Comparison of the Psychometric Properties of Three Fatigue Scales in Persian-Speaking Patients with Multiple Sclerosis

Shabnam Behrangrad, MSc; Amin Kordi Yoosofinejad, PhD

Background: Fatigue is the most disabling symptom in patients with multiple sclerosis (MS). Although there is no standard tool to evaluate fatigue in clinical settings, the Fatigue Impact Scale (FIS), Fatigue Severity Scale (FSS), and Multidimensional Assessment of Fatigue (MAF) scale are popular instruments for this purpose. The aim of this study was to compare the psychometric properties of the Persian versions of these scales.

Methods: One hundred thirty adult patients with MS and 60 controls participated in this study. They completed the scales on two occasions 3 days apart. Reproducibility and internal consistency were evaluated as intraclass correlation coefficients (ICCs) and Cronbach α. Convergent validity was assessed by evaluating the association of the fatigue scales with age, sex, Expanded Disability Status Scale (EDSS) score, disease duration, and sleep quality. Dimensionality was evaluated using confirmatory factor analysis. Acceptability and known-group validity were investigated. The effect size of each scale was computed.

Results: The ICC of all instruments was 0.99. Internal consistency was 0.97 for the MAF scale, 0.93 for FSS, and 0.83 for FIS. The instruments showed moderate-to-good correlations with Pittsburgh Sleep Quality Index, EDSS score, and disease duration. Acceptability was acceptable. The FIS had three dimensions, and the FSS and MAF scale were unidimensional. All scales were able to discriminate patients with MS from controls.

Conclusions: The Persian version of the MAF scale seems to be the most suitable instrument to evaluate fatigue in patients with MS based on its time efficiency, effect size, and detailed data about various aspects of fatigue. Int J MS Care. 2021;23:8-15.

Up to 92% of patients with multiple sclerosis (MS) report chronic fatigue as their most disabling symptom.1 According to published research, MS-related fatigue is a multidimensional, non-specific subjective symptom2 with a widespread effect on several aspects of the patient’s life, including physical activity3,4 and the level of participation in activities of daily3,5 and social living.1,4 Fatigue is known to be one of the main factors leading to decreased quality of life in patients with MS.6

The etiology and pathophysiology of fatigue are not fully understood.7,8 Because no definite, objective factors in its pathogenesis have been identified, it is difficult to define an appropriate, objective gold standard for assessing fatigue in patients with MS. Hence, finding an appropriate screening tool to evaluate fatigue in these patients is challenging. Because fatigue is a nonspecific subjective symptom, its objective assessment is complex.9
and most measurement tools are based on the individual’s own judgment. Previously, the correlation between fatigue and variables such as disease duration, sleep quality, disability level, age, and sex has been confirmed.\(^9,10,11\)

Many comprehensive self-reported scales have been developed in an attempt to assess the nature, severity, and impact of fatigue in various clinical situations. Among the most popular ones are the Fatigue Impact Scale (FIS), Fatigue Severity Scale (FSS), and Multidimensional Assessment of Fatigue (MAF) scale. The FIS was designed by Fisk et al\(^1\) in 1994 to evaluate the impact of fatigue on activities of daily living and is among the most frequently used generic fatigue scales in patients with MS. The FSS is a valid and reliable scale developed by Krupp et al\(^12\) with high internal consistency for assessing fatigue in people with MS and other neurologic and medical conditions. The MAF scale is a comprehensive compilation of questions that measure fatigue aspects in patients with chronic illnesses. It was first developed by Belza et al\(^1\) to measure fatigue in patients with rheumatoid arthritis,\(^13\) and it was further developed for rapid fatigue assessment in both clinical screening and clinical trials.\(^14\)

The previously mentioned instruments were validated and cross-culturally adapted to the Persian language\(^15-17\); however, it is not clear which of them is the most suitable for evaluating different aspects of fatigue in patients with MS in terms of validity, reliability, acceptability, and time needed to complete. Obviously, asking patients to complete all three instruments simultaneously would be time-consuming and unfeasible.

There are approximately 110 million Persian speakers worldwide; the language holds official status in countries such as Iran, Afghanistan, and Tajikistan. Identifying a suitable instrument to evaluate fatigue in Persian-speaking patients with MS can provide neurologists and other related clinicians with valuable information during the evaluation of fatigue in these patients. In this study, we mainly focused on the relapsing-remitting subtype of MS (RRMS) because this is the most common subtype (reported prevalence, 70% to 90%).\(^18\) The prevalence of this type of MS is increasing rapidly in Persian-speaking countries and has been reported as 81% of patients with MS in Pakistan\(^19\) and 65.8% to 87.8% of those in Iran.\(^20\)

At present, more than half of Iran’s population is younger than 35 years, born since 1987, and the largest age group is reported to be 20- to 24-years-old. Given that MS mostly affects young adults (20-40 years old), the public health system faces an increasing disease burden.\(^20\) Accordingly, the aim of this study was to compare psychometric characteristics such as reliability (reproducibility, internal consistency), validity (convergent validity, known-group validity), acceptability, effect size, and cutoff scores in the self-administered instruments noted previously herein to evaluate fatigue in Persian-speaking patients with RRMS who were able to walk.

