Clinical Significance of the Multiple Sclerosis Functional Composite

Relationship to Patient-Reported Quality of Life

Deborah M. Miller, PhD; Richard A. Rudick, MD; Gary Cutter, PhD; Monika Baier, MS; Jill S. Fischer, PhD

Background: The Multiple Sclerosis Functional Composite (MSFC) was recommended by a task force of the National Multiple Sclerosis Society as a new clinical outcome measure for clinical trials. The task force recommended that the MSFC be validated against other measures of the disease, such as patient-reported quality of life.

Methods: Three hundred patients with multiple sclerosis (MS) representing the spectrum of disease severity were included in this cross-sectional study. The MSFC and Kurtzke Expanded Disability Status Scale (EDSS) were used as measures of disease severity. Clinical relevance of the disease severity scores was analyzed using measures included in the Multiple Sclerosis Quality of Life Inventory. The MSFC and EDSS scores were correlated with self-reported employment status, the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36), and the Sickness Impact Profile (SIP).

Results: The MSFC and EDSS scores were strongly correlated ($r = -0.80, P < .001$). The MSFC scores were correlated with patient-reported physical functioning (SIP Physical Summary Scale: $r = -0.71, P < .001$; SF-36 Physical Component Score: $r = -0.41, P < .001$). The MSFC scores were significantly but more weakly correlated with emotional functioning (SIP Psychosocial Summary Scale: $r = -0.34, P < .01$). After controlling for EDSS scores, there were significant residual correlations between the MSFC scores and measures of health-related quality of life, suggesting that the MSFC accounts for the variability in health-related quality of life measures not reflected by the EDSS.

Conclusions: The observed strong correlations between MSFC scores and validated measures of self-reported quality of life indicate that the MSFC scores are clinically relevant. This study supports a recommendation by the National Multiple Sclerosis Society Task Force to use the MSFC as a clinical outcome measure.

Arch Neurol. 2000;57:1319-1324
SUBJECTS AND METHODS

STUDY DESIGN

This is a secondary analysis of a prospective, multicenter, cross-sectional study that is intended to assess the relationship of the MSFC with HRQoL. In the original study, 300 subjects from 4 sites in the United States and Canada completed generic and disease-specific HRQoL measures for a cross-sectional study devised to develop a patient-based battery. Other data collected for this protocol included measures of physical and cognitive impairment. Among these were the 9-HPT, PASAT-3, and timed 25-foot walk, which make up the MSFC. Instructions for conducting and scoring the MSFC have been published elsewhere. For the MSFC validation described in this article, hypotheses regarding relationships between the MSFC and HRQoL measures were generated before data analysis was initiated.

SUBJECTS

Three hundred patients with clinically definite MS from 4 clinical sites were entered into the study. The sites included the Mellen Center (Cleveland, Ohio), St Agnes Hospital (White Plains, NY), Dalhousie University (Halifax, Nova Scotia), and St Michael's Hospital (Toronto, Ontario). The study used a stratified sampling plan based on disease severity and sex. The sex ratio was 2 females to 1 male to reflect the distribution in the general MS population. Additionally, subjects were selected to provide an even representation of mild (EDSS score, 0-3.0 inclusive), moderate (EDSS score, 3.5-6.5 inclusive), and severe (EDSS score, 7.0-8.5) neurologic impairment.

MEASURES AND RESEARCH QUESTIONS

The EDSS and timed 25-foot walk test were administered by neurologists at each of the 4 clinical sites. Neuropsychology technicians administered the 9-HPT and PASAT-3 and assisted subjects who were unable to independently complete the self-reported HRQoL measures. The HRQoL measures were selected for this study from a larger set of HRQoL measures included in the original MSQLI project prior to data analysis. Table 1 provides details on these measures and their predicted relation to the MSFC. The 2 generic HRQoL measures included in the MSQLI were the Sickness Impact Profile (SIP) and the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36). The Medical Outcomes Study–Social Support Questionnaire (MOS-SSS) was included as a measure of perceived social support. The Fatigue Impact Scale (FIS) was included as a disease-specific measure. Self-reported employment status was included.

