Test–Retest Reliability and Convergent Validity of the Fatigue Impact Scale for Persons With Multiple Sclerosis

Virgil Mathiowetz

KEY WORDS
• energy conservation
• fatigue assessment
• rehabilitation

OBJECTIVE. The test–retest reliability and the convergent validity of the Fatigue Impact Scale (FIS) were evaluated using secondary data from 54 persons with multiple sclerosis (MS).

METHODS. This reliability and validity study used FIS data from before and after two control periods to evaluate test–retest reliability. Convergent validity of the FIS with the Fatigue Severity Scale and with subscales of the SF-36 Health Survey was evaluated using data collected before the first control period.

RESULTS. No significant differences between before and after FIS measurements and intraclass correlation coefficients ranging from .68 to .85 indicate that the FIS has good test–retest reliability except for the physical subscale. The expected moderate correlations between the FIS and several subscales of the SF-36 support its convergent validity. In contrast, the unexpected low correlation between the FIS and Fatigue Severity Scale does not support convergent validity.

CONCLUSION. The FIS has adequate reliability and validity and is recommended to evaluate the effectiveness of fatigue management interventions such as energy conservation education for persons with MS.

Mathiowetz, V. (2003). Test–retest reliability and convergent validity of the Fatigue Impact Scale for persons with multiple sclerosis. *American Journal of Occupational Therapy, 57,* 389–395.

Virgil Mathiowetz, PhD, OTR, FAOTA, is Associate Professor, Program in Occupational Therapy, Department of Physical Medicine & Rehabilitation, University of Minnesota, MMC 388, 420 Delaware Street SE, Minneapolis, Minnesota 55455; mathi003@umn.edu

Fatigue is widely recognized as the most common symptom for individuals with multiple sclerosis (MS). For example, Krupp, Alvarez, LaRocca, and Scheinberg (1988) found that 87% of individuals with MS reported fatigue to be a problem and that 28% described it as their most troubling symptom. The Multiple Sclerosis Council for Clinical Practice Guidelines (MSCCPG, 1998) defined fatigue as, "A subjective lack of physical and/or mental energy that is perceived by the individual or caregiver to interfere with usual and desired activities" (p. 2). They suggested that the Fatigue Impact Scale (FIS) (Fisk, Pontefract, Ritvo, Archibald, & Murray, 1994a) was "most appropriate for assessing the impact of MS-related fatigue on quality of life" (MSCCPG, 1998, p. 2). However, there is a lack of test–retest reliability and limited validity data for this instrument. The purpose of this study was to evaluate the test–retest reliability and the convergent validity of the FIS using secondary data from a study by Mathiowetz, Matuska, and Murphy (2001).

Multiple Sclerosis Fatigue

Persons with MS have described fatigue as a frustrating and overwhelming symptom, which can be disabling (McLaughlin & Zeeberg, 1993). Krupp et al. (1988) provided evidence that MS fatigue has unique characteristics that include: (1) occurs more frequently and is more severe than normal fatigue; (2) prevents sustained physical functioning; (3) comes on quickly and recovery from it takes much longer than normal fatigue; (4) exacerbates other MS symptoms; (5) is worsened by heat; (6) is chronic; and (7) its severity is not always related to neurologic status or other MS symptoms. Thus, MS fatigue is likely to interfere with the per-

formance of everyday activities. One of the most common instruments used to measure MS fatigue is the FIS.

Fatigue Impact Scale (FIS)

The FIS was developed "to evaluate the perceived impact of fatigue on the lives of MS patients, the factors that affect patients' perceptions of fatigue impact, and how fatigue may affect the mental health and general health status of MS patients" (Fisk et al., 1994a, p. 10). The FIS consists of 40 statements that measure fatigue in three areas: physical, cognitive, and social. In addition to the FIS: Total score, Physical, Cognitive, and Social subscale scores can be calculated. Respondents rate the statements on a Likert scale ranging from 0 (no problem) to 4 (extreme problem). Unfortunately there are no reports on FIS test–retest reliability or stability over time. Internal consistency of the FIS and subscales was high to very high (i.e., Cronbach's alpha .88 to .98) (Fisk et al., 1994b). The FIS was moderately correlated (r = .53) with the Sickness Impact Profile providing some evidence of convergent validity (i.e., FIS is related to a similar construct as expected). The FIS discriminated between MS patients and hypertensive patients whose primary complaint was not fatigue. In addition, patients with chronic fatigue syndrome had significantly higher FIS scores than MS patients as was predicted (Fisk et al., 1994b). These results support the discriminant validity of the FIS. The fact that there were significant changes in FIS scores as a result of a therapeutic intervention (i.e., energy conservation course) (Mathiowetz et al., 2001) provides evidence of the sensitivity of the FIS.

