

PAPER

Psychometric evaluation of the multiple sclerosis impact scale (MSIS-29) for proxy use

F A H van der Linden, J J Kragt, M Klein, H M van der Ploeg, C H Polman, B M J Uitdehaag

J Neurol Neurosurg Psychiatry 2005;76:1677–1681. doi: 10.1136/jnnp.2005.065227

See end of article for authors' affiliations

Correspondence to:
Ms F A H van der Linden,
Department of Neurology,
VU University Medical
Centre, Amsterdam,
Netherlands, PO Box
7057, 1007 MB
Amsterdam, The
Netherlands; fah.
vdlinden@vumc.nl

Received 9 February 2005
In revised form 4 May 2005
Accepted 5 June 2005

Background: There may be difficulties in the use of self report measurements in patients with cognitive impairment or serious mood disturbances which interfere with reliable self assessment, as may be the case in multiple sclerosis (MS). In such cases proxies may provide valuable information. However, before using any questionnaires in a proxy sample, the questionnaire should be evaluated for proxy use.

Objective: To evaluate the psychometric properties of the 29 item Multiple Sclerosis Impact Scale (MSIS-29) when used by proxies of MS patients.

Methods: A sample of 62 partners of MS patients completed the MSIS-29. The data were evaluated for the psychometric criteria of the MSIS-29, including data quality, scaling assumptions, acceptability, reliability, validity, and responsiveness.

Results: Psychometric evaluation was satisfactory; data quality was high, and scaling assumptions and acceptability were good. Reliability was high ($\alpha > 0.80$). Findings were consistent with results of a psychometric evaluation in a patient sample.

Conclusions: The MSIS-29 can be used reliably in proxies of patients with MS. As a next step the relation between data obtained from patients and proxies needs to be studied, focusing on factors that may affect agreement and discrepancies.

In recent years, there has been increasing use of self report measurements for assessing quality of life, disease impact, or disability. The underlying assumption when using self report measurements is that the patient fully understands the questions and can give a reliable judgment on their situation. Problems may arise when the cognitive or communication abilities are insufficient. Moreover, emotional factors may interfere with self assessment, which will affect reliability. Therefore self report measurements may be less suitable in cases of cognitive dysfunction or severe mood disturbance. When applying self report measurements, the presence of such factors could lead to unreliable information or loss of information. It has been suggested that in these situations the use of proxies (for example, partners, relatives, or close friends) to assess the situation of the patient should be considered.¹ These considerations are relevant for multiple sclerosis (MS) because self report measures are often used in MS, and both cognitive decline and severe mood disturbances may be present during the course of the disease.^{2,3} These disorders could lead to invalid self reporting of quality of life,^{4–8} although there are also papers suggesting that in cognitively impaired and depressed patients self reports may be valid.^{9–12} Nevertheless, proxy measurements could be important in MS and useful under certain circumstances. We are not aware of any studies that have assessed systematically the value and limitations of proxy measurements in MS. Of all available self report measurements that can be used in MS, the Multiple Sclerosis Impact Scale (MSIS-29) is disease specific and has been rigorously evaluated for its psychometric properties.¹³ We therefore evaluated the use of the MSIS-29, which measures physical and psychological disease impact, in proxies. However, the MSIS-29 was not developed for use in proxies of MS patients. Thus it is essential that the questionnaire is validated first by using standard psychometric tests before using the scale to evaluate differences and agreement between patients and proxies.¹

The focus of the present study was to evaluate the psychometric properties of the MSIS-29 when used by proxies of patients with MS.

METHODS

Study sample

For this study, partners of MS patients were asked to complete the MSIS-29 as proxies. Partners were recruited in two ways. The first was through an ongoing study of MS patients at the outpatient clinic which required the presence of a healthy control. Those controls who were partners of the patients were asked to participate in the present study by completing the MSIS-29 during the initial visit and at the time of a six month follow up visit. Secondly, a group of partners of patients who visited the outpatient clinic was asked to complete the MSIS-29 twice, with a two week interval, to measure test-retest reliability. The Expanded Disability Status Scale (EDSS), the Guy's Neurological Disability Scale (GNDS), and the MS subtype were available for all patients.