**Methods**

This cross-sectional study recruited a convenience sample of 170 patients with RRMS through printed advertisements on notice boards at various sites in five outpatient physiotherapy clinics, two outpatient rehabilitation centers in hospitals, and a self-help MS community organization between May 2016 and May 2017. Patients were included if they were between 20 and 40 years old and had a diagnosis of MS in accordance with the McDonald diagnostic criteria,\(^21,22\) an Expanded Disability Status Scale (EDSS) score between 2 and 4.5,\(^15\) and a history of RRMS for at least 2 years.\(^9,23\) The exclusion criteria were pregnancy; acute relapse during the previous 6 weeks; use of sedatives, fatigue-inducing drugs, or immunosuppressive agents; comorbid conditions\(^9,23,24\); and secondary fatigue confirmed by an expert neurologist. Sample size was calculated as a minimum of 117 patients with MS. The calculation of sample size was based on a pilot study, with an alpha level of 0.05, a beta level of 0.2, and a dropout rate of 10%. To estimate the sample size, the pilot group filled out the Persian version of the MAF scale. The primary outcome measure used to calculate the sample size was the intraclass correlation coefficient (ICC) obtained from test-retest of the pilot study of the Persian version of the MAF scale.

Of 170 recruited patients diagnosed as having MS, 130 were eligible to participate in the study. The main reasons for exclusion were relapse remission during the previous 6 months (15 patients), potential effects of medications (12 patients), and refusal to participate in the study (13 patients). All 130 patients completed the instruments on two occasions 3 days apart.\(^9,24\)

Medical and neurologic examination findings, disease duration (time since diagnosis), and EDSS score were determined by a neurologist specialized in MS who was blinded to the study process. A signed written informed consent form was obtained from all participants. The study was approved by the ethics committee of Shiraz University of Medical Sciences (Shiraz, Iran) in accordance with the standards of the Helsinki Declaration.

To investigate the known-group validity of the instruments, 60 healthy people aged 20-40 years completed the questionnaires as a control group. The control group was recruited by convenience sampling among healthy volunteers with no history of neurologic or orthopedic disorders. The control group was recruited through advertisement and personal recruitment at public places.
Procedures
After verification of the inclusion criteria, including EDSS score and type and duration of disease, each participant was asked to complete a study booklet that contained the Persian versions of the MAF scale, FSS, and FIS and also the Persian version of the Pittsburgh Sleep Quality Index (PSQI). All participants completed all instruments on two occasions at a 3-day interval. The order of the questionnaires in the booklets was randomized before being handed to the participants. Note that the patients were asked to report on their fatigue for only the previous week.

Instruments
Multidimensional Assessment of Fatigue Scale
The MAF scale is a reliable, valid, and responsive 16-item self-administered questionnaire developed to assess fatigue in patients with chronic disorders. It was designed to measure five dimensions of fatigue, including degree, severity, distress, timing, and effect on various activities of daily living.\(^15\) To score the MAF scale, a numerical rating scale is used: 1 to 10 for item 1, 1 = mild to 10 = severe for item 2, 1 = no distress to 10 = a great deal of distress for item 3, and 1 = not at all to 10 = a great deal for items 4 to 14. A categorical response (score range, 1-4) for timing was used for items 15 and 16. Items 15 and 16 have multiple-choice responses. The 10-point numerical rating scale ranges from 1 (not at all) to 10 (a great deal). For all respondents, the score for item 15, which asks about the frequency of fatigue, was converted from 1-4 to 2.5-10 by multiplying the responses by 2.5. This conversion then allowed items measuring degree of fatigue (item 1), severity of fatigue (item 2), distress of fatigue (item 3), the average effect on activities of daily living items (items 4-14), and the newly scored frequency of fatigue item (item 15) to be summed to create the total score of the MAF scale. The MAF scale score ranges from 1 (no fatigue) to 50 (severe fatigue). It takes less than 5 minutes to complete the scale. The MAF scale was originally developed by Belza in 1991 and is a modification of the Piper Fatigue Scale developed in 1989.\(^15\) The Persian version of the MAF scale was developed by Behrangrad and Kordi Yoosefinejad.\(^15\)