The following HRQoL subscales were selected to assess general physical and psychosocial impact of health on well-being: (1) total score of the SIP (SIP Total); (2) Physical Summary Scale of the SIP (SIP Physical); (3) Psychosocial Summary Scale of the SIP (SIP Psychosocial); (4) Physical Component Score of the SF-36 (SF-36 PCS); and (5) Mental Component Score of the SF-36 (SF-36 MCS).

The total score of the FIS (FIS Total) and the physical subscale score of the FIS (FIS Physical) were included because fatigue is one of the most common and most incapacitating symptoms of MS. Because this measure and its subscales assess only the impact of fatigue, which is not assessed by the MSFC, it was anticipated that only weak correlations with the MSFC would be demonstrated. Employment status was included as a surrogate measure of health status impact and was expected to be moderately correlated with the MSFC. The Tangible Support subscale of the MOS-SSS (SSS-Tan) was included as an indicator of the level of personal assistance required to compensate for increasing disability. Because this social support measure would be relevant only for the most severely disabled patients, its correlation with MSFC was expected to be weak.

ANALYSIS PLAN

Using the means and SDs from the baseline visit for all patients in the study (internal standards), MSFC z scores were calculated as recommended in the Administration and Scoring Manual for the Multiple Sclerosis Functional Composite Measure, using the following formula:

\[
\text{MSFC Score} = \frac{\left( \frac{\text{Average 9-HPT} - \text{Baseline Mean 9-HPT}}{\text{Baseline SD 9-HPT}} \right) + \left( \frac{\text{Average Timed 25-ft Walk} - \text{Baseline Mean Timed 25-ft Walk}}{\text{Baseline SD Timed 25-ft Walk}} \right) + \left( \frac{\text{PASAT-3} - \text{Baseline Mean PASAT-3}}{\text{Baseline SD PASAT-3}} \right)}{3.0}
\]

Spearman rank correlations between MSFC scores and MSQLI measures were constructed to test the strength of the predicted relationships. Because it was anticipated that the magnitude of the correlations between the MSFC and EDSS measures would be large, a second step in the analysis plan involved calculation of partial correlations between the MSFC and HRQoL measures, controlling for EDSS. Finally, a comparison of MSFC-HRQoL correlations was planned to assess the relative performance of the MSFC and HRQoL measures for the total study group and for 3 subgroups based on severity of neurological impairment.

and MSFC scores in the pooled data set. (2) Individual component scores from the MSFC (eg, timed 25-ft walk and 9-HPT) were only moderately correlated with each other, suggesting that they represented partially independent dimensions of the MS process. (3) Change in MSFC scores was correlated in the expected direction with change in EDSS scores during a 1-year period. (4) Worsening in MSFC scores during the first year of observation predicted worsening of EDSS scores in the subsequent year for those patients who had stable EDSS scores during the first year.

The MSFC was recommended for use as a clinical outcome measure for MS trials. Observation 4 above was used to support the predictive validity of the MSFC, at least in terms of subsequent changes in EDSS scores. This was considered important in shortening the duration or decreasing the sample size of future MS studies. Because of the potential precision of quantitative mea-

(REPRINTED) ARCH NEUROL/VOL 57, SEP 2000 WWW.ARCHNEUROL.COM

©2000 American Medical Association. All rights reserved.
measure has undergone extensive validity and reliability testing and is being used in a variety of clinical studies (R. Philip Kinkel, MD, personal communication, April 2, 2000; and Nicholas G. LaRocca, PhD, personal communication, February 2, 2000).

Although the NMSS Clinical Outcomes Assessment Task Force had not yet presented its recommendations, the component measures that were ultimately included in the MSFC were included as measures of disease severity in the MSQLI development work. The inclusion of these measures now permits a secondary analysis of the data to evaluate the relationship between MSFC scores and MSQLI scores.

It is generally accepted that HRQoL data are distinct from objective measures of disease severity. However, significant correlations have been demonstrated between clinical and HRQoL data. The authors of at least 3 MS-specific HRQoL measures established validity of their measures by demonstrating that the instruments correlate in the anticipated manner with objective measures of disability. In the present study, the clinical significance of the MSFC was determined by analyzing its relationship with the MSQLI.

