 or higher indicates a moderate to high level of fatigue (26).

The CIS20 (13) is a 20-item questionnaire that contains 4 sub-scales (subjective experience of fatigue, reduction in concentration, reduction in motivation and reduction in physical activity) related to fatigue. For the purpose of this study, only the sub-scale CIS20-F was used. The CIS20-F consists of 8 statements, and participants have to indicate on a 7-point scale (ranging from 1 to 7) to what extent the particular statement applies to them. The total score is calculated as the sum of the responses to the 8 statements. A higher CIS20-F score indicates a higher degree of fatigue. Based on research with CFS patients, a score of 35 or higher indicates severe fatigue (23).

Content comparison of FSS with CIS20-F

The content of the FSS and the CIS20-F were compared by identifying, quantifying and comparing the concepts contained in both questionnaires, using the ICF (27) as a reference. This method of linking outcome measures to the ICF, based on standardized linking rules developed by Cieza et al. (28, 29), is widely used to explore the content of a measurement instrument. Following the linking rules, each meaningful concept (MC), i.e. a unit of text that conveys a single theme within an instrument item, was linked to the most precise corresponding ICF category, identified with its alphanumeric code that indicates the component of the ICF: Body Functions (b), Body Structures (s), Activities and Participation (d), Environmental Factors (e), and Personal Factors. Perceptions were not separately coded if they were inextricably bound up with other MCs (30). For example, in item 1 of the CIS20-F "I feel tired" *"feeling tired"* was considered as 1 MC. In item 6 of the CIS20-F "physically I feel I am in a bad condition" *"feeling"* and *"physical condition"* were considered different MCs, and therefore coded separately.

Two health professionals working in rehabilitation research (FK, MB) independently extracted and linked MCs contained in FSS and CIS20-F items. Disagreements (defined as the identification of different MCs or the assignment of different ICF categories) were resolved by a consensus procedure, if necessary, with a third author trained in the ICF linking rules (YH).

Analyses

Returned questionnaires were checked for completeness, and they were excluded from the analysis if one or more of the items were missing. Patient characteristics and mean scores of the FSS and CIS20-F were analysed with descriptive statistics.

Content analysis. For both questionnaires, the results of the linking process were presented at the item level (i.e. a presentation of MCs contained in the items of both questionnaires and the linked ICF categories) and at the ICF component level (i.e. a representation of the ICF components in both questionnaires). Furthermore, the content density and content diversity were determined for each questionnaire (31). The *content density* is a measure of multidimensionality within the item structure of an instrument and is calculated as the ratio of the total number of MCs and the total number of items in the questionnaire. Ratios close to 1 mean that each item contains 1 MC, while higher values show that there are several MCs contained within each item. The *content diversity*, which is a measure of reach or bandwidth of a questionnaire with regard to ICF categories covered, is the ratio of the number of different ICF codes and the number of MCs. Values close to 1 indicate that MCs of the measure correspond to a different ICF category. Values closer to zero indicate that several MCs in the questionnaire relate to the same ICF category.

Reliability of the FSS and CIS20-F. Internal consistency (32) was assessed using Cronbach's α coefficient and item-total correlation (ITC). An α coefficient > 0.70 is considered sufficient, an α coefficient > 0.80 is considered good for the purpose of group comparisons, and an ITC > 0.30 is considered acceptable (33).

Test-retest reliability was assessed with the intraclass correlation coefficient (ICC) and 95% confidence intervals (95% CI) of the ICC. A 2-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 (32). An ICC value > 0.70 is considered acceptable (33) and a lower boundary of the 95% CI of the ICC > 0.75 is considered to indicate excellent reliability (34). Systematic differences between the 2 assessments together with the 95% CI were analysed for the total scores on the FSS and the CIS20-F using paired *t*-tests.

Measurement error was expressed with the standard error of measurement (SEM), calculated as $\sqrt{(var_o + var_{res})}$, where var_o is the variance due to systematic differences between occasions and var_{res} is the random error variance (32). From the SEM, the smallest detectable change (SDC), i.e. the amount of change between 2 assessment scores that is reliably detectable above measurement error) (35) was derived, calculated as $1.96 \times \sqrt{2} \times SEM$ (32).

Analyses were carried out using the Statistical Program for Social Sciences (SPSS) for Windows, version 18. For all tests, a *p*-value < 0.05 was considered significant.