Fatigue Impact Scale
The FIS is a 40-item questionnaire comprising three sub-scales to assess the impact of perceived fatigue on cognitive functioning (ten items), physical functioning (ten items), and psychosocial functioning (20 items). Respondents are required to rate items for which fatigue has caused problems during the previous month on a scale from 0 (no problem) to 4 (extreme problem). The FIS score ranges from 0 (no fatigue) to 160 (severe fatigue). It has been validated for different languages and used in several studies of MS.\(^31-33\) The Persian version of this scale was validated for patients with MS and showed good reliability with high internal consistency.\(^16\)

Fatigue Severity Scale
The FSS is a self-administered nine-item questionnaire measuring the severity of fatigue and its effect on a person’s activities and lifestyle during the previous week. The items are scored from 1 (strong disagreement) to 7 (strong agreement). A total score of 4 or higher represents a significant effect of fatigue on the respondent’s life. The FSS score ranges from 1 to 7. The higher the score, the more severe the fatigue is and the more it affects the person’s activities. The reliability and validity of the Persian version of the FSS in patients with MS were confirmed by Azimian et al.\(^17\)

Pittsburgh Sleep Quality Index
The PSQI is widely used to evaluate sleep latency, sleep duration, and habits leading to good sleep quality. Its seven components provide an overall picture of sleep quality and quantity. The psychometric properties of the Persian version of the PSQI were confirmed previously.\(^31\)

Statistical Analysis
To evaluate the reproducibility of the scales, the ICC was calculated (ICC <0.4 indicates poor reliability; 0.4-0.74, moderate-to-good reliability; and 0.75, excellent reliability). Cronbach α was calculated to investigate the internal consistency of the scales. Values exceeding 0.70 are considered acceptable. Item-total correlation adjusted for specific items was also calculated as a measure of internal consistency. A correlation of at least 0.4 was assumed as a standard to provide internal consistency.\(^32,33\) Convergent validity was evaluated using the Spearman correlation coefficient among the evaluated fatigue scales and variables such as age, EDSS score, disease duration, and PSQI. To evaluate the association between sex and the fatigue questionnaires, the point-biserial correlation coefficient was used. The correlation of age, EDSS score, disease duration, sleep quality, and sex with fatigue in patients with MS was evaluated using the Pearson correlation coefficient. The r values were interpreted as follows: 0 to 0.19 = very weak correlation, 0.2 to 0.39 = weak correlation, 0.4 to 0.69 = moderate correlation, 0.7 to 0.89 = strong correlation, and 0.9 to 1.00 = very strong correlation. The known-group validity was evaluated for each scale using an independent-samples t test to compare results between the MS and control groups. To evaluate the dimensionality of the questionnaires, confirmatory factor analysis was used. Acceptability of the questionnaires was evaluated using the following indices: possible and observed change, mean ± SD score, floor or ceiling effects, and skewness.\(^34\) Floor and ceiling effects are based on the assumption that more than 15% of participants obtain either the lowest or highest score on the questionnaire. An absolute skew value of lower than 2 indicates normality.\(^35\) The effect size of the scales was also calculated using Cohen’s d. Calculation of the cutoff scores of the questionnaires was based on receiver operating characteristic curve analysis. Considering the specificity and sensitivity values, the cutoff point was the point on the curve where sensitivity and specificity values were equal.\(^36\)

Results
A total of 130 patients with RRMS (87 women, 43 men) and 60 controls (33 women, 27 men) participated in this study. Patients with MS had a mean ± SD age of 33.65 ± 4.18 years. In the MS group, 60 patients were

International Journal of MS Care
10
married and 33 were divorced or widowed. The mean ± SD EDSS score was 2.17 ± 1.43 (range, 1-2.5) and disease duration was 7.55 ± 3.2 (range, 4-13) years. The mean ± SD age of the control group was 35.01 ± 5.91 years. In the control group, 31 participants were married and the others were either divorced or widowed. There was no significant difference in age (P = .52) between the groups. Known-group validity was evaluated for all three scales. The Levene test for equality of variances was significant (P < .001) for all instruments, showing that the variances were nonhomogeneous.

Test-retest reliability and internal consistency results are shown in Table 1. Item-total analysis of the total score for the three subscales of the FIS (social, cognitive, and physical) and individual items yielded a Cronbach α greater than 0.80. Item-to-scale analysis of the FSS showed a corrected item-total correlation of 0.63 to 0.99 and a Cronbach α of 0.92 to 0.93 for the items. Item-to-scale analysis of the MAF scale revealed a corrected item-total correlation of 0.96 to 0.99 and a Cronbach α of 0.99 for the items.

All instruments were able to differentiate patients with MS from the people in the control group (Table 2).

