

The patient knows best: significant change in the physical component of the Multiple Sclerosis Impact Scale (MSIS-29 physical)

Lisa Costelloe, Killian O'Rourke, Hugh Kearney, Christopher McGuigan, Lisa Gribbin, Marguerite Duggan, Leslie Daly, Niall Tubridy, Michael Hutchinson

J Neurol Neurosurg Psychiatry 2007;**78**:841–844. doi: 10.1136/jnnp.2006.105759

See end of article for authors' affiliations

Correspondence to: Professor M Hutchinson, Newman Clinical Research Professor, Department of Neurology, St Vincent's University Hospital, Dublin 4, Ireland; mhutchin@iol.ie

Received 1 September 2006
Revised 15 January 2007
Accepted 18 February 2007
Published Online First
1 March 2007

Aim: The aims of this study were to determine the reliability, responsiveness and minimally important change score of the Multiple Sclerosis Impact Scale (MSIS)-29 physical using the Expanded Disability Status Scale (EDSS) as an anchor measure.

Methods: 214 patients with multiple sclerosis (MS) (EDSS 0–8.5) had concurrent MSIS-29 and EDSS assessments at baseline and at up to 4 years of follow-up.

Results: 116 patients had unchanged EDSS scores. Stability of the MSIS-29 physical (mean change 0.1 points) was better in the 85 patients with EDSS 0–5.0 than in the 31 patients with EDSS 5.5–8.5 in whom the MSIS-29 physical score fell by 8 points, a response shift phenomenon. A floor effect for the MSIS-29 was observed in 5% of stable patients at both time points. 98 patients experienced EDSS change with moderately strong statistically significant correlations between change scores in the EDSS and the MSIS-29 physical ($r=0.523$, $p<0.0001$). Effect sizes for MSIS-29 physical change were moderate to large. Using receiver operating characteristic curves, the MSIS-29 change score which produced a combination of optimal sensitivity and specificity was chosen for both EDSS ranges. For EDSS range 5.5–8, a change score of 8 had a sensitivity of 87% and specificity of 67%. For EDSS 0–5.0, a change score of 7 had a sensitivity of 78% and a specificity of 51%.

Conclusions: The MSIS-29 physical performs well over time, and is suitable for use in trials; a minimal change score of 8 points in the MSIS-29 is clinically significant.

In the past 10 years, patients' perspectives on the effect of multiple sclerosis (MS) on their well being has been addressed by the development of a number of disability and quality of life (QoL) self report scales.^{1–3} The properties of such new MS scales have been examined using the Expanded Disability Status Scale (EDSS)⁴ or the Multiple Sclerosis Functional Composite (MSFC)⁵ as comparators. Surprisingly, despite moderately good cross sectional correlations, the longitudinal correlations of change scores have been weak.^{6–8} Part of this difficulty might relate to relatively short periods of follow-up or differing perceptions between the patient and the neurologist as to whether change had taken place.⁶ This discrepancy might in part be accounted for by a phenomenon called "response shift" which is well recognised in QoL studies but has not been addressed in MS self report literature.⁹

The multiple sclerosis impact scale (MSIS-29) is a psychometrically designed patient reported measure of the effect of MS on activities of daily living³; it has been validated in a number of patient groups and shows responsiveness to change in patients treated with steroid therapy for relapses and in the rehabilitation setting.^{10–14} There are two parts to the scale: one part (questions 1–20) addresses physical impact and is termed MSIS-29 physical; the other part (questions 21–29) relates to psychological concerns and is termed MSIS-29 psychological. The stability, sensitivity to change and minimally important change of the MSIS-29 physical using repeated assessments in patients with accumulating disability in MS over years has not been established. Responsiveness to change is an important aspect of scale performance, particularly in its possible use as an endpoint in drug trials. The concept of minimally important change in the MSIS-29 is important in both the clinical setting and in research studies; it is necessary to know whether a

change in the MSIS-29 score reported by the patient is clinically relevant.

The aims of this study were to examine the reliability of the MSIS-29 physical over 4 years in patients with stable disease (absence of relapses and unchanged EDSS scores) and to examine the long term responsiveness to change of the MSIS-29 physical in patients with worsening EDSS scores. In addition, we aimed to determine minimally important change in the MSIS-29 physical score using the EDSS as an anchor measure over a 4 year period in patients with EDSS scores 0–8.5.