RESULTS

Content comparison of FSS with CIS20-F

The 2 authors (FK, MB) reached full consensus regarding identification of different MCs from the items of both questionnaires (30 MCs for both questionnaires together). There was disagreement in the assignment of ICF codes for 9 MCs (30%), and a third author (YH) determined the ICF code for these MCs.

Tables I and II summarize the results of the content comparison. The total number of identified MCs was 20 for the FSS and 10 for the CIS20-F. With a high number of MCs per item, the FSS had a higher content density ratio than the CIS20-F (2.2 and 1.25, respectively). Also, the content diversity ratio of the FSS was higher compared with the CIS20-F (0.45 and 0.30, respectively). The proportion of MCs that could be linked to an ICF category was 80% for both questionnaires. The MCs of the CIS20-F were linked to 3 different categories, all within the ICF component "Body Functions". Ten MCs (50%) of the FSS were linked to the ICF component "Body Functions" (to 3 different categories) and 6 MCs (30%) were linked to the ICF component "Activities and Participation" (to 6 different categories). The ICF components "Environmental Factors" and "Personal Factors" were not represented in the questionnaires.

MCs in both questionnaires were found to be linked to 2 different fatigue-related ICF categories: "b1300 Energy level", which is defined as "mental functions that produce vigour and stamina" and "b4552 Fatigability", defined as "exercise tolerance functions related to susceptibility to fatigue, at any level of exertion". Most items (6 out of 8) of the CIS20-F contained only 1 MC, which were all linked to 1 of the 2 fatigue-related ICF categories. Most items (8 out of 9) of the FSS contained

Table II. *Summarizing statistics of International Classification of Functioning, Disability and Health (ICF)-linking for Fatigue Severity Scale (FSS) and CIS20-F*

	FSS *n* (%)	CIS20-F *n* (%)
Scale items	9	8
MCs	20	10
Content density (MCs per item)	2.2	1.25
Different ICF codes	9	3
Content diversity (ICF codes per MC)	0.45	0.30
Representation of the ICF components		
Body functions (b) and Body structures (s) (MCs)	10 (50)	8 (80)
Activities and Participation (d) (MCs)	6 (30)	0 (0)
Environmental factors (e) (MCs)	0 (0)	0 (0)
Personal factors (MCs)	0 (0)	0 (0)
Not definable (MCs)	4 (20)	2 (20)

ICF: International Classification of Functioning Disability and Health; CIS20-F: sub-scale subjective experience of fatigue of the Checklist Individual Strength; FSS: Fatigue Severity Scale; MCs: number of meaningful concepts assigned to the items in the questionnaires (duplicates included); Content density: the number of MCs divided by the number of items; Different ICF codes: number of different ICF codes assigned to the MCs; Content diversity: the number of different ICF codes used divided by the number of MCs.

Table I. *International Classification of Functioning Disability and Health (ICF)-linking for Fatigue Severity Scale (FSS) and CIS20-F items*

	Meaningful concept	ICF category
FSS item		
1. My motivation is lower when I am fatigued	Motivation	b1301 Motivation
	Fatigued	b1300 Energy level
2. Exercise brings on my fatigue	Exercise	d Activities
	Fatigue	b4552 Fatiguability
3. I am easily fatigued	Fatigued	b4552 Fatiguability
4. Fatigue interferes with my physical functioning	Fatigue	b1300 Energy level
	Physical functioning	nd-ph
5. Fatigue causes frequent problems for me	Fatigue	b1300 Energy level
	Problems	nd
6. My fatigue prevents sustained physical functioning	Fatigue	b1300 Energy level
	Physical functioning	nd-ph
7. Fatigue interferes with carrying out certain duties and responsibilities	Fatigue	b1300 Energy level
	Carrying out duties	d230 Carrying out daily routine
	Carrying out responsibilities	d2400 Handling responsibilities
8. Fatigue is among my three most disabling symptoms	Fatigue	b1300 Energy level
	Disabling symptoms	nd-gh
9. Fatigue interferes with my work, family or social life	Fatigue	b1300 Energy level
	Work	d850 Remunerative employment
	Family life	d760 Family relationships
	Social life	d9 Community, social and civic life
CIS20-F item		
1. I feel tired	Feeling tired	b1300 Energy level
2. Physically I feel exhausted	Feeling physically exhausted	b1300 Energy level
3. I feel fit	Feeling fit	b1300 Energy level
4. I feel weak	Feeling weak	b1300 Energy level
5. I feel rested	Feeling rested	b1300 Energy level
6. Physically I feel I am in a bad condition	Feeling	b152 Emotional Functions
	Physical condition	nd-ph
7. I get tired very quickly	Getting tired	b4552 Fatiguability
8. Physically I feel in a good shape	Feeling	b152 Emotional Functions
	Physical shape	nd-ph