Although the results showed that none of the instruments had an acceptable association with age or sex, a strong correlation was found between the MAF scale and EDSS score (r = 0.94), PSQI score (r = 0.91), and disease duration (r = 0.89). Likewise, the FSS showed a very strong correlation with EDSS score (r = 0.93) and PSQI score (r = 0.91) and a strong correlation with disease duration (r = 0.88). Total FIS score had a strong correlation with EDSS score (r = 0.84) and PSQI score (r = 0.77) and a moderate correlation with disease duration (r = 0.67). The MAF scale demonstrated very strong correlations with the FSS and FIS total (r = 0.99). Also, a very strong correlation (r = 0.98) was observed between the FSS and FIS total. It also yielded a strong correlation with the physical (r = 0.86) and cognition (r = 0.76) subcomponents of the FIS and a moderate correlation with the social subcomponent of the FIS (r = 0.59). The FSS had very strong correlation with the FIS total (r = 0.99), a strong correlation with the physical (r = 0.86) and cognition (r = 0.74) subcomponents of the FIS, and a moderate correlation (r = 0.57) with the FIS social subcomponent. The correlations between each instrument and PSQI score, sex, age, disease duration, and EDSS score are summarized in Table 3.

Acceptability of the indices was evaluated based on score range, mean scores, floor and ceiling effects, and skewness. The results are summarized in Table 4. None of the instruments showed floor or ceiling effects in the present study. The scores were not notably skewed.

The results of confirmatory factor analysis supported the unidimensionality of the MAF scale (Kaiser-Olkin = 0.965, χ² = 2681.457, df = 91, P < .001) and the FSS (Kaiser-Olkin = 0.899, χ² = 791.688, df = 36, P < .001) and identified three dimensions for the FIS (Kaiser-Olkin = 0.781, χ² = 2671.694, df = 780, P < .001).

To evaluate the effect sizes of the scales, Cohen’s d was computed and found to be 15.82 for the MAF scale, 8.03 for the FSS, and 11.14 for the FIS.

---

**Table 1. ICC reliability measures, administration times, cutoff scores, and effect sizes of the instruments compared**

| Instrument | Test-retest reliability (ICC[3,1]) | Internal consistency (Cronbach α) | Administration time, min | Cutoff score | Effect size (Cohen’s d) |
|------------|-----------------------------------|----------------------------------|--------------------------|-------------|-----------------------|
| MAF scale  | 0.99                              | 0.97                             | 5.12 ± 0.92              | 25.50       | 15.82                 |
| FSS        | 0.99                              | 0.93                             | 5.02 ± 0.85              | 2.65        | 8.03                  |
| FIS        | 0.99                              | 0.83                             | 29.97 ± 3.39             | 78          | 11.14                 |

Note: Values are given as number or mean ± SD unless otherwise indicated.

Abbreviations: FIS, Fatigue Impact Scale; FSS, Fatigue Severity Scale; ICC, intraclass correlation coefficient; MAF, Multidimensional Assessment of Fatigue.

---

**Table 2. Evaluation of known-group validity of the instruments**

| Instrument and group | Independent-samples t test | P value |
|----------------------|---------------------------|---------|
| MAF scale            |                           | <.001   |
| MS (n = 130)         | 44.47 ± 2.16              |         |
| Control (n = 60)     | 3.58 ± 2.95               |         |
| FSS                  |                           | <.001   |
| MS (n = 130)         | 5.15 ± 0.72               |         |
| Control (n = 60)     | 0.95 ± 0.17               |         |
| FIS                  |                           | <.001   |
| MS (n = 130)         | 128.4 ± 11.0              |         |
| Control (n = 60)     | 41.0 ± 1.41               |         |

Note: Values are given as mean ± SD unless otherwise indicated.

Abbreviations: FIS, Fatigue Impact Scale; FSS, Fatigue Severity Scale; MAF, Multidimensional Assessment of Fatigue; MS, multiple sclerosis.
Table 3. Associations between the fatigue instruments and sex, age, EDSS score, disease duration, and PSQI

| Variable          | MAF scale | FIS total | FIS social | FIS cognition | FIS physical | FSS  |
|------------------|-----------|-----------|------------|--------------|--------------|------|
| Age              | 0.19      | 0.16      | 0.09       | 0.07         | 0.2          | 0.19 |
| EDSS score       | 0.94      | 0.84      | 0.47       | 0.64         | 0.74         | 0.93 |
| PSQI             | 0.91      | 0.77      | 0.42       | 0.61         | 0.70         | 0.91 |
| Sex 1            | 0.02      | 0.02      | -0.12      | 0.001        | -0.09        | 0.02 |
| Disease duration | 0.89      | 0.67      | 0.41       | 0.56         | 0.60         | 0.88 |

Note: Spearman correlation coefficient was used to evaluate correlation between fatigue scales and all variables except sex. Abbreviations: EDSS, Expanded Disability Status Scale; FIS, Fatigue Impact Scale; FSS, Fatigue Severity Scale; MAF, Multidimensional Assessment of Fatigue; PSQI, Pittsburgh Sleep Quality Index.