The FIS is more relevant to occupational therapy settings than other fatigue assessments such as the Fatigue Severity Scale (Krupp, LaRocca, Muir-Nash, & Steinberg, 1989) or Fatigue Assessment Instrument (Schwartz, Jandorf, & Krupp, 1993), because the items included in the FIS measure the construct, the impact of fatigue on the performance of everyday activities. As occupational therapists, we are more concerned about this construct (i.e., how an impairment such as fatigue influences functional performance) than about the severity of fatigue itself. The FIS would be a useful assessment to measure clinical outcomes of occupational therapy interventions designed to reduce the impact of fatigue. However, the lack of test–retest reliability data and the limited data on convergent validity of the FIS limit its usefulness.

Purpose and Significance

This study evaluated the test–retest reliability or stability of the Fatigue Impact Scale using secondary data from a study (Mathiowetz et al., 2001) in which the efficacy of an energy conservation course for persons with MS was evaluated using a one-group, repeated measures design. The FIS was used to measure the impact of fatigue before and after a placebo control (i.e., support group) condition, an experimental (i.e., energy conservation course) condition, and a no intervention control condition. The two control conditions provided test–retest reliability data on the FIS for this study because there were no significant changes in the FIS scores during these two conditions as was hypothesized.

In addition, this study explored the convergent validity of the FIS by comparing it to the Fatigue Severity Scale, an alternative measure of fatigue, and to the subscales of the Short Form-36 (SF-36) Health Survey (Ware, Snow, Kosinski, & Gandek, 1993), a quality of life measure, that were also used in the Mathiowetz et al. (2001) study. In this study, the FIS was hypothesized to have moderate correlations with the Fatigue Severity Scale because the two scales measure similar but not identical aspects of fatigue. Based on the correlations between subscales of the SF-36 and the Modified Fatigue Impact Scale (a shortened and modified version of the FIS) (Ritvo et al., 1997b), it was hypothesized that the FIS would correlate moderately with SF-36 subscales: vitality (i.e., feels tired and worn out all of the time), social functioning (i.e., frequent interference with normal social activities), role-emotional (i.e., problems with daily activities as a result of emotional problems), role-physical (i.e., problems with daily activities as a result of physical health), and mental health (i.e., feelings of nervousness and depression) subscales, and the mental component summary score.

Methods

Design

In the Mathiowetz et al. (2001) study, the FIS was administered four times: during weeks #1, #7, #13, and #19 (see diagram below). This study evaluated test–retest reliability of the FIS using data from before and after the placebo (weeks #1 & #7) and no intervention (weeks #13 & #19) control conditions, which were analyzed separately. Convergent validity of the FIS with the Fatigue Severity Scale and with subscales of the SF-36 was evaluated using data collected before the placebo control condition (week #1).

Interventions	Placebo Control	Experimental	No Intervention Control	
(one group, repeated measures)				
Assessments during Weeks:	1	7	13	19

Test–Retest Reliability Periods

Placebo Control Condition. The placebo-control condition consisted of 6 weekly, 2-hour support group sessions involving education on and discussion of topics that are commonly addressed in support groups for individuals with MS and other chronic diseases. The group generated a list of prioritized topics of interest and the occupational therapist that led the group gathered educational materials, videos, and resources as a basis for discussions. Topics included basic information on MS, medications, financial issues, estate planning, the Americans With Disabilities Act, reasonable accommodation, dealing with others' expectations, nutrition, exercise, memory problems, hiring an aide, and community resources. This condition was considered a placebo control because most aspects of it were similar to the experimental condition except that issues related to energy conservation or fatigue management were not discussed.