The medical ethics committee of the VU University Medical Centre approved the study protocol. Informed consent was obtained from all participants.

Measures and procedures

The MSIS-29 is a 29 item measurement, which assesses the physical and psychological impact of MS on affected individuals; it is a self report questionnaire consisting of 20 and 9 items, respectively. Scores on the individual items are added and then transformed to a 0–100 scale, thereby generating two summary scores (for physical and psychological impact). Higher scores indicate worse health.¹³ For this study the Dutch version of the MSIS-29 was used, which is an in-house translation of the original English version that was subsequently validated in a large study across eight European countries.¹⁴ The partners were asked to complete the MSIS-29 after being instructed to keep the following question in mind: "How do you think the patient experiences the impact of MS on his/her life?" Standard psychometric

Abbreviations: EDSS, Expanded Disability Status Scale; ES, effect size; GNDS, Guy's Neurological Disability Scale; ICC, intraclass correlation coefficient; MS, multiple sclerosis; MSIS, Multiple Sclerosis Impact Scale

Table 1 Psychometric criteria

Psychometric property	Criteria
Data quality	Missing items <10% High percentage of computable scale scores High item test-retest reliability: $r \geq 0.50$
Scaling assumptions	Similar response option frequency distribution Skewness between -1 and +1 Similar means and standard deviations Similar and substantial item-total correlations: $r > 0.30$
Acceptability	Scores should span the full scale range Mean scores should be around the midpoint Floor and ceiling effect <20% Skewness between -1 and +1
Reliability	Cronbach's $\alpha > 0.80$ Scale test-retest reliability (ICC) >0.80
Responsiveness	Correlation between change score and transition score should be high Clinically important difference: 0.5 = small, 1.0 = moderate, 1.5 large
Validity	
Internal validity	Moderate intercorrelation: $r = 0.30-0.70$
External validity	High correlation between MSIS-29 physical scale and EDSS Low correlation between MSIS-29 psychological scale and EDSS High correlation between MSIS-29 physical scale and GNDS Low correlation between MSIS-29 psychological scale and GNDS
Group differences validity	Similar and low correlations between men and women Similar and low correlations for age
Hypotheses	Mean physical MSIS-29 score for different groups of the EDSS score: MSIS-29 score (EDSS 0.0-3.5) < MSIS-29 score (EDSS 4.0-6.0) MSIS-29 score (EDSS 0.0-3.5) < MSIS-29 score (EDSS ≥ 6.5) MSIS-29 score (EDSS 4.0-6.0) < MSIS-29 score (EDSS ≥ 6.5) Mean physical MSIS-29 score for different groups of the GNDS score: MSIS-29 score (GNDS <1.5) < MSIS-29 score (GNDS ≥ 1.5) Mean physical and psychological MSIS-29 scores for different MS subtypes: MSIS-29 score (RR) < MSIS-29 score (SP) MSIS-29 score (RR) < MSIS-29 score (PP) MSIS-29 score (SP) < MSIS-29 score (PP)

EDSS, Expanded Disability Status Scale; GNDS, Guy's Neurological Disability Scale; MS, multiple sclerosis; MSIS-29, Multiple Sclerosis Impact Scale; PP, primary progressive; RR, relapsing-remitting; SP, secondary progressive.

methods were used to evaluate the following psychometric properties: data quality, scaling assumptions, acceptability, reliability, responsiveness, and validity. Evaluation was done according to the methods used in the health technology assessment report of Hobart *et al.*¹³ Table 1 gives a summary of the psychometric criteria that were applied.¹³

Reliability

The MSIS-29 and instructions were given to the patient for the proxy to complete at home. After a two week interval the proxy received the MSIS-29 for the second time by postal survey. Test-retest reliability was calculated by intraclass correlation coefficient (ICC).