PATIENTS AND METHODS

In total, 214 patients with MS (baseline EDSS range 0–8.5; 139 had relapsing remitting MS, 68 had secondary progressive MS and 7 patients had primary progressive MS) attending the neurology outpatient clinics were assessed by four neurologists who trained together in the use of the EDSS. Data collection began in February 2001 and ended in June 2005; the shortest period between the first and last collection was 2 years and the longest was 4 years (median follow-up 2 years 9 months). The frequency of attendances varied between six monthly and yearly; 214 patients were assessed on two occasions, 104 on three, 61 on four and 33 on five. When the EDSS was determined, patients were asked to complete the MSIS-29. Data were not collected from patients during a relapse or within 3 months of a relapse. The MSIS-29 physical (questions 1–20 of

Abbreviations: EDSS, Expanded Disability Status Scale; ES, effect sizes; MS, multiple sclerosis; MSFC, Multiple Sclerosis Functional Composite; MSIS, Multiple Sclerosis Impact Scale; QoL, quality of life; ROC, receiver operating characteristic

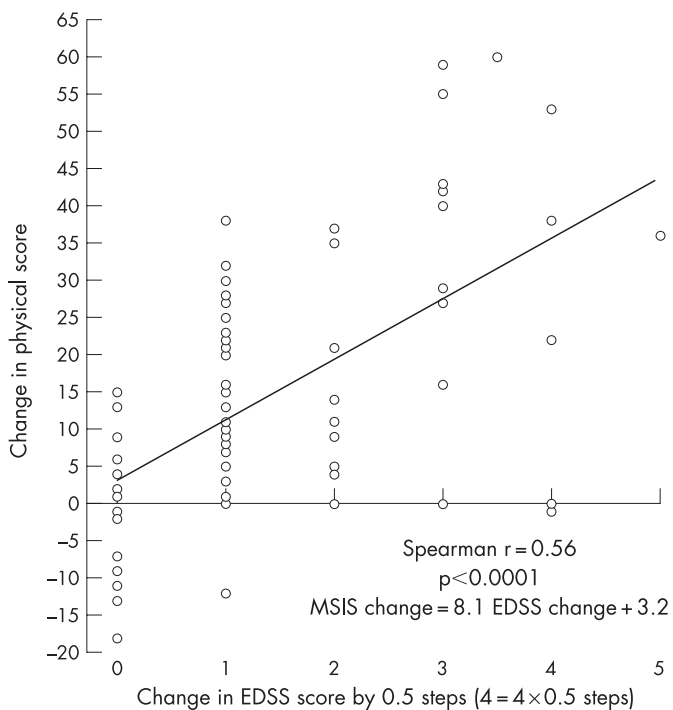


Figure 1 Correlation of change in the Multiple Sclerosis Impact Scale (MSIS)-29 physical scores in relation to change in the Expanded Disability Status Scale (EDSS) in 44 patients with an initial EDSS of 5.5 or more. There were 76 separate observations in the 44 patients. Each 1 point change on the x axis represents 0.5 change in the EDSS score. The correlation between change in the MSIS-29 physical and change in EDSS was moderate ($r=0.56$, $p<0.0001$).

the MSIS-29) was scored by the research registrar and collected prospectively with the EDSS. After data were collected for up to 4 years they were then analysed.

Analysis and statistical methods

The stability of the MSIS-29 was assessed in 116 patients with unchanged EDSS throughout the period of follow-up. First and final visit scores were used to assess stability. Responsiveness was assessed in 98 patients with EDSS change. Clinically important change in the EDSS was deemed to be 1 point change in the range 0–5.5 and 0.5 point change in the range 5.5–8.5. When a patient experienced EDSS changes over more than two assessments, the separate changes in MSIS-29 physical scores were counted—for example, a patient with three MSIS-29 scores at EDSS levels 5.5, 6.0 and 6.5 was counted as having two changes with corresponding MSIS-29 change scores from EDSS 5.5 to 6.0 and from 6.0 to 6.5. The relationship between changed EDSS scores and changed MSIS-29 physical scores was examined by the Spearman rank correlation coefficient (r). Effect sizes (ES) were calculated as mean change score divided by the SD of the initial score¹⁵ and interpreted using Cohen's arbitrary criteria (0.2 = small, 0.5 = moderate and 0.8 = large).¹⁶

In order to examine the sensitivity and specificity of change in the MSIS-29 physical score, a receiver operating characteristic (ROC) plot was calculated. Because the lower part of the EDSS measures impairments and the upper part measures ambulation limitations, the ROC characteristics of MSIS-29 physical change were examined separately for the lower and upper parts of the EDSS range. Minimally important change of the MSIS-29 physical for the two EDSS ranges was derived from the ROC curves by using the MSIS-29 change score which produced a combination of optimal sensitivity and specificity.

Accuracy and positive and negative predictive values for the change scores were calculated.

Data were analysed using Prism 4 software (GraphPad Software Inc., California, USA).

Ethics approval for this study was granted by the Research Ethics Committee of St Vincent's University Hospital, Dublin, Ireland.