No Intervention Control. There was no intervention in the 6 weeks following the experimental (i.e., energy conservation course) condition to determine if the benefits of the course were maintained over that time period. Because there was no intervention during this control period, it was considered a slightly better estimate of the test–retest reliability of the FIS than the placebo-control condition. Both control periods were 6 weeks in length to be consistent with the 6-week length of the energy conservation course. Although these were relatively long test–retest reliability periods, the length was acceptable because fatigue impact is a construct that is expected to change slowly and the FIS measures fatigue impact over the prior 4 weeks.

Participants

Participants were recruited through a mailing by the Minnesota Chapter of the National MS Society to its members in a large metropolitan area. Interested candidates contacted the project director who did an informal screening on the phone. Potential participants who appeared to meet the inclusion criteria were invited to a formal screening session. To be included in the study, participants had a diagnosis of MS, were18 years of age or older, were functionally literate (i.e., able to read course materials), had a Fatigue Severity Scale (Krupp et al., 1989) score of four or greater (i.e., moderate to high fatigue severity), lived in the community, and were independent in the majority of self-care and daily activities. Fifty-four participants completed the study (i.e., met all the inclusion criteria throughout the 19 weeks of the study). Participants were excluded if they did not attend at least five out of six support group and energy conservation sessions, experienced an exacerbation of MS symptoms, had fatigue medication changes, or had other major illnesses, hospitalizations, or rehabilitation during the course of the study.

Table 1. Demographic and MS-related Characteristics of Participants (*N* = 54)

Characteristics	*n*	(%)
Gender		
Male	18	(33%)
Female	36	(67%)
Type of Multiple Sclerosis		
Chronic progressive	12	(22%)
Relapsing or remitting	20	(36%)
Exacerbating or remitting	7	(13%)
Benign	4	(7%)
Unknown	12	(22%)
Employment Status		
Full time outside home	22	(41%)
Part-time outside home	16	(30%)
Disability status	6	(11%)
Fulltime homemakers	4	(7%)
Retired	3	(5.5%)
Unemployed	3	(5.5%)
Other Variables	*M*	range
Age	50 years	31–74 years
Years diagnosed	9.5 years	1–34 years
FSS score	5.55	4–7

Note. FSS = Fatigue Severity Scale. From "Efficacy of an Energy Conservation Course for Persons with Multiple Sclerosis" by V. Mathiowetz, K. M. Matuska, and M. E. Murphy, 2001, *Archives of Physical Medicine & Rehabilitation, 82*, p. 451. Copyright 2001 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine & Rehabilitation. Reprinted with permission of the author.

Demographic and MS-related information on the 54 participants who completed the study are shown in Table 1. Eighty-three percent of the participants reported fatigue as one of their primary symptoms.

Assessment Instruments

FIS (Fisk et al., 1994a, 1994b). This assessment, described above, was used to measure the impact of fatigue on participants' lives before and after the 6-week control conditions.

Fatigue Severity Scale. Krupp et al. (1989) developed this scale to assess disabling fatigue in MS and systemic lupus erythematosus. It consists of nine statements, which patients rate on a Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). Fatigue Severity Scale scores of 11 patients not being treated for fatigue were not significantly different before and after an average 10-week period and were correlated highly (r = .84), providing some support for the test–retest reliability of the instrument (Krupp et al., 1989). Internal consistency was high (i.e., Cronbach's alpha ranged from .81 to .89). Fatigue Severity Scale scores correlated r = .47 with a Visual Analogue Scale of fatigue, providing limited evidence of convergent validity. Fatigue Severity Scale scores distinguished the two patient groups from the control group but not from each other. This provided some evidence of discriminant validity. In patients

being treated for fatigue (n = 8), clinical improvement in fatigue was associated with lower Fatigue Severity Scale scores, providing some support for the sensitivity of the instrument. This assessment was used as a screening instrument in the Mathiowetz et al. (2001) study to determine if participants had moderate to high fatigue severity.