*Point-biserial correlation coefficient was used to evaluate association between fatigue scales and sex.

Discussion
The objective of this study was to compare the psychometric characteristics of three instruments used to evaluate fatigue in patients with MS. The reliability and validity of the Persian version of each instrument were confirmed previously, and the present results suggested that each of them is useful in measuring fatigue in ambulatory patients with MS. We also show that the instruments were able to discriminate patients with MS from people in the control group. This finding confirms those of Amtmann et al,37 who reported significant known-group validity for the FSS and modified FIS in discriminating patients with MS with different EDSS scores from healthy people. In addition, Flensner et al40 showed that the Swedish version of the FIS discriminated between groups of individuals with differences in perceived impact of fatigue. Also, the Finnish version of the FSS yielded a higher mean fatigue score for patients with greater disease severity than for those with less severity.34 Because all participants in the present study were ambulatory (mean ± SD patient EDSS score, 2.17 ± 1.43), we compared known-group validity between patients with MS and a control group.

Table 4. Acceptability of the fatigue instruments

| Acceptability index | MAF scale | FSS | FIS |
|---------------------|-----------|-----|-----|
| Possible score range| 1-50      | 1-7 | 0-160 |
| Observed score range| 41-49     | 4-6.7 | 110-153 |
| Scale score         | 44.47 ± 2.16 | 5.14 ± 0.72 | 128.37 ± 10.9 |
| Floor effect        | 1 (0.8)   | 2 (1.5) | 2 (1.5) |
| Ceiling effect      | 1 (0.8)   | 3 (2.3) | 1 (0.8) |
| Skewness            | 0.37      | 0.40 | 0.29 |

Note: Values are given as number, mean ± SD, or number (percentage). Abbreviations: FIS, Fatigue Impact Scale; FSS, Fatigue Severity Scale; MAF, Multidimensional Assessment of Fatigue.

Reliability
The Persian version of all three instruments showed acceptable test-retest reliability (ICC = 0.99). The ICC was 0.88 for the Greek version of the FSS and 0.81 for the Turkish version38,39 and test-retest reliability for the Finnish version was not available.34

The internal consistency of the Persian version of the MAF scale, FSS, and FIS according to Cronbach α was 0.97, 0.93, and 0.83, respectively. To our knowledge, only two studies have investigated the psychometric characteristics of the MAF scale in patients with MS. In one study, Schwartz et al40 evaluated the MAF scale in 139 patients with MS (EDSS score, 4.71 ± 1.85) and reported Cronbach α as 0.93. In agreement with the present results, the internal consistency of the MAF scale score was reported to be high (Cronbach α, 0.85). The present findings also confirmed those of Fraser et al41 who evaluated 79 patients with MS (mostly RRMS). They reported a Cronbach α for the MAF scale of 0.93.41

Amtmann et al37 reported a Cronbach α for the FSS of 0.93. In addition, Cronbach α for the Greek (0.96), Turkish (0.89), and Swiss (0.93) versions of the FSS were consistent with the present results.38 Moreover, high internal consistency (Cronbach α = 0.95) was reported for the Finnish version of the FSS when used to evaluate fatigue in patients with MS.34

In line with the present findings, the Hungarian version of the FIS yielded an acceptable Cronbach α (0.98) in patients with MS.29 Also, the dimensions of the French FIS showed high internal consistency (Cronbach α, >0.80).28 In the present study the item-total correlation for the MAF scale was 0.97 to 0.98, which confirmed the homogeneity of the Persian version of this instrument. The item-total correlation was 0.93 for the FSS and 0.73 to 0.74 for the FIS total. The item-total correlation ranged from 0.56 to 0.89 for the FSS and...
from 0.66 to 0.80 for the modified version of the FIS.\textsuperscript{37} In addition, the Finnish version of the FSS had a high item-to-total correlation ($r = 0.62-0.87$).\textsuperscript{34}