ICF: International Classification of Functioning Disability and Health; CIS20-F: sub-scale subjective experience of fatigue of the Checklist Individual Strength; FSS: Fatigue Severity Scale; nd: not definable; nd-gh: not definable – general health; nd-ph: not definable – physical health.

Table III. *Baseline characteristics (*n = *61)*

Patient characteristics	
Sex, male/female, *n* (%)	25 (41)/36 (59)
Age, years, mean (SD) [range]	59 (10) [31–83]
Age at acute polio, years, mean (SD) [range]	3.8 (3.4) [0–19]
Diagnosis of PPS according to the criteria of March of Dimes (36), *n* (%)	
Yes	50 (82)
No	11 (18)
Time since onset of PPS, years, mean (SD) [range]	14.2 (11.5) [1–60]

SD: standard deviation; PPS: post-poliomyelitis syndrome.

both MCs linked to 1 of the 2 fatigue-related ICF categories, as well as 1 or more MCs linked to non-fatigue-related ICF categories (within the ICF components "Body Functions" and "Activities and Participation"). For example, item 9 of the FSS contains the fatigue-related ICF category "b1300 Energy level" as well as the non-fatigue related ICF categories "d850 Remunerative employment", "d760 Family relationships" and "d9 Community, social and civic life".

Descriptive data of the FSS and CIS20-F

From the 80 questionnaires that were sent out, 61 were returned (76%). Baseline characteristics of these 61 patients are shown in Table III. Test-retest response rate was 90%. On the first and second assessment, respectively, 3 (5%) FSS questionnaires and 3 (5%) CIS20-F questionnaires and 1 (2%) FSS questionnaire and 2 (4%) CIS20-F questionnaires were excluded, due to missing values on 1 or more items. Data from 49 FSS questionnaires and 47 CIS20-F questionnaires were available for test-retest analyses.

The item scores on the first assessment are shown in Table IV. The mean (standard deviation (SD)) total fatigue score was 5.3 (1.3) on the FSS and 38 (12) on the CIS20-F. The highest item scores were seen on question 6 of the FSS ("My fatigue prevents sustained physical functioning": median score 6.5 (IQR 6.0–7.0)), and on question 7 of the CIS20-F ("I get tired very quickly"; median score 6.0 (IQR 5.0-7.0)). Highest possible total scores were found in 5% and 3% of the patients on the FSS and the CIS20-F, respectively. Lowest possible total scores were only seen on the CIS20-F, in 2% of the patients.

Table IV. *Item scores on the Fatigue Severity Scale (FSS) and CIS20-F on the first measurement*

	Median (IQR)
FSS (range 1–7) (*n* = 58)	
1. My motivation is lower when I am fatigued	6.0 (4.0–7.0)
2. Exercise brings on my fatigue	6.0 (4.0–6.3)
3. I am easily fatigued	6.0 (5.0–7.0)
4. Fatigue interferes with my physical functioning	6.0 (5.0–7.0)
5. Fatigue causes frequent problems for me	5.0 (4.0–7.0)
6. My fatigue prevents sustained physical functioning	6.5 (6.0–7.0)
7. Fatigue interferes with carrying out certain duties and responsibilities	6.0 (5.0–6.3)
8. Fatigue is among my 3 most disabling symptoms	6.0 (4.8–7.0)
9. Fatigue interferes with my work, family or social life	5.0 (4.0–6.0)
CIS20-F (range 1–7) (*n* = 58)	
1. I feel tired	5.0 (4.0–6.0)
2. Physically I feel exhausted	4.0 (2.0–5.3)
3. I feel fit[a]	5.0 (4.0–6.0)
4. I feel weak	4.0 (2.0–5.0)
5. I feel rested[a]	6.0 (4.0–6.0)
6. Physically I feel I am in a bad condition	4.5 (2.8–6.0)
7. I get tired very quickly	6.0 (5.0–7.0)
8. Physically I feel in a good shape[a]	6.0 (4.0–7.0)

[a]Items formulated in positive terms are re-coded.
ICF: International Classification of Functioning Disability and Health; CIS20-F: sub-scale subjective experience of fatigue of the Checklist Individual Strength; FSS: Fatigue Severity Scale; IQR: inter quartile range.