SF-36 Health Survey (Ware, Snow, Kosinski, & Gandek, 1993). This assessment was used to measure the participants' perceived quality of life. The SF-36 is considered a generic measure of health-related quality of life because it represents values that are relevant to the functional status and well being of persons with various diagnoses. It consists of eight subscales: physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional, and mental health. In addition, two component summary subscales: physical and mental can be computed. The internal consistency of the scales ranged from .80–.92 for patients with chronic conditions. Test–retest reliability of the scales over a 2-week interval ranged from .60–.81 for general practice patients. This level of reliability is not considered acceptable for comparisons of individual patients but is considered acceptable for group-level analyses (Ware et al., 1993). Numerous studies support the validity of the SF-36. For example, the Vitality Scale correlates moderately (r = -.68) with the energy scale of the Nottingham Health Profile (Ware et al., 1993). The SF-36 health survey is used widely in clinical outcome studies.

Procedures

After screening, each participant who met the inclusion criteria was given the FIS and SF-36 Health Survey. All eligible participants were assigned to groups of 8 to 10, based on time preference of participants. One week after the screening session, all groups began their placebo-control period of 6 weekly, 2-hour support group sessions. At the end of the sixth support group session, participants completed a second FIS. One week after that, each group began the experimental intervention, the 6 weekly, 2-hour energy conservation course. At the end of the sixth experimental session, participants completed the FIS assessments for the third time. There was no further contact with participants for the next 6 weeks (no intervention control period). At the end of the sixth week, most participants returned to complete their fourth and final FIS. Those participants unable to attend this session completed their final assessments at home and returned them through the mail. The University's Institutional Review Board approved the study for human subjects.

Data Analysis

Test–Retest Reliability. The stability of the FIS: Total and subscale scores were evaluated using paired-data t tests and intraclass correlation coefficients (ICC)(3,1) (Shrout & Fleiss, 1979). Currently, the ICC is the preferred measure of test–retest reliability because it reflects both correlation and agreement in one index. Nonsignificant (p < .05) differences between test and retest occasions and an ICC of .75 or higher indicate good reliability, .50–.75 moderate reliability, and .50 and below poor reliability (Portney & Watkins, 2000).

Convergent Validity. Pearson product-moment correlation coefficients were calculated to correlate the FIS with the Fatigue Severity Scale and subscales of the SF-36. High (r = .70–.89) or very high (r = .90–1.00) correlations were expected between assessments that appear to measure the same constructs; moderate (r = .50–.69) correlations (Munro, 1997) were expected between assessments that measure similar or related constructs. Munro described another way of determining the meaningfulness of r by using the coefficient of determination (r^2), which is the amount of shared variance or overlap between two variables. Thus, if two tests are correlated r = .50, then r^2 = .25, which means the two tests have 25% shared variance between them.

Results

Test–Retest Reliability

Descriptive data on the three subscales of the FIS and the FIS: Total before and after the two control periods are reported in Tables 2 and 3. In all cases, FIS scores decreased slightly between the before and after scores, reflecting a slight decrease in fatigue impact and a potential practice or learning effect. However, paired-data t tests indicated that there were no significant differences between the before and after scores. These results support the stability of the FIS scores over the 6-week control periods. In all cases, ICCs(3,1) were higher for the no intervention-control period than for the placebo-control period. In both control periods, the physical subscales of the FIS had the lowest

Table 2. Test–Retest Reliability: Descriptive Data, *t* tests, and Intraclass Correlation Coefficients (ICC)(3,1) for the Three Subscales and Total Score of the Fatigue Impact Scale (FIS) (*N* = 54) Before and After the 6-Week Placebo Control Period (i.e., Support Group)

FIS Subscales & Total Score	Before		After				
	M	*(SD)*	*M*	*(SD)*	*t* test	*p*	ICC
FIS: Cognitive	14.7	(9.3)	14.0	(8.6)	.92	.361 *ns*	.76
FIS: Physical	22.0	(7.7)	20.7	(7.3)	1.52	.134 *ns*	.68
FIS: Social	32.2	(14.0)	31.8	(14.5)	.29	.771 *ns*	.72
FIS: Total	68.9	(26.2)	66.4	(26.5)	.98	.334 *ns*	.76

Note. Lower scores reflect decreased impact of fatigue.
ns = not significant at p < .05