Table 2 Characteristics of proxies and patients

	Patients	Proxies
Female (n)	42	20
Age (years) (mean (SD))	48 (8.9)	49 (9.0)
Time since MS onset (years) (mean (SD))	13 (7.2)	-
Type of MS (n)		
Relapsing-remitting	28	-
Secondary progressive	21	-
Primary progressive	11	-
Other	2	-
EDSS		
0.0-3.5	17	-
4.0-6.5	32	-
≥ 6.5	10	-
Missing	3	-
GNDS		
<1.5	26	-
≥ 1.5	32	-
Missing	4	-

EDSS, Expanded Disability Status Scale; GNDS, Guy's Neurological Disability Scale; MS, multiple sclerosis.

Table 3 Psychometric properties for the physical and psychological MSIS-29 impact scales

Psychometric property	MSIS-29 physical	MSIS-29 psychological
Data quality		
Item missing data	0.5%	0.1%
Computable scale scores	93.5%	98.4%
Item test-retest reliability (ICC)	0.87	0.83
Scaling assumptions		
Item mean scores range	2.1 to 3.3	1.8 to 2.9
Item SD range	1.0 to 1.5	1.1 to 1.4
Item skewness range	-0.499 to +0.742	0.201 to +1.235
Item-total correlation range	0.55 to 0.85	0.41 to 0.86
Acceptability		
Observed scale score range	0 to 91.3	0 to 91.7
Scale score (mean (SD))	43.9 (24.2)	32.7 (23.2)
Floor/ceiling effect (%)	4.7/1.6	1.6/1.6
Skewness	-0.231	0.459
Reliability		
Cronbach's α	0.96	0.90
Scale test-retest reliability (ICC)	0.87	0.83
Responsiveness		
Correlation	0.07	0.24
Effect size	0.05	0.02
Clinically important difference:		
improved patients (n=30)	0.6	0.8
deteriorated patients (n=7)	0.2	1.1

ICC, intraclass correlation coefficient; MSIS-29, Multiple Sclerosis Impact Scale.

Responsiveness

A method to detect clinically important change over time is by comparing change scores (baseline score minus follow up score) with an external criterion of change, such as a transition question. This is also referred to as the retrospective method.^{13 16} In this study we compared the change scores of the MSIS-29 for the two domains, using the following transition question: "When you compare the health status of the patient at this moment with that of six months ago (baseline study), would you say that it is: better, the same, or worse?" Although patients did not have an intervention, change could have been induced by the passage of time.¹³ Responsiveness can then be determined in different ways. One method is to calculate the correlation between the change scores and the transition score; a high correlation indicates a greater responsiveness.¹³ Another method is by calculating the minimum clinically important difference. This is done by dividing the mean change score for improved/deteriorated patients by the mean change score of unchanged patients. A change of 0.5 is considered small, 1.0 moderate, and 1.5 large.^{13 15} Finally, effect sizes (ES) were calculated by dividing the mean change score by the standard deviation of the admission score. Effect sizes are interpreted as small at ES <0.20, medium at ES = 0.50, and moderate at ES >0.80.¹³

Validity

Internal validity was determined by calculating the inter-correlations between the physical and the psychological domain. External validity was examined by correlations between the two MSIS-29 domains and the EDSS and GNDS. Group differences validity was assessed by comparing both scales with variables such as age and sex.

Validity was also determined by comparing the mean MSIS-29 proxy score between MS subtypes and weighing against patient scores on the GNDS and the EDSS. The mean scores on the GNDS, EDSS, and the MS subtype were divided into different groups and it was hypothesised that the

corresponding mean MSIS-29 proxy scores for these different groups should differ significantly. The GNDS score was divided into two groups based on the median: ≤ 15 and >15 . The EDSS score was divided into three groups: EDSS 0.0–3.5, EDSS 4.0–6.0, and EDSS ≥ 6.5 . MS subtype was also divided into three groups: relapsing remitting (RR), secondary progressive (SP), and primary progressive (PP). Independent *t* tests were done to compare the different groups. The mean MSIS-29 proxy scores for GNDS and EDSS were compared only for the physical domain as the GNDS and the EDSS do not have a psychological domain. The MS subtype was compared for both domains, given that the subtype might influence the psychological impact of MS. Table 1 shows the hypotheses, which were defined prospectively. Correlations were calculated for age and sex.