Reliability of FSS and CIS20-F

Both the FSS and the CIS20-F showed acceptable internal consistency (Cronbach's $\alpha = 0.90$ for the FSS on the first assessment with ITCs ranging from 0.38 to 0.82; Cronbach's $\alpha = 0.93$ for the CIS20-F on the first assessment with ITCs ranging from 0.70 to 0.87).

Table V shows the results regarding test-retest reliability and measurement error. Total fatigue scores on the FSS and the CIS20-F did not differ between the first and second assessment (mean difference: –0.04 (95% CI –0.26–0.19) for FSS and 1.2 (95% CI –0.55–2.93) for CIS20-F). The ICCs (95% CI) for the FSS and the CIS20-F were 0.80 (0.67–0.88) and 0.85 (0.75–0.91), respectively, indicating acceptable and excellent test-retest reliability. Measurement error, expressed by the SDC, was 1.55 (representing 28.7% of the mean of the 2 assessments) for the FSS and 11.6 (29.4% of the mean) for the CIS20-F.

Table V. *Test-retest reliability and measurement error for Fatigue Severity Scale (FSS) and CIS20-F*

	ICC [95% CI]	Mean T1 and T2 (SD)	Δ T2–T1 (SD) [95% CI]	SEM	SDC
FSS (range 1–7) (*n* = 49)	0.80 [0.67–0.88]	5.4 (1.2)	–0.04 (0.79) [–0.26–0.19]	0.56 (10.4%)	1.55 (28.7%)
CIS20-F (range 8–56) (*n* = 47)	0.85 [0.75–0.91]	39.5 (10.5)	1.2 (5.9) [–0.55–2.93]	4.23 (10.7%)	11.6 (29.4%)

T1: first assessment; T2: second assessment; ICC: intraclass correlation coefficient = $var_p/(var_p + var_o + var_{res})$, were var_p is the variance due to systematic difference between "true" scores of patients, var_o is the variance due to systematic differences between occasions, var_{res} is the random error variance; 95% CI: 95% confidence interval; SEM: standard error of measurement = $\sqrt{(var_o + var_{res})}$, were var_o is the variance due to systematic differences between occasions; var_{res} is the random error variance. The SEM is expressed in original scale points and as percentage of the mean of the 2 assessments; SDC: smallest detectable change = $1.96 \times \sqrt{2} \times$ SEM. The SDC is expressed in original scale points and as percentage of the mean of the 2 assessments; CIS20-F: sub-scale subjective experience of fatigue of the Checklist Individual Strength; FSS: Fatigue Severity Scale; SD: standard deviation.

DISCUSSION

This study shows that the FSS and CIS20-F differ in content with regard to the ICF categories they represent. Our results further show that the ability of the FSS and CIS20-F to discriminate between polio survivors with different levels of fatigue is acceptable (test-retest reliability), while the measurement error of both questionnaires is large. Clinicians and researchers should be aware of the differences in content metrics of these questionnaires and the insufficient sensitivity to detect real changes beyond measurement error in individual patients.

The results of the current study with respect to content metrics (i.e. content density and content diversity) and the percentages of ICF components covered in the FSS are comparable to those reported in the study of Gradinger (19). Yet, it is important to note that, in the Gradinger study, only the content metrics and the representation of different ICF components were reported and not the results of the linking process on item level, which prevents detailed insight into the construct of the questionnaire.