Table 3. Test–Retest Reliability: Descriptive Data, *t* tests, and Intraclass Correlation Coefficients (ICC)(3,1) for the Three Subscales and Total Score of the Fatigue Impact Scale (FIS) (*N* = 54) Before and After the 6-Week No Intervention-Control Period

FIS Subscales & Total Score	Before *M*	Before *(SD)*	After *M*	After *(SD)*	*t* test	*p*	ICC
FIS: Cognitive	12.0	(9.1)	11.9	(8.8)	.11	.913 *ns*	.85
FIS: Physical	17.8	(8.8)	17.3	(8.2)	.49	.628 *ns*	.69
FIS: Social	26.1	(15.5)	25.2	(14.0)	.74	.446 *ns*	.83
FIS: Total	55.8	(29.7)	54.5	(27.3)	.58	.564 *ns*	.81

Note. Lower scores reflect decreased impact of fatigue.
ns = not significant at $p < .05$

ICCs, which were slightly lower than .75 desired. In contrast, the FIS: Cognitive and FIS: Social subscales, and FIS: Total had the highest ICCs. The latter demonstrated good test–retest reliability.

Convergent Validity

The correlations between the FIS, Fatigue Severity Scale, and subscales of SF-36 are displayed in Table 4. The FIS: Total score had a low correlation ($r = .44$) with the Fatigue Severity Scale, which meant that they share only 19% of their variance. The FIS: Total score had moderate correlations ($r = -.54$–$.62$) with the vitality, social functioning, and mental health subscales of the SF-36 and the mental component summary of the SF-36. The FIS: Cognitive subscale had moderate correlations with the mental health and mental component summary subscales of the SF-36. The FIS: Physical subscale had moderate correlations with the physical functioning and physical component summary subscales of the SF-36. Finally, the FIS: Social subscale had a moderate correlation with the vitality, social functioning, mental health, and mental component summary subscales of the SF-36. These moderate correlations between similar subscales of the FIS and SF-36 were expected and provided evidence of convergent validity. In contrast, the relatively low correlation ($r = .44$) between the FIS and Fatigue Severity Scale was lower than expected.

Discussion

Test–Retest Reliability

The cognitive and social subscales of the FIS, and the FIS: Total score demonstrated good test–retest reliability. The physical subscale of the FIS had slightly lower than desired stability and should be used with more caution. It is possible that the relatively long test–retest time periods used in this study combined with the fluctuating nature of MS may have affected the reliability data. It is speculated that 2- or 4-week control periods might have resulted in higher ICCs than the relatively long, 6-week control periods used in this study. Most assessment instruments have higher test–retest reliability when shorter test–retest periods are used. This may be especially true given the fluctuating nature of the disease process in MS. The no intervention-control period had higher test–retest reliability as expected and was considered the better time period for evaluating test–retest reliability because there was no intervention during that time.

The test–retest reliability of the FIS excluding the physical subscale is comparable to the Fatigue Severity Scale ($r = .84$). However, it should be noted that the reliability data on the Fatigue Severity Scale were based on a very small sample ($N = 11$) (Krupp et al., 1989). FIS test–retest reliability is clearly stronger than the Fatigue Assessment Instrument (Schwartz, Jandorf, & Krupp, 1993), which has lower than desirable test–retest reliability ($r = .29$-$.65$). The Modified Fatigue Impact Scale (Ritvo et al., 1997a; 1997b) and its three subscales have no reports of test–retest reliabil-

Table 4. Convergent Validity: Pearson Correlations Between the Fatigue Impact Scale (FIS), Fatigue Severity Scale (FSS), and the Short Form-36 (SF-36) Subscales and Summary Scales of Persons with Multiple Sclerosis (*N* = 52[a])