**Convergent Validity**

This study investigated the convergent validity of the Persian version of the MAF scale, FSS, FIS, and FIS subscales according to age, EDSS score, PSQI, sex, and disease duration. All the fatigue scales showed very weak correlation with age or weak association with sex. Very strong to strong correlations were found between the scales and EDSS score, PSQI, and disease duration. The FIS total score showed moderate correlation with disease duration ($r = 0.67$). Amtmann et al\textsuperscript{37} reported moderate-to-strong correlations for the FSS and a modified version of the FIS and its subscales ($\rho = 0.55-0.77$). The present results are comparable with those of previous studies. Ghajarzadeh et al\textsuperscript{42} found that patients with a longer duration of MS had higher FSS scores. In line with the present findings, they reported that patients with higher EDSS scores experienced more fatigue. Mills and Young,\textsuperscript{43} however, found no correlation between fatigue and disease duration or patient age. However, similar to us, they found a close relationship between fatigue scores and sleep. Fazli and Shayesteh-Azar\textsuperscript{44} found a weak correlation between fatigue measures, ie, a visual analogue scale of fatigue status and the Fatigue Impact Scale for Daily Use, and MS duration, time from symptom initiation, and sex. We also observed a very weak association between fatigue scores and patient age and sex but a strong correlation between fatigue scores and disease duration. The mean ± SD duration of MS was 3.80 ± 3.35 years in the study by Fazli and Shayesteh-Azar,\textsuperscript{44} whereas it was 7.55 ± 3.2 years in the present study. Note that the type of MS was not determined in the study by Fazli and Shayesteh-Azar. The MAF scale had very strong correlations with the FSS and FIS total in the present study; in addition, we found a very strong correlation between the FSS and FIS total. Both FSS and FIS scores had moderate correlation ($r = 0.56-0.68$) with the Hamilton Depression Rating Scale and the Beck Depression Inventory in a study by Smarr and Keefer.\textsuperscript{45} In contrast to the present findings, the Finnish version of the FSS showed weak correlation with EDSS score ($r = 0.33$).\textsuperscript{34} Considering the mean ± SD EDSS score in their study (4 ± 2.5) and that in the present study (2.17 ± 1.43), greater correlation was expected in the Finish version; however, note that many factors other than the variability range of the variables would affect the value of correlation. Some of these factors are the shape of the distribution of the variables, lack of linearity, presence of some outliers, characteristics of the sample, and measurement error.\textsuperscript{46} So, the greater correlation in the present study might be attributed to one or a combination of the previously mentioned factors. Another study found a strong correlation between MAF scale score and the Profile of Mood States fatigue scale.\textsuperscript{41} In contrast to the present findings, Schwartz and colleagues\textsuperscript{48} found a very weak correlation between MAF scale and EDSS scores. They investigated a heterogeneous sample of patients with MS (EDSS score, 1–8), whereas the present participants were all ambulatory.

No ceiling or floor effects were observed in the present study. This is in line with the findings of previous studies that reported minimal floor and ceiling effects for the FSS.\textsuperscript{34,37} The Persian version of the MAF scale was also found to have no floor or ceiling effects.\textsuperscript{24} Confirmatory factor analysis showed that the Persian versions of the MAF scale and FSS were unidimensional and the Persian FIS was a three-dimensional scale. The unidimensionality of the FSS was previously confirmed in similar studies. To our knowledge, the dimensionality of the MAF scale has not been reported previously in patients with MS. Fairbrother et al\textsuperscript{47} confirmed the unidimensionality of the MAF scale in a group of pregnant and postpartum women. Wolfe\textsuperscript{48} found three dimensions for the MAF scale in patients with rheumatoid arthritis. There are disagreements in the literature about the unidimensionality or multidimensionality of the MAF scale. The other limitation was the availability of studies evaluating the MAF scale in populations with MS to be considered for comparison.

**Instrument Selection**

Although the MAF scale, FSS, and FIS seem to be gold standard instruments in fatigue assessment, some considerations can influence which to choose in practice. First, when clinicians need to assess many patients with MS and require rapid processing of the results and a general view of their fatigue, the FSS and MAF scale may be more appropriate because they need less time to complete (approximately 5 minutes for the MAF scale and FSS vs 30 minutes for the FIS). Multidisciplinary clinical practitioners or researchers may perceive that the FIS and MAF scale can provide more comprehensive and detail-oriented assessments.

Questionnaire content reflects points of differentiation among instruments. For example, the FSS does not
differentiate among functions leading to fatigue, such as walking, daily activity, or sexual relations, and measures only the severity of fatigue symptoms with a few items, which may overlook subjectivity or exclusive analysis. In addition, although the MAF was developed to measure chronic fatigue, it was not developed specifically for MS and does not include an explicit cognitive or psychosocial subscale. However, the FIS has more content on psychosocial functioning than on fatigue related to physical activities. Another difference among instruments is the time patients need to complete them. The average time to complete the MAF scale and FSS was 5 minutes, whereas it took 30 minutes for patients to complete the FIS. Despite its being multidimensional, the FIS may have limited clinical utility because it is apparently more time-consuming (40 items). All three instruments had a large effect size, but Cohen’s $d$ was greater for the MAF scale than for the other two instruments and was greater for the FIS than for the FSS.