The lower content density and content diversity of the CIS20-F compared with the FSS indicates less multidimensionality within the item structure of the CIS20-F, with a smaller coverage of the questionnaire with respect to different ICF categories. This study has found that the CIS20-F focuses only on fatigue-related concepts (within the ICF component "Body Functions"), while the FSS measures both fatigue-related concepts and non-fatigue related concepts (within the ICF components "Body Functions" and "Activities and Participation"). Examining the items of the FSS and the CIS20-F in more detail and taking into account the differences in content density and content diversity, one might suggest that the CIS20-F quantifies the concept of fatigue, while the FSS seems mainly to measure the interference of fatigue with functioning. However, this conclusion should be regarded with caution, because information regarding the causal relationship contained in items was not systematically analysed. This limitation has been described by the developers of the ICF linking rules (29). Nonetheless, the method of ICF linking has been shown to be helpful in clarifying the conceptual differences between the FSS and the CIS20-F.

Regarding reliability, both the FSS and the CIS20-F show acceptable internal consistency and test-retest reliability, with excellent test-retest reliability for the CIS20-F, based on the lower 95% CI boundary of the ICC. For the FSS, our results are in accordance with those reported in previous studies in polio survivors (with Cronbach's α ranging from 0.80 to > 0.95 (20–22) and ICCs ranging from 0.83 to 0.97 (20, 22)). No previous data on reliability of the CIS20-F in polio survivors have been published, but the present results are comparable to reliability parameters found in CFS (Cronbach's $\alpha = 0.88$) (23) and multiple sclerosis (ICC = 0.84 (95% CI 0.72–0.91)) (24).

Despite acceptable internal consistency and test-retest reliability, the smallest detectable changes (SDC) found for both questionnaires were rather high. For the FSS this is in line with the results of the study by Horemans, who reported an SDC of 1.5 (27%) (20). Considering that changes in FSS scores following different therapeutic interventions in PPS were found to range from 0.6 to 2.7 (7–10) it seems that the FSS is insufficiently sensitive to detect changes beyond measurement error in single individuals. For applications at the group level, much smaller changes can be detected, since measurement error is reduced by $\sqrt{n}$ for a group of n patients (32). For both questionnaires we found highest and lowest possible scores in less than 15% of the patients, which means that floor and ceiling effects do not negatively influence the reliability of the questionnaires (32).

The large measurement error might, in part, result from day-to-day fluctuations in fatigue in polio survivors over time (20), although there are no studies available to support this hypothesis. For detecting changes in fatigue in individual patients with PPS, one might consider using multiple repeated measures over time in order to overcome the problem of large measurement error (32).

Generalization of our results to PPS patients is considered to be good, since we used a consecutive series of polio survivors known to an outpatient rehabilitation clinic. The level of fatigue in our population is comparable to that in earlier studies in polio survivors, with FSS scores ranging from 5.1 to 5.4 and corresponding SDs ranging from 1.1 to 1.5 (5, 20, 21). Furthermore, according to the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN), we used a fair sample size (37). The response rate of 76% found in our study is considered sufficient and in line with response rates in previous surveys in polio survivors (varying from 74% to 88%) (4, 38, 39).

Study limitations

This study has some limitations. First, we examined the conceptual properties of only 2 fatigue questionnaires. Future research should focus on investigating the conceptual properties of other fatigue instruments commonly used in polio research, so that a core set of instruments can be selected for use in research and clinical practice. The present study might serve as an example for this purpose. Secondly, despite the standardization of the ICF linking rules, this method is not simple and straightforward (28, 29). An evaluation of our linkage process showed that agreement between the 2 coders was 70%, which is in line with the results of other studies reporting this measure (varying from 64% to 79%) (19, 30). Extensive training of the ICF linking procedure might further improve the reliability of the method. Finally, although this study gave insight into the content validity and reliability of the FSS and CIS20-F in polio survivors, studies aimed at determining the construct validity are needed to gain more insight into the internal relationships of items within the instrument and the relationships with scores of other instruments or differences between relevant groups (for the CIS20-F). The results of our content analyses provide the basis for formulating hypotheses regarding the construct validity and facilitate the interpretation of the results of this process. Furthermore, the responsiveness of both questionnaires should also be studied to determine their adequateness to detect meaningful changes in fatigue over time (32).

Conclusion

This study shows that the FSS and CIS20-F differ in content with regard to the ICF categories they represent. It is suggested that the CIS20-F primarily quantifies the concept of fatigue, while the emphasis of the FSS is on measuring the interference of fatigue with functioning. Considering the acceptable internal consistency and test-retest reliability, as well as the potential of the 2 instruments to detect changes at group level, we conclude that both the FSS and the CIS20-F can be applied in PPS research when assessing fatigue, although they cannot be used interchangeably.