RESULTS

Baseline

The 214 patients had a range of disability from EDSS 0 to 8.5; the MSIS-29 physical scores ranged from 0 to 99. At baseline, the cross sectional correlation between the MSIS-29 physical score and the EDSS was 0.79 (95% CI 0.74 to 0.83; $p<0.0001$).

Stability

There were 116 MS patients who remained relapse free on follow-up, with a stable EDSS (range 0–8.5). The group median MSIS-29 physical scores showed no significant change (first visit median: 20 (interquartile range 6–49.5); last visit median: 21 (interquartile range 9–43)). Although there was no ceiling effect, a floor effect was noted in 6 (5%) of the 116 patients at both times; all of these patients scored between EDSS 0 and 1.5 at both the first and last visit.

The stability of the MSIS-29 was better for lower grades of disability than higher grades. In the 31 stable patients with an EDSS in the range 5.5–8.5, the mean MSIS-29 physical score fell by 8 points from the first to the last assessment whereas in the 85 patients with stable EDSS in the range 0–5.0, the mean change in MSIS-29 physical was only an increase of 0.1 points. The ES in stable patients with EDSS 0–5.0 was negligible at 0.01 but for those with a stable EDSS score in the range 5.5–8.5 there was a moderate negative ES of -0.57 .

Responsiveness in the total cohort

Ninety-eight patients with a baseline EDSS of 0–8.5 experienced 198 changes in EDSS scores over the follow-up period. Using an EDSS change of 1 point from EDSS 0 to 5.0 and 0.5 points from EDSS 5.5 to 8.5 as an index of minimally important change, there was a moderately significant correlation between change in the EDSS and change in the MSIS-29 physical score ($r=0.523$, $p<0.0001$). The regression equation for this relationship was: MSIS-29 physical change score = 6.9 (95% CI 5.5 to 8.4) EDSS change score -0.21 .

Responsiveness in patients with a baseline EDSS of 5.5 or more

Forty-four patients experienced 76 changes in their EDSS (taking 0.5 point change as significant). The correlation between change in EDSS and change in MSIS-29 physical was moderate ($r=0.56$, $p<0.0001$) (fig 1). The regression equation for this relationship was: MSIS-29 physical change score = 8.1 (95% CI 5.5 to 10.7) EDSS change score $+3.2$.

Responsiveness in patients with a baseline EDSS of 0–5.0

Fifty-four patients experienced 96 changes in EDSS scores from a baseline EDSS of 0–5.0 (using an EDSS change of 1.0 point as significant). The correlation between change in EDSS and change in MSIS-29 physical was only mild ($r=0.33$, $p=0.0002$). The regression equation for this relationship was: MSIS-29 physical change score = 6.7 (95% CI 3.3 to 10.1) EDSS change score -1.8 .

Effect sizes

For the 98 patients with EDSS change, ES were large: for 1 point EDSS change (71 patients) the ES was 0.93, for 2 point

Table 1 Behaviour of change in the Multiple Sclerosis Impact Scale-29 physical score in relation to change in Expanded Disability Status Scale

	EDSS range 0–5	EDSS range 5.5–8.5
Stability: no change in EDSS effect size	0.01	–0.57*
Responsiveness		
Effect size : EDSS+1 step	0.57	1.4
Effect size : EDSS+2 steps	1.1	1.8
Effect size : EDSS+3 steps	Not available	3.4
Responsiveness : slope of linear correlation of change in MSIS-29 physical v change in EDSS	6.7 (95% CI 3.3 to 10.1)	8.1 (95% CI 5.5 to 10.7)
Correlation of change in MSIS-29 vs change in EDSS	0.33 (p<0.0002)	0.56 (p<0.0001)
Minimally important MSIS-29 physical change score	7	8
ROC: area under curve	0.725 (p<0.0001)	0.854 (p<0.0001)
Sensitivity	78%	87%
Specificity	51%	67%
Accuracy	69%	80%
Positive predictive value	50%	82%
Negative predictive value	77%	78%

EDSS, Expanded Disability Status Scale; MSIS, Multiple Sclerosis Impact Scale; ROC, receiver operating characteristic. Summary of the behaviour of the physical component of the MSIS (MSIS-29 physical) in a cohort of 214 patients followed for up to 4 years in relation to their EDSS status. The correlation of MSIS-29 change scores with EDSS change was stronger in the upper range of EDSS disability with more robust sensitivity, specificity and predictive values. However, the MSIS-29 scores in stable patients in the range EDSS 5.5–8.5 tended to improve at follow-up, consistent with the phenomenon of response shift*.

EDSS change (37 patients) the ES was 1.4 and for 3 point EDSS change (9 patients) the ES was 2.2 (table 1). For the 54 patients with EDSS change in the range 0–5.0, ES were moderate: for 1 point EDSS change the ES was 0.57 and for 2 points the ES was 1.1.