Limitations and Conclusion

The main limitation of this study was possible selection bias due to convenience sampling. In addition, data were not available for the 40 patients who were excluded or who declined to participate in the study. Therefore, included individuals may have been biased toward more highly motivated patients. It is also possible that asking patients to consider their fatigue only during the week before completing the instruments may have limited the accuracy of the measurements. However, this is unlikely because, in our experience, most patients tend to report their fatigue status in a short-term time frame whether or not they are specifically directed to do so. Another limitation of this study was the lack of a gold standard to compare fatigue scales in patients with MS. As far as we know, no gold standard has been determined for comparing the fatigue scales used in patients with MS. This might be attributed to the complexity and subjective nature of fatigue. The other limitation was the lack of investigating and comparing the responsiveness of the fatigue scales and divergent validity. A final limitation was that we did not control for whether the participants had depression.

Although a careful choice of instrument is important, clinicians should have alternative or multiple strategies for evaluating fatigue, and they should be aware of the limitations of each method. To our knowledge, this is the first study to compare the Persian versions of three instruments used to investigate fatigue in patients with MS. Nevertheless, these instruments should be tested further in patients with other types of MS (ie, primary progressive and secondary progressive subtypes) in larger samples of patients. In addition, future studies should expand the age range to provide greater generalizability of the results. The participants in the present study were able to walk (EDSS score, 1.0-3.5); hence, the results we obtained cannot be generalized to patients whose EDSS scores are outside this range.

In conclusion, comparison of the Persian versions of the MAF scale, FIS, and FSS showed that all instruments were suitable for evaluating fatigue in ambulatory patients with RRMS, but when factors such as time efficiency, effect size, and ability to provide detailed and comprehensive information are considered, patients may prefer the MAF scale to other instruments.

Acknowledgments: The authors thank all the participants for their kind cooperation and K. Shashok (AuthorAID in the Eastern Mediterranean) for improving the use of English in the manuscript.

Financial Disclosures: The authors declare no conflicts of interest.

Funding/Support: None.

References

1. Fisk JD, Pontefract A, Ritvo PG, Archibald CJ, Murray T. The impact of fatigue on patients with multiple sclerosis. *Can J Neurol Sci.* 1994;21:9-14.
2. Giovannoni G. Multiple sclerosis related fatigue. *J Neurol Neurosurg Psychiatry.* 2006;77:23.
3. Meads D, Doward L, McKenna S, Fisk J, Twiss J, Eckert B. The development and validation of the Unidimensional Fatigue Impact Scale (U-FIS). *Mult Scler.* 2009;15:1228-1238.
4. Zifko UA. Management of fatigue in patients with multiple sclerosis. *Drugs.* 2004;64:1295-1304.
5. Mills R, Young C. A medical definition of fatigue in multiple sclerosis. *QJM.* 2008;101:49-60.
6. Amato M, Ponziani G, Rossi F, Liedl C, Stefanile C, Rossi L. Quality of life in multiple sclerosis: the impact of depression, fatigue and disability. *Mult Scler.* 2001;7:340-344.

PRACTICE POINTS

- The Persian versions of three questionnaires—the Fatigue Impact Scale, Fatigue Severity Scale, and Multidimensional Assessment of Fatigue scale—have been validated to evaluate fatigue in patients with MS, and the present study suggests that each of them is useful in measuring fatigue in ambulatory patients with MS.
- When considering factors such as time efficiency, effect size, and ability to provide detailed and comprehensive information, patients may prefer the Multidimensional Assessment of Fatigue scale to other instruments.
Piper BF, Dibble SL, Dodd MJ, Weiss MC, Slaughter RE, Paul SM. The Fatigue Impact Scale in multiple sclerosis: mechanisms, evaluation, and treatment. Sleep. 2010;33:1061-1067.

Bahrani H, Rostam S, Bahiri R, Hakkou J, Aissaoui N, Hajjaj-Hassouni N. Psychometric evaluation of the Arabic version of the Multidimensional Assessment of Fatigue scale (MAF) for use in patients with amyotrophic spondylitis. Rheumatol Int. 2012;32:3969-3976.

Aldughmi M, Huisings J, Lynch SG, Siengukon CF. The relationship between fatigueability and sleep quality in people with multiple sclerosis. Mult Scler J Exp Transl Clin. 2016;2:2055217316682774.

Ayginolgu SK, Celebi A, Vardar N, Gursoy E. Correlation of fatigue with depression, disability level and quality of life in patients with multiple sclerosis. Nöro Psikiyatr Ar. 2015;52:247-251.

Krupp LB, Alvarez LA, LaRocca NG, Scheinberg LC. Fatigue in multiple sclerosis. Arch Neurol. 1988;45:435-437.

Belza B, Miyawaki C, Liu M, Zhang X, Fessel M. The Multidimensional Assessment of Fatigue Scale: a 25-year review and evaluation. Arthritis Rheum. 2015;67:3077-3078.