REFERENCES

1. Kalkman JS, Zwarts MJ, Schillings ML, van Engelen BG, Bleijenberg G. Different types of fatigue in patients with facioscapulohumeral dystrophy, myotonic dystrophy and HMSN-I. Experienced fatigue and physiological fatigue. Neurol Sci 2008; 29 Suppl 2: S238–S240.
2. Halstead LS, Rossi CD. New problems in old polio patients: results of a survey of 539 polio survivors. Orthopedics 1985; 8: 845–850.
3. Nollet F, Beelen A, Prins MH, Visser M.de, Sargeant AJ, Lankhorst GJ, et al. Disability and functional assessment in former polio patients with and without postpolio syndrome. Arch Phys Med Rehabil 1999; 80: 136–143.
4. Schanke AK, Stanghelle JK. Fatigue in polio survivors. Spinal Cord 2001; 39: 243–251.
5. Tersteeg IM, Koopman FS, Stolwijk-Swuste JM, Beelen A, Nollet F. A 5-year longitudinal study of fatigue in patients with late-onset sequelae of poliomyelitis. Arch Phys Med Rehabil 2011; 92: 899–904.
6. On AY, Oncu J, Atamaz F, Durmaz B. Impact of post-polio-related fatigue on quality of life. J Rehabil Med 2006; 38: 329–332.
7. Farbu E, Rekand T, Vik-Mo E, Lygren H, Gilhus NE, Aarli JA. Post-polio syndrome patients treated with intravenous immunoglobulin: a double-blinded randomized controlled pilot study. Eur J Neurol 2007; 14: 60–65.
8. Horemans HL, Nollet F, Beelen A, Drost G, Stegeman DF, Zwarts MJ, et al. Pyridostigmine in postpolio syndrome: no decline in fatigue and limited functional improvement. J Neurol Neurosurg Psychiatry 2003; 74: 1655–1661.
9. On AY, Oncu J, Uludag B, Ertekin C. Effects of lamotrigine on the symptoms and life qualities of patients with post polio syndrome: a randomized, controlled study. NeuroRehabilitation 2005; 20: 245–251.
10. Strumse YA, Stanghelle JK, Utne L, Ahlvin P, Svendsby EK. Treatment of patients with postpolio syndrome in a warm climate. Disabil Rehabil 2003; 25: 77–84.
11. Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. Arch Neurol 1989; 46: 1121–1123.
12. Koopman FS, Beelen A, Gerrits KH, Bleijenberg G, Abma TA, de Visser, et al. Exercise therapy and cognitive behavioural therapy to improve fatigue, daily activity performance and quality of life in postpoliomyelitis syndrome: the protocol of the FACTS-2-PPS trial. BMC Neurol 2010; 10: 8.
13. Vercoulen JHHM, Alberts M, Bleijenberg G. [The Checklist Individual Strength (CIS).] Gedragstherapie 1999; 32: 131–136 (in Dutch).
14. Reeves WC, Lloyd A, Vernon SD, Klimas N, Jason LA, Bleijenberg G, et al. Identification of ambiguities in the 1994 chronic fatigue syndrome research case definition and recommendations for resolution. BMC Health Serv Res 2003; 31: 25.
15. Gielissen MF, Verhagen S, Witjes F, Bleijenberg G. Effects of cognitive behavior therapy in severely fatigued disease-free cancer patients compared with patients waiting for cognitive behavior therapy: a randomized controlled trial. J Clin Oncol 2006; 24: 4882–4887.
16. Prins JB, Bleijenberg G, Bazelmans E, Elving LD, de Boo TM, Severens JL, et al. Cognitive behaviour therapy for chronic fatigue syndrome: a multicentre randomised controlled trial. Lancet 2001; 357: 841–847.
17. Kalkman JS, Schillings ML, van der Werf SP, Padberg GW, Zwarts MJ, van Engelen BG, et al. Experienced fatigue in facioscapulohumeral dystrophy, myotonic dystrophy, and HMSN-I. J Neurol Neurosurg Psychiatry 2005; 76: 1406–1409.
18. Voet NB, Bleijenberg G, Padberg GW, van Engelen BG, Geurts AC. Effect of aerobic exercise training and cognitive behavioural therapy on reduction of chronic fatigue in patients with facioscapulohumeral dystrophy: protocol of the FACTS-2-FSHD trial. BMC Neurol 2010; 10: 56.