	FIS Cogn	FIS Phys	FIS Soc	FIS Total	FSS	SF-36 PhysF	SF-36 Role-P	SF-36 Pain	SF-36 GenH	SF-36 Vitality	SF-36 SocF	SF-36 Role-E	SF-36 MenH	SF-36 PCS	SF-36 MCS
FIS: Cognitive	-	.44**	.68**	.84**	.27*	.01	-.12	-.22	-.31*	-.34*	-.32*	-.33*	-.50**	-.02	-.51**
FIS: Physical		-	.55**	.73**	.34*	-.57**	-.20	-.42**	-.38**	-.48**	-.33*	-.17	-.30*	-.53**	-.20
FIS: Social			-	.93**	.46**	-.22	-.42**	-.39**	-.47**	-.55**	-.62**	-.25	-.68**	-.34*	-.56**
FIS: Total				-	.44**	-.28*	-.32*	-.41**	-.47**	-.55**	-.54**	-.30*	-.62**	-.34*	-.54**
FSS					-	-.28*	-.08	.04	-.26	-.36**	-.29*	-.09	-.22	-.16	-.22
SF-36: Physical Function						-	.08	.25	.20	.35*	.30*	.04	.10	.65**	-.07
SF-36: Role-Physical							-	.38**	.27	.38**	.38**	.18	.25	.50**	.22
SF-36: Bodily Pain								-	.40**	.49**	.22	.10	.18	.70**	.08
SF-36: General Health									-	.53**	.53**	.09	.38**	.55**	.33*
SF-36: Vitality										-	.42**	.21	.48**	.51**	.45**
SF-36: Social Function											-	.07	.58**	.38**	.51**
SF-36: Role-Emotional												-	.48**	-.22	.76**
SF-36: Mental Health													-	-.02	.87**
SF-36: Physical Component Summary (PCS)														-	-.22
SF-36: Mental Component Summary (MCS)															-

[a]Number of participants that completed the SF-36.
*$p < .05$; **$p < .01$

ity, which is a limitation of this instrument. In conclusion, test–retest reliability data for the FIS is comparable to or stronger than other fatigue assessment scales.

Convergent Validity

The unexpected low correlation between the FIS and Fatigue Severity Scale suggests that fatigue impact and fatigue severity are somewhat different constructs, indicating that the two assessments measure different aspects of fatigue and should not be used interchangeably. Therapists selecting a measure of fatigue need to consider what aspects of fatigue they want to measure (i.e., severity of fatigue vs. impact of fatigue). For most occupational therapy purposes, the impact of fatigue on everyday life would be the most relevant construct to measure.

Most of the data support the convergent validity of the FIS given the many moderate correlations between the FIS and similar subscales of the SF-36. Higher correlations were not expected because these instruments do not measure identical constructs. Most of the moderate correlations were similar to the correlations between the Modified Fatigue Impact Scale and subscales of the SF-36. This was expected given that the Modified Fatigue Impact Scale is a shortened form of the FIS. The low correlations between the FIS: Total score and Role-Physical & Role-Emotional subscales of the SF-36 were exceptions to this. It is possible that FIS questions that had low correlations with these subscales of the SF-36 were eliminated when the FIS was shortened to make the Modified Fatigue Impact Scale. If so, the correlations between the Modified Fatigue Impact Scale and these subscales of the SF-36 would be higher.

Limitations

The fact that secondary data were used in this study may have influenced the results of this study. The use of secondary data meant that variables other than time (e.g., the support group used in the placebo-control condition) could have influenced the test–retest reliability results. However, the fact that there were nonsignificant differences between FIS scores before and after the control periods (Mathiowetz et al., 2001) suggests other potential variables did not affect the FIS scores. Thus, the data were usable but not ideal for a test–retest study. In an ideal study of the test–retest reliability of the FIS, there would have been no other interventions happening and the length between test and retest would have been shortened to 4 weeks. The fact that persons with MS and severe disability were excluded from the original study means that the results of this study should be generalized to persons with MS and mild–moderate disability only.

Recommendations for Future Research

Fatigue is a common disabling symptom of many other medical conditions such as arthritis, fibromyalgia, cancer, chronic fatigue syndrome, and post-polio syndrome. Research on the reliability and validity of the FIS with these other medical conditions would broaden the usefulness of the FIS.