### RESULTS

All 300 subjects entered into the MSQLI project were included in this study. Descriptive statistics on the total sample and the 3 strata are provided in Table 2 and Table 3. The mean EDSS score was 4.61, mean age was 44.7 years, 66% of the patients were female, 93% were white, and 62% were married. One hundred thirteen pa-

### Table 1. Summary of Measures Used in the Study and Predicted Relationship With MSFC Scores

<table>
<thead>
<tr>
<th>Measure</th>
<th>Range</th>
<th>Characteristics</th>
<th>Prediction of Correlation With MSFC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSFC†</td>
<td>NA</td>
<td>Higher scores, better function</td>
<td>NA</td>
</tr>
<tr>
<td>SIP Total‡</td>
<td>0-100</td>
<td>Higher scores, worse function</td>
<td>Moderate negative correlation</td>
</tr>
<tr>
<td>SIP Physical‡</td>
<td>0-100</td>
<td>Higher scores, worse function</td>
<td>Strong negative correlation</td>
</tr>
<tr>
<td>SIP Psychosocial‡</td>
<td>0-100</td>
<td>Higher scores, worse function</td>
<td>Weak negative correlation</td>
</tr>
<tr>
<td>SF-36 PCS§</td>
<td>0-100</td>
<td>Higher scores, better function</td>
<td>Strong positive correlation</td>
</tr>
<tr>
<td>MSQLI§</td>
<td>0-100</td>
<td>Higher scores, better function</td>
<td>Weak positive correlation</td>
</tr>
<tr>
<td>FIS Total¶</td>
<td>0-160</td>
<td>Higher scores, worse function</td>
<td>Weak negative correlation</td>
</tr>
<tr>
<td>FIS Physical¶</td>
<td>0-40</td>
<td>Higher scores, worse function</td>
<td>Moderate negative correlation</td>
</tr>
<tr>
<td>Employment¶</td>
<td>0-2</td>
<td>Higher scores, better function</td>
<td>Moderate positive correlation</td>
</tr>
<tr>
<td>SSS-Tan#</td>
<td>0-100</td>
<td>Higher scores, better function</td>
<td>Positive correlation</td>
</tr>
</tbody>
</table>

* NA indicates not applicable.
† MSFC indicates the Multiple Sclerosis Functional Composite. The MSFC is expressed as an average of the z scores of the individual components. A z score of 0 is the average of the scores from the reference population; –1 is 1 SD below the reference population; +1 is 1 SD above the reference population, etc.
‡ SIP indicates the Sickness Impact Profile. The SIP Total is a behaviorally based health status measure. It consists of 136 items that describe activities associated with everyday living. It is scored according to the number and type of items that are endorsed. The 136 items are aggregated into 12 categories that represent specific areas of activity, including Rest and Sleep, emotional behavior, body care, and movement, home management, mobility, social interaction, ambulation, alertness behavior, communication, work, recreation and pastimes, and eating. The SIP Physical is made up of the 3 categories Body Care and Movement, Mobility, and Ambulation. The SIP Psychosocial is made up of the 4 categories Emotional Behavior, Social Interaction, Alertness Behavior, and Communication.
§ SF-36 indicates the Medical Outcomes Study 36-Item Short-Form Health Survey. It is a well-validated generic measure of quality of life that is intended for use in a diverse population. As the result is intense psychometric development, it reduces the respondent's burden by requiring only 36 items to assess 8 domains of daily functioning. The SF-36 Physical Component Score (PCS) is made up of 4 scales, including Physical Functioning, Role Physical, Bodily Pain, and General Health. The SF-36 Mental Component Score (MCS) is made up of the remaining 4 scales: Vitality, Social Functioning, Role Emotional, and Mental Health.
¶ FIS indicates the Fatigue Impact Scale. The FIS Total is a validated, 21-item, symptom-specific measure that assesses the impact of fatigue on daily functioning in terms of cognition, physical abilities, and social functioning. The FIS Physical subscale includes 10 items that address physical discomfort, weakness, and ability to sustain physical activity.
# SSS-Tan indicates the Tangible Support subscale of the Social Support Survey; a brief multidimensional measure for perceived social support based on the Medical Outcomes Study. The SSS-Tan subscale inquires about the availability of others to help with errands and other types of concrete help.