RESULTS

In all, 62 partners were recruited as proxy for the study. Table 2 shows the characteristics of the partners and the patients.

Data quality

These results are shown in table 3. The percentage of missing data was low for both scales, and the computable scale scores were high. Scales could be calculated for at least 93.5%. Item test-retest was 0.87 for the physical domain and 0.83 for the psychological domain.

Scaling assumptions

Scaling assumptions are shown in table 3. Similar item means and standard deviations were found for both scales. Frequency distributions showed to be symmetrical for both scales. Skewness for the physical scale was between -1 and 1 , but skewness for the psychological scale was slightly out of range (+1.235). Item-total correlations for both scales were high: the physical scale ranged from 0.55 to 0.85 and the psychological scale from 0.41 to 0.86, both therefore fulfilling the correlation criterion (>0.30). There was a definite scaling

Table 4 Pearson correlation for age, sex, EDSS, and GNDS scores with the physical and psychological MSIS-29 impact scales

Variable	MSIS-29 physical	MSIS-29 psychological
Age	0.07	0.29
Sex	0.09	0.13
EDSS	0.66	0.25
GNDS	0.69	0.56
MSIS-29 psychological	0.65	1.0

EDSS, Expanded Disability Status Scale; GNDS, Guy's Neurological Disability Scale; MSIS-29, Multiple Sclerosis Impact Scale.

success of 50% and 44% for the physical and psychological scale, respectively. Consequently the criterion of 65% or larger was not satisfied.

Acceptability

The score ranged from 0–91.3 for the physical scale and 0–91.7 for the psychological scale (table 3). The physical scale was slightly skewed (–2.31). Floor and ceiling effects were small and did not exceed the predefined maximum of 20%. Scale midpoints were near the middle and standard deviations were almost equal.

Reliability

The internal consistency was high (table 3): 0.96 for the physical scale and 0.90 for the psychological scale. In all, 30 proxies completed both questionnaires for the test–retest study, which resulted in a test–retest reproducibility of 0.87 for the physical scale and 0.83 for the psychological scale.

Responsiveness

The correlations between the change score and the transition score were 0.07 for the physical scale and 0.24 for the psychological scale (table 3). The minimum clinically important change was small to moderate. The effect sizes were minimal for both domains.

Validity

Validity assessments are shown in tables 4 and 5. Internal validity was moderate, with an intercorrelation of 0.65 (table 4). The EDSS score showed a higher correlation of

0.66 with the physical domain than with the psychological domain. Correlation for GNDS with the physical domain was slightly higher (0.69) than for the EDSS score. Both MSIS-29 scales showed moderate correlations with age and sex.

Table 5 shows the hypothesised group differences with accompanying *p* values. Independent *t* tests were significant except for the difference between the mean MSIS-29 physical score (SP) and mean MSIS-29 physical score (PP). This was also true for the group differences of the mean MSIS-29 psychological scores per MS subtype.

DISCUSSION

Our aim in this study was to evaluate the psychometric properties of the MSIS-29 when used by proxies. Standard techniques were applied to evaluate data quality, scaling assumptions, acceptability, reliability, responsiveness, and validity. This was done according to psychometric criteria that were also used for the psychometric evaluation of the MSIS-29 in a patient sample.¹³ Psychometric properties were satisfactory for most of the criteria. Data quality was high. Scaling assumptions and acceptability were good, although the score did not span the full range for both domains and the physical scale was slightly skewed. High values for reliability were found. Responsiveness showed minimal effect sizes and poor correlations. This might have reflected the fact that change was caused by natural progression over time rather than by a treatment effect following an intervention. Natural progression may be harder to detect by proxies than the effect of a treatment. Taking into account the relatively small sample size we think that the responsiveness results should be interpreted with caution.