The FIS was a relatively lengthy assessment instrument for participants to complete (10–15 minutes) and for the researcher to score (5 minutes). A lengthy assessment instrument itself can create fatigue problems for some persons with MS and given the time constraints that therapists experience in clinical practice, a shorter version of the FIS would be desirable. Research to develop a shorter version of the FIS is needed. One strategy would be to further evaluate the psychometric properties of the Modified Fatigue Impact Scale (Ritvo et al., 1997a, 1997b), an existing shorter version of the FIS, by evaluating its test–retest reliability, convergent validity with the FIS, and discriminant validity. A second strategy would be to develop a new version of the FIS that might be even shorter than the Modified Fatigue Impact Scale and would retain the original FIS Likert scaling. It is not clear whether the modified scaling used in the Modified Fatigue Impact Scale affects the reliability and validity of the instrument. Thus, there would be some advantage to retaining the original FIS scaling.

Conclusion

Considering the results of this study and previous studies, the test–retest reliability and convergent validity of the FIS is adequate for research and clinical use for persons with mild–moderate multiple sclerosis. Until a shorter version of the FIS is available, the FIS is recommended for clinical use because it has stronger psychometric properties than other fatigue assessment instruments. If clinicians have clients complete the FIS before or after therapy time, the 10–15 minute completion time is not a limitation. Therapists can also save time by computing the FIS: Total score only. Computing the three subscales makes scoring more complex and time consuming and is not essential for most clinical purposes. The FIS is useful to evaluate the effectiveness of interventions used to manage fatigue such as energy conservation education.▲

References

Fisk, J. D., Pontefract, A., Ritvo, P. G., Archibald, C. J., & Murray, T. J. (1994a). The impact of fatigue on patients with multiple sclerosis. *Canadian Journal of Neurological Sciences, 21*(1), 9–14.

Fisk, J. D., Ritvo, P. G., Ross, L., Haase, D. A., Marrie, T. J., & Schlech, W. F. (1994b). Measuring the functional impact of fatigue: Initial validation of the fatigue impact scale. *Clinical Infectious Diseases, 18*(Suppl. 1), S79–S83.

Krupp, L. B., Alvarez, L. A., LaRocca, N. G., & Scheinberg, L. C. (1988). Fatigue in multiple sclerosis. *Archives of Neurology, 45*(4), 435–437.

Krupp, L. B., LaRocca, N. G., Muir-Nash, J., & Steinberg, A. D. (1989). The Fatigue Severity Scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. *Archives of Neurology, 46*(10), 1121–1123.

Mathiowetz, V., Matuska, K. M., & Murphy, M. E. (2001). Efficacy of an energy conservation course for persons with multiple sclerosis. *Archives of Physical Medicine & Rehabilitation, 82*(4), 449–456.

McLaughlin, J., & Zeeberg, I. (1993). Self-care and multiple sclerosis: A view from two cultures. *Social Science & Medicine, 37*(3), 315–329.

Multiple Sclerosis Council for Clinical Practice Guidelines (MSCCPG). (1998). *Fatigue and multiple sclerosis: Evidence-based management strategies for fatigue in multiple sclerosis.* Washington, DC: Paralyzed Veterans of America.

Munro, B. (1997). *Statistical methods for health care research,* (3rd ed.). Philadelphia: Lippincott.

Portney, L., & Watkins, M. (2000). *Foundations of clinical research: Applications to practice,* (2nd ed.). Upper Saddle River, NJ: Prentice Hall Health.

Ritvo, P. G., Fischer, J. S., Miller, D. M., Andrews, H., Paty, D., & LaRocca, N. G. (1997a). *Multiple Sclerosis Quality of Life Inventory (MSQLI): A user's manual.* New York: National Multiple Sclerosis Society.

Ritvo, P. G., Fischer, J. S., Miller, D. M., Andrews, H., Paty, D., & LaRocca, N. G. (1997b). *Multiple Sclerosis Quality of Life Inventory (MSQLI): Technical supplement.* New York: National Multiple Sclerosis Society.

Schwartz, J. E., Jandorf, L., & Krupp, L. B. (1993). The measurement of fatigue: A new instrument. *Journal of Psychosomatic Research, 37*(7), 753–762.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlation: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–428.

Ware, J. E., Snow, K., Kosinski, M., & Gandek, B. (1993). *SF-36 Health Survey: Manual & interpretation guide.* Boston: Health Institute.