19. Gradinger F, Glassel A, Bentley A, Stucki A. Content comparison of 115 health status measures in sleep medicine using the International Classification of Functioning, Disability and Health (ICF) as a reference. Sleep Med Rev 2011; 15: 33–40.
20. Horemans HL, Nollet F, Beelen A, Lankhorst GJ. A comparison of 4 questionnaires to measure fatigue in postpoliomyelitis syndrome. Arch Phys Med Rehabil 2004; 85: 392–398.
21. Burger H, Franchignoni F, Puzic N, Giordano A. Psychometric properties of the Fatigue Severity Scale in polio survivors. Int J Rehabil Res 2010; 33: 290–297.
22. Oncu J, Atamaz F, Durmaz B, On A. Psychometric properties of fatigue severity and fatigue impact scales in postpolio patients. Int J Rehabil Res 2013; 36: 339–345.
23. Vercoulen JH, Swanink CM, Fennis JF, Galama JM, van der Meer JW, Bleijenberg G. Dimensional assessment of chronic fatigue syndrome. J Psychosom Res 1994; 38: 383–392.
24. Rietberg MB, Van Wegen EE, Kwakkel G. Measuring fatigue in patients with multiple sclerosis: reproducibility, responsiveness and concurrent validity of three Dutch self-report questionnaires. Disabil Rehabil 2010; 32: 1870–1876.
25. Beurskens AJ, Bultmann U, Kant I, Vercoulen JH, Bleijenberg G, Swaen GM. Fatigue among working people: validity of a questionnaire measure. Occup Environ Med 2000; 57: 353–357.
26. Valko PO, Bassetti CL, Bloch KE, Held U, Baumann CR. Validation of the fatigue severity scale in a Swiss cohort. Sleep 2008; 31: 1601–1607.
27. World Health Organization (WHO). International Classification of Functioning, Disability and Health (ICF). Geneva: WHO; 2001.
28. Cieza A, Brockow T, Ewert T, Amman E, Kollerits B, Chatterji S, et al. Linking health-status measurements to the international classification of functioning, disability and health. J Rehabil Med 2002; 34: 205–210.
29. Cieza A, Geyh S, Chatterji S, Kostanjsek N, Ustun B, Stucki G. ICF linking rules: an update based on lessons learned. J Rehabil Med 2005; 37: 212–218.
30. van der Mei SF, Dijkers MP, Heerkens YF. Participation as an outcome measure in psychosocial oncology: content of cancer-specific health-related quality of life instruments. Qual Life Res 2011; 20: 1617–1627.
31. Geyh S, Cieza A, Kollerits B, Grimby G, Stucki G. Content comparison of health-related quality of life measures used in stroke based on the international classification of functioning, disability and health (ICF): a systematic review. Qual Life Res 2007; 16: 833–851.
32. Vet de HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine. Cambridge: University Press; 2011.
33. Nunnally JC, Bernstein IH. Psychometric theory. 3th ed. New York: McGraw-Hill; 1994.
34. Lee J, Koh D, Ong CN. Statistical evaluation of agreement between

two methods for measuring a quantitative variable. Comput Biol Med 1989; 19: 61–70.
35. Roebroeck ME, Harlaar J, Lankhorst GJ. The application of generalizability theory to reliability assessment: an illustration using isometric force measurements. Phys Ther 1993; 73: 386–395.
36. March of Dimes Birth Defects Foundation. Identifying best practices in diagnosis & care. Proceedings of the International conference on Post Polio Syndrome; 2000, May 19–20; Warm Springs, GA. New York, USA: March of Dimes; 2001. Available from: http://www.polioplace.org/sites/default/files/files/MOD-%20Identifying.pdf.
37. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res 2012; 21: 651–657.
38. Ivanyi B, Nollet F, Redekop WK, Haan R de, Wohlgemuht M, Van Wijngaarden JK, et al. Late onset polio sequelae: disabilities and handicaps in a population-based cohort of the 1956 poliomyelitis outbreak in the Netherlands. Arch Phys Med Rehabil 1999; 80: 687–690.
39. Bickerstaffe A, Beelen A, Nollet F. Circumstances and consequences of falls in polio survivors. J Rehabil Med 2010; 42: 908–915.