From 1994 to 1997, a period overlapping the development of the MSFC (1992–1995), an independent research group sponsored by the Consortium of Multiple Sclerosis Centers and funded by the NMSS developed an MS-specific HRQoL measure, the Multiple Sclerosis Quality of Life Inventory (MSQLI). The battery was developed to assess patient perceptions of response to treatment in clinical trials. Development of the MSQLI involved analysis of generic and disease-specific measures in relation to manifestations of the MS disease process. The measure has undergone extensive validity and reliability testing and is being used in a variety of clinical studies (R. Philip Kinkel, MD, personal communication, April 2, 2000; and Nicholas G. LaRocca, PhD, personal communication, February 2, 2000).
tients (37.7%) scored 0 to 3.5 on the EDSS, 131 (43.7%) scored 3.5 to 6.5, and 56 (18.7%) scored 7.0 to 8.5.

For patients in this study, the mean SIP Total score was 23.3, the mean SIP Physical score was 22.6, and the mean SIP Psychosocial score was 20.6. Subjects also demonstrated poor quality of life (QoL) on the SF-36. The mean SF-36 PCS score (34.7) was below the 25th percentile of the general population, and the mean SF-36 MCS score (47.9) was between the 25th and 50th percentiles of the general population.

As expected, employment declined with MS severity as measured by the EDSS. Sixty-one percent of those in the total group were unemployed, ranging from 37.5% of the patients in the lowest disability group to 85.7% of the patients in the highest disability group. Mean ± SD age was 44.5 ± 9.3 years (range, 23-65 years). Perceived need for tangible social support increased as expected in the highest disability group. Fatigue severity was significant in the population but did not differ among the 3 EDSS subgroups.

The Figure shows the relationship between the MSFC and EDSS scores. The Spearman correlation coefficient (~0.80) indicates a strong correlation between the 2 measures. Thus, EDSS scores accounted for 64% of the variance in MSFC scores. As EDSS scores increased, there was a progressive decrease in the MSFC scores, indicating poorer performance on the functional

### Table 2. Demographic Characteristics of the Study Population

<table>
<thead>
<tr>
<th>EDSS Score*</th>
<th>Total Sample (N = 300)</th>
<th>0-3.0 (n = 113)</th>
<th>3.5-6.5 (n = 131)</th>
<th>7.0-8.5 (n = 56)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD), y</td>
<td>44.7 (9.3)</td>
<td>41.9 (8.9)</td>
<td>46.1 (8.7)</td>
<td>47.1 (10.3)</td>
</tr>
<tr>
<td>Female/male, %</td>
<td>66/34</td>
<td>70.8/29.2</td>
<td>63.4/36.6</td>
<td>62.5/37.5</td>
</tr>
<tr>
<td>Race, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>93</td>
<td>94.7</td>
<td>92.4</td>
<td>91.1</td>
</tr>
<tr>
<td>African American</td>
<td>0.7</td>
<td>0.9</td>
<td>0</td>
<td>1.8</td>
</tr>
<tr>
<td>Asian</td>
<td>0.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>American Indian</td>
<td>6</td>
<td>3.5</td>
<td>7.6</td>
<td>7.1</td>
</tr>
<tr>
<td>Employment status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>62.8</td>
<td>37.5</td>
<td>74.6</td>
<td>85.7</td>
</tr>
<tr>
<td>Part-time</td>
<td>13.1</td>
<td>20.5</td>
<td>10.0</td>
<td>5.4</td>
</tr>
<tr>
<td>Full-time</td>
<td>24.2</td>
<td>42.0</td>
<td>15.4</td>
<td>8.9</td>
</tr>
</tbody>
</table>

*EDSS indicates Expanded Disability Status Scale.