The validity of the MSIS-29 was established in different ways. Correlation between the EDSS and the physical scale was high, which could be expected as the EDSS measures disability. Correlations between the two scales and the GNDS were both moderate, which is probably explained by the fact that the GNDS contains questions on both physical and psychological topics. Correlations for both MSIS-29 scales and the variables age and sex were low, indicating that neither scale was biased by these variables. Significant differences in mean MSIS-29 scores were found for the different groups defined on the basis of GNDS, EDSS, and MS subtype. Of the 10 prior hypotheses, six were significantly confirmed, especially those referring to the physical domain. Partners of patients with an EDSS score ≥ 6.5 scored significantly higher on the MSIS-29 physical score than partners of patients with a low EDSS score of 0.0 to 3.5.

Table 5 Validity

Hypothesised group differences for corresponding MSIS-29 scores in groups of EDSS, GNDS, and MS subtype	<i>p</i> Value
EDSS	
MSIS-29 score (EDSS 0.0–3.5) < MSIS-29 score (EDSS 4.0–6.0)	0.001
MSIS-29 score (EDSS 0.0–3.5) < MSIS-29 score (EDSS ≥ 6.5)	0.000
MSIS-29 score (EDSS 4.0–6.0) < MSIS-29 score (EDSS ≥ 6.5)	0.000
GNDS	
MSIS-29 score (GNDS <15) < MSIS-29 score (GNDS ≥ 15)	0.000
MS subtype (physical)	
MSIS-29 score (RR) < MSIS-29 score (SP)	0.001
MSIS-29 score (RR) < MSIS-29 score (PP)	0.001
MSIS-29 score (SP) < MSIS-29 score (PP)	0.584
MS subtype (psychological)	
MSIS-29 score (RR) < MSIS-29 score (SP)	0.052
MSIS-29 score (RR) < MSIS-29 score (PP)	0.195
MSIS-29 score (SP) < MSIS-29 score (PP)	0.706

EDSS, Expanded Disability Status Scale; GNDS, Guy's Neurological Disability Scale; MS, multiple sclerosis; MSIS-29, Multiple Sclerosis Impact Scale; PP, primary progressive; RR, relapsing–remitting; SP, secondary progressive.

Proxies believed that patients with a high EDSS score experience a greater physical impact of MS than patients with a lower EDSS score. The same can be said about patients with a high GNDS score: partners of patients with a GNDS score >15 scored significantly higher on the MSIS-29 physical score than partners of patients with a GNDS score ≤15. This shows that differences in the EDSS score and the GNDS score are reflected in the MSIS-29 score of proxies, and that the MSIS-29 measures what it is supposed to measure when completed by proxies. The MSIS-29 physical score only differed between the subtypes RR and SP and between RR and PP, indicating that proxies do not make a distinction between the subtypes SP and PP. No significant differences were found between the MS subtypes on the psychological domain. Whether this is in line with the opinion of the patient remains to be investigated. A limitation of the present study in this respect is the lack of other quality of life measurements with psychological domains such as the 36 item short form health survey (SF-36) and the general health questionnaire (GHQ), which precludes direct comparisons. Other research showed that proxies are better at detecting the more observable domains, such as physical problems, than the subjective domains, such as psychological problems.¹⁷

Overall, the results showed resemblance with the psychometric properties when used by patients.^{13 18 19} In summary, our results indicate that the MSIS-29 also is a reliable and valid instrument when used by proxies. This creates a solid basis for further use of MSIS-29 proxy measurements in MS. As a next step the relation between data on proxy versus self reports needs to be studied, focusing on factors that may affect agreement and discrepancies.

ACKNOWLEDGEMENTS

We wish to thank all the patients and their partners who participated in this study. The MS Centre of the VU Medical Centre is partially funded by a programme grant of the Dutch MS Research Foundation.