Behrangrad S, Kordi Yoosofinejad A. Validity and reliability of the multidimensional assessment of fatigue scale in Iranian patients with relapsing-remitting subtype of multiple sclerosis. Disabil Rehabil. 2018;40:673-677.

Heidari M, Akbarfahimi M, Salehi M, Nabavi SM. Validity and reliability of the Persian version of Fatigue Impact Scale in multiple sclerosis patients in Iran. Koomesh. 2014;15:295-301.

Azimian M, Shahvaraghi Farahani A, Daddkhah A, Fallahpour M, Karimlu M. Fatigue Severity Scale: the psychometric properties of the Persian-version in patients with multiple sclerosis. Res J Biol Sci. 2009;4:974-977.

Sumelehto ML, Tienari PJ, Hakama M, Wikström J. Multiple sclerosis in Finland: incidence trends and differences in relapsing remitting and primary progressive disease courses. J Neurol Neurosurg Psychiatry. 2003;74:25-28.

Wassay M, Ali S, Khatri I, et al. Multiple sclerosis in Pakistan. Mult Scler. 2007;13:686-689.

Ettemadifar M, Sajjadi S, Naaz Z, et al. Epidemiology of multiple sclerosis in Iran: a systematic review. Eur Neurol. 2013;70:356-363.

McDonald WI, Compston A, Edan G, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. Ann Neurol. 2001;50:121-127.

Polman CH, Reingold SC, Edan G, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the “McDonald Criteria.” Ann Neurol. 2005;58:840-846.

Larson RD. Psychometric properties of the Modified Fatigue Impact Scale. Int J MS Care. 2013;15:15-20.

Behrangrad S, Kordi Yoosofinejad A. Validity and reliability of the Multidimensional Assessment of Fatigue Scale in Iranian patients with relapsing-remitting subtype of multiple sclerosis. Disabil Rehabil. 2018;40:673-677.

Rosenthal TC, Majeroni BA, Pretorius R, Malik K. Fatigue: an overview. J Clin Epidemiol. 2008;61:1173-1179.

Pouchot J, Kherani RB, Brant R, et al. Determination of the minimal appropriate test cut-off value: the case of tests with continuous results. Biochem Med. 2016;26:297-307.

Ammann D, Bamer AM, Noonan V, Lang N, Kim J, Cook KF. Comparison of the psychometric properties of two fatigue scales in multiple sclerosis. Rehabil Psychol. 2012;57:159-166.

Bakolioudou D, Skordilis EK, Giannopoulos S, Stamboulis E, Voumvourakis K. Validity and reliability of the FSS in Greek MS patients. Springer Plus. 2013;2:304.

Ozturk EK, Kicer BG, Gundogdu I, Umay E, Cakic FA. Reliability and validity study of a Turkish version of the fatigue severity scale in Parkinson’s disease patients. Int J Rehabil Res. 2017;40:185-190.

Schwarz CE, Coulthard-Morris L, Zeng G. Psychosocial correlates of fatigue in multiple sclerosis. Arch Phys Med Rehabil. 1996;77:165-170.

Fraser R, Johnson E, Clemmons D, Getter A, Johnson K, Gibbons L. Vocational rehabilitation in multiple sclerosis (MS): a profile of clients seeking services. Work. 2003;21:69-76.

Ghajarzadeh M, Jalilian R, Eskandari G, Sahraian MA, Azimi A, Mohammadifar M. Fatigue in multiple sclerosis: relationship with disease duration, physical disability, disease pattern, age and sex. Acta Neurol Belg. 2013;113:411-414.

Mills RJ, Young CA. The relationship between fatigue and other clinical features of multiple sclerosis. Mult Scler. 2011;17:604-612.

Fazli M, Shayesteh-Azaz M. Correlation between the fatigue with gender, age and disease duration in multiple sclerosis patients. Int J Med Invest. 2013;2:206-209.

Smarr KL, Keefer AL. Measures of depression and depressive symptoms: Beck Depression Inventory-II (BDI-II), Center for Epidemiologic Studies Depression Scale (CES-D), Geriatric Depression Scale (GDS), Hospital Anxiety and Depression Scale (HADS), and Patient Health Questionnaire-9 (PHQ-9). Arthritis Care Res. 2011;63:1211-1217.

Goodwin LD, Leech NL. Understanding correlation: factors that affect the size of r. J Exp Educ. 2006;74:249-266.

Fairbrother N, Hurtten EK, Stoll K, Hall W, Kluza S. Psychometric evaluation of the Multidimensional Assessment of Fatigue Scale for use with pregnant and postpartum women. Psychol Assess. 2008;20:150-158.

Wolfe F. Fatigue assessments in rheumatoid arthritis: comparative performance of visual analog scales and longer fatigue questionnaires in 7760 patients. J Rheumatol. 2004;31:1896-1902.