### Table 3. Health-Related Quality-of-Life Scores for the Study Population*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Total Sample (N = 300)</th>
<th>0-3.0 (n = 113)</th>
<th>3.5-6.5 (n = 131)</th>
<th>7.0-8.5 (n = 56)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIP Total</td>
<td>23.3 (13.3)</td>
<td>14.4 (10.5)</td>
<td>25.2 (10.8)</td>
<td>36.7 (10.3)</td>
</tr>
<tr>
<td>SIP Physical</td>
<td>22.6 (15.9)</td>
<td>9.5 (8.3)</td>
<td>25.0 (11.6)</td>
<td>43.2 (10.8)</td>
</tr>
<tr>
<td>SIP Psychosocial</td>
<td>20.6 (15.1)</td>
<td>16.5 (14.6)</td>
<td>21.4 (14.8)</td>
<td>26.8 (14.6)</td>
</tr>
<tr>
<td>SF-36 PCS</td>
<td>24.7 (10.4)</td>
<td>41.4 (9.9)</td>
<td>32.1 (8.9)</td>
<td>27.1 (5.6)</td>
</tr>
<tr>
<td>SF-36 MCS</td>
<td>47.9 (11.6)</td>
<td>46.7 (10.8)</td>
<td>47.8 (11.7)</td>
<td>50.5 (12.7)</td>
</tr>
<tr>
<td>FIS Total</td>
<td>39.1 (15.9)</td>
<td>37.0 (16.5)</td>
<td>40.5 (15.3)</td>
<td>40.4 (16.1)</td>
</tr>
<tr>
<td>FIS Physical</td>
<td>20.4 (7.4)</td>
<td>18.8 (7.4)</td>
<td>21.7 (7.0)</td>
<td>20.8 (8.0)</td>
</tr>
<tr>
<td>SSS-Tan</td>
<td>74.6 (24.6)</td>
<td>73.3 (26.1)</td>
<td>72.5 (24.8)</td>
<td>82.0 (19.4)</td>
</tr>
</tbody>
</table>

*All values are expressed as mean (SD). EDSS indicates Expanded Disability Status Scale; SIP, Sickness Impact Profile; SF-36, Medical Outcomes Study 36-Item Short-Form Health Survey; PCS, Physical Component Score; MCS, Mental Component Score; FIS, Fatigue Impact Scale; and SSS-Tan, Tangible Support subscale of the Social Support Survey.

The relationship between the Multiple Sclerosis Functional Composite (MSFC) and Expanded Disability Status Scale (EDSS) scores, ± SD. The Spearman correlation coefficient (~0.80) indicates a strong correlation between the 2 measures. The dash at 4.5 indicates that there were no subjects with this EDSS score.
measures within the MSFC. As indicated in the Figure, MSFC scores varied somewhat at each EDSS level, indicating that there is information within the MSFC scores that is not accounted for by the EDSS score.

**Table 4** shows the correlations between MSFC scores and HRQoL measures as well as correlations between HRQoL measures and EDSS scores. For all but 2 of the HRQoL measures, correlations with MSFC scores were statistically significant in the predicted direction. The MSFC correlations were stronger with the SIP than with the SF-36. This may be explained by a greater number of items in the SIP, which provides a more detailed assessment, or by greater focus on physical functioning in the SIP compared with the SF-36. While the correlations between the MSFC and HRQoL are similar to the correlations between the EDSS and HRQoL, examination of the partial correlations of the MSFC and HRQoL, which control for the correlation between the MSFC and EDSS, indicates that the MSFC is unique in its correlations with the HRQoL.

Correlations between the MSFC and SIP Psychosocial scores and between the MSFC score and employment status were stronger than predicted. Because these scales reflect the impact of disease on ability to perform routine daily activities, it appears that the MSFC measures dimensions of the disease with relevance for daily functioning.

There was not a statistically significant difference between the SF-36 MCS and MSFC scores or between the perceived need for tangible social support and MSFC score (P > .05)—in the latter case, presumably because the SSS-Tan scores increased only in the subgroup with the highest EDSS scores. Also, there were only weak correlations between the MSFC and fatigue scores, probably because fatigue is not measured as a component of the MSFC.