Authors' affiliations

F A H van der Linden, J J Kragt, C H Polman, Department of Neurology, VU University Medical Centre, Amsterdam, the Netherlands

M Klein, H M van der Ploeg, Department of Medical Psychology, VU University Medical Centre

B M J Uidehaag, Department of Clinical Epidemiology and Biostatistics, VU University Medical Centre

Competing interests: none declared.

REFERENCES

- Riemsma RP**, Forbes CA, Glanville JM, *et al*. General health status measures for people with cognitive impairment: learning disability and acquired brain injury. *Health Technol Assess* 2001;**5**:1–100.
- Rao SM**, Leo GJ, Bernardin L, *et al*. Cognitive dysfunction in multiple sclerosis. I. Frequency, patterns, and prediction. *Neurology* 1991;**41**:685–91.
- Janssens AC**, van Doorn PA, de Boer JB, *et al*. Anxiety and depression influence the relation between disability status and quality of life in multiple sclerosis. *Mult Scler* 2003;**9**:397–403.
- Kenealy PM**, Beaumont GJ, Lintern T, *et al*. Autobiographical memory, depression and quality of life in multiple sclerosis. *J Clin Exp Neuropsychol* 2000;**22**:125–31.
- Hanita M**. Self-report measures of patient utility: should we trust them? *J Clin Epidemiol* 2000;**53**(5):469–476.
- Benedict RH**, Cox D, Thompson LL, *et al*. Reliable screening for neuropsychological impairment in multiple sclerosis. *Mult Scler* 2004;**10**(6):675–678.
- Maor Y**, Olmer L, Mozes B. The relation between objective and subjective impairment in cognitive function among multiple sclerosis patients—the role of depression. *Mult Scler* 2001;**7**:131–5.
- Goverover Y**, Chiaravalloti N, DeLuca J. The relationship between self-awareness of neurobehavioral symptoms, cognitive functioning, and emotional symptoms in multiple sclerosis. *Mult Scler* 2005;**11**:203–12.
- Marrie RA**, Miller DM, Chelune GJ, *et al*. Validity and reliability of the MSQOL in cognitively impaired patients with multiple sclerosis. *Mult Scler* 2003;**9**:621–6.
- Benedict RH**, Fishman I, McClellan MM, *et al*. Validity of the Beck Depression Inventory—Fast Screen in multiple sclerosis. *Mult Scler* 2003;**9**:393–6.
- Benedict RH**, Munschauer F, Linn R, *et al*. Screening for multiple sclerosis cognitive impairment using a self-administered 15-item questionnaire. *Mult Scler* 2003;**9**:95–101.
- Gold SM**, Schulz H, Monch A, *et al*. Cognitive impairment in multiple sclerosis does not affect reliability and validity of self-report health measures. *Mult Scler* 2003;**9**:404–10.
- Hobart JC**, Riazi A, Lamping DL, *et al*. Improving the evaluation of therapeutic interventions in multiple sclerosis: development of a patient-based measure of outcome. *Health Technol Assess* 2004;**8**.
- Hobart JC**, Cano S, O'Conner R, *et al*. Multiple Sclerosis Impact Scale-29 (MSIS-29): measurement stability across eight European countries. *Mult Scler* 2003;**9**(suppl 1):23.
- Norman GR**, Stratford P, Regehr G. Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach. *J Clin Epidemiol* 1997;**50**:869–79.
- Juniper EF**, Guyatt GH, Willan A, *et al*. Determining a minimal important change in a disease-specific Quality of Life Questionnaire. *J Clin Epidemiol* 1994;**47**:81–7.
- Sneeuw KC**, Sprangers MA, Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease. *J Clin Epidemiol* 2002;**55**:1130–43.
- Riazi A**, Hobart JC, Lamping DL, *et al*. Multiple Sclerosis Impact Scale (MSIS-29): reliability and validity in hospital based samples. *J Neurol Neurosurg Psychiatry* 2002;**73**:701–4.
- Riazi A**, Hobart JC, Lamping DL, *et al*. Evidence-based measurement in multiple sclerosis: the psychometric properties of the physical and psychological dimensions of three quality of life rating scales. *Mult Scler* 2003;**9**:411–19.