Table 4 also shows a correlation analysis controlling for the EDSS. There were significant residual correlations between the MSFC and 5 of the MS QoL measures included in the study. In particular, there were highly significant residual correlations with the SIP total scores. This indicates that the MSFC explains variance in the SIP total score not accounted for by the EDSS scores.

**Table 5** shows correlations between the MSFC and the MS QoL measures for each of the 3 disability severity subgroups. Generally, correlations were larger with smaller P values in the 2 subgroups with less severe disability compared with the subgroup with the most severe disability. As with the whole group analysis, correlations were strongest with SIP scores and employment status. Interestingly, small but statistically significant correlations were observed between the MSFC and fatigue scores in the subgroup with the least disability.

The lack of correlation between MSFC scores and employment status was expected in the subgroup with the most disability, as there were very few employed patients in this subgroup. However, we found that the correlation between MSFC scores and employment status was stronger in the group with midrange disability compared with the group with lower disability. This finding is unexplained.

**COMMENT**

In its published recommendations, the NMSS task force called for further prospectively designed studies of the MSFC to explore its significance and to quantify its value as a clinical trial outcome measure. The results of this cross-sectional comparison of the MSFC with HRQoL and the EDSS support the clinical relevance of the MSFC.

Significant correlations between MSFC scores and MS QoL measures for the total study population indi-
cated that the MSFC reflects the severity of MS as perceived by patients. The MSFC scores best reflect general physical well-being, as indicated by strong correlations with the SIP physical scores. However, the study also demonstrated significant correlations with measures of psychosocial functioning, suggesting that MSFC scores capture important psychosocial consequences of physical impairments. Finally, the study demonstrated moderate correlations between MSFC scores and employment status, indicating that the MSFC assesses dimensions relevant to everyday functioning.

Results of this study could direct future improvements in the MSFC measurement method in several ways. As an example, lack of correlation with a disease-specific measure of fatigue used in this study suggests that including a fatigue measure in the MSFC might improve overall correlation with disease as perceived by the patients. Furthermore, correlations between MSFC and HRQoL measures were lower in the high-disability subgroup. This may be explained by inability of the patients in this subgroup to perform the walking task that is included in the MSFC. Each patient in this subgroup was assigned a constant severe score for the walking component of the MSFC, which had the effect of truncating biological variability relating to leg function in this subgroup. This may have accounted for the lower correlations with HRQoL measures. Adding an informative measure of lower-extremity function for nonambulatory patients would probably improve correlations with HRQoL measures in patients with severe impairment. Finally, as noted by the task force, it may be useful to include a visual score.1,3

A limitation of this study is the lack of a universally accepted criterion standard for MS disease severity. In this study, the Kurtzke EDSS was used because it is a widely accepted disease severity measure. However, the EDSS has been criticized for being imprecise, using an ordinal scoring system, being insensitive to change over time, and not reflecting important components of the disease, such as arm and cognitive function. However, in the absence of a true criterion standard, the EDSS remains a reasonable comparison measure. Similarly, even though HRQoL measures are useful—their relevance to the patient is self-evident—they do not serve as a criterion standard. Therefore, this study does not resolve the question of the relative value of the MSFC scores compared with other disease measures in actually quantifying disease severity. Rather, it demonstrates significant correlations between MS QoL measures and the MSFC, and it shows significant residual correlations between MS QoL measures and the MSFC after controlling for EDSS scores. These findings demonstrate the clinical relevance of MSFC scores across the range of MS disease severity. In this regard, the study strongly supports the use of the MSFC as a clinical trials outcome measure because of the robust correlations with physician-derived and patient-derived comparison measures. Further validation of the MSFC using longitudinal methods will clarify the significance of MSFC changes and determine its utility as a primary outcome measure in clinical trials.

Accepted for publication January 28, 2000.

Corresponding author: Deborah M. Miller, PhD, Mellen Center for Multiple Sclerosis Treatment and Research, U-10, Cleveland Clinic Foundation, 9500 Euclid Ave, Cleveland, OH 44195.

REFERENCES