The most marked ES were seen in the 44 patients with change in the EDSS range 5.5–8.5; for 1 step of EDSS change of 0.5 (for example 6.5 to 7.0) the ES was 1.4, for 2 steps change the ES was 1.8 and for 3 steps EDSS change (for example 6.0 to 7.5) the ES was 3.4.

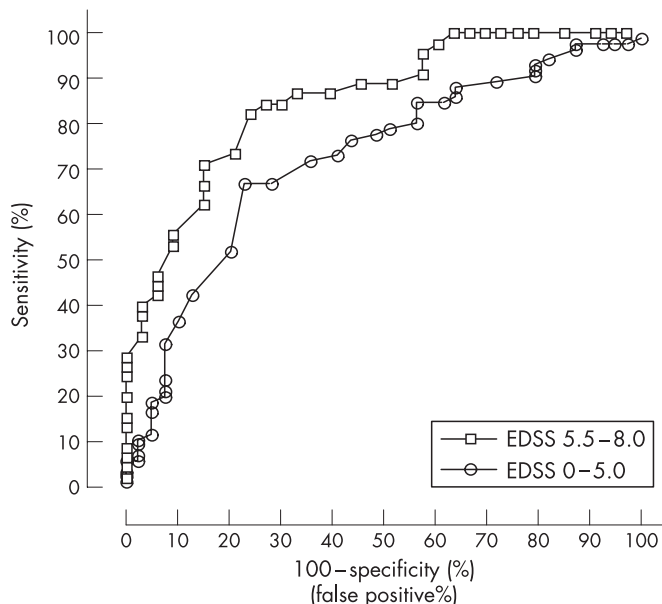


Figure 2 Receiver operating characteristics of the change in Multiple Sclerosis Impact Scale (MSIS)-29 physical scores for patients with 1 step change in Expanded Disability Status Scale (EDSS) scores compared with the change in MSIS-29 physical scores for patients with stable EDSS scores. The area under the curve (AUC) for MSIS-29 physical change for EDSS scores in the range 0 to 5.0 was 0.73 (95% CI 0.63 to 0.82) and the AUC for MSIS-29 physical change for EDSS scores in the range 5.5 to 8.0 was 0.85 (95% CI 0.77 to 0.94).

What is the minimally important change? For EDSS range 5.5–8.0

This was addressed by constructing an ROC curve comparing the changed MSIS-29 physical scores in the 33 patients who worsened by only 0.5 of an EDSS point (1 step of clinically important change) between EDSS 5.5 and EDSS 8.5 and using as controls the change in MSIS-29 physical scores in the 45 patients in the same EDSS range who experienced no worsening in the study period. The 33 patients had a mean MSIS-29 physical change score of 12.8 (95% CI 8.6 to 17.1). The ROC area under the curve was 0.854 (SE 0.42) (95% CI 0.77 to 0.94) (fig 2). From the ROC curve, a change in the MSIS-29 physical score of 8 had a sensitivity of 87% (95% CI 73 to 95) and a specificity of 67% (95% CI 48 to 82). The accuracy of a change score of 8 in this EDSS range was 80%, the positive predictive value was 82% and negative predictive value 78% (table 1).

For EDSS range 0–5.0

Similarly, an ROC curve was generated from the changed MSIS-29 physical scores in the 39 patients who worsened by 1 EDSS point between EDSS 0 to 5.0 using as controls the change in MSIS-29 physical scores in the 85 patients who experienced no worsening in the study period. The ROC area under the curve was 0.725 (SE 0.049) (95% CI 0.63 to 0.82) (fig 2). Taking a change in MSIS-29 physical score of 7 had a sensitivity of 78% (95% CI 67% to 86%) and specificity of 51% (95% CI 35% to 68%). The accuracy of a change score of 7 in this EDSS range was 69%, positive predictive value was 50% and negative predictive value was 77% (table 1).

DISCUSSION

We have demonstrated that the MSIS-29 physical has a number of characteristics and few limitations which make it suitable both as a self-report measure in long term studies of patients with MS and in phase III studies of treatment efficacy. The change in MSIS-29 physical scores showed moderately good correlation with change in the EDSS scores in the ranges 0–8.5 and 5.5–8 and weaker correlation with change in the range 0–5. In addition, the MSIS-29 physical score change demonstrated moderate to large effect sizes and importantly had moderate discriminatory power based on the ROC area under the curve of 0.85 for EDSS range 5.5–8.5. Minimally important MSIS-29

physical change scores derived from the ROC curves were similar to the slope of the regression equations derived from the linear regression analysis. Thus two independent statistical approaches to gauging a minimally important change in the MSIS-29 gave similar results. The minimally important change score for the upper part of the EDSS range showed particularly good clinimetric features.

A feature of the newer scales developed in the past 10 years is that although they purport to measure similar dimensions, when compared over several years there were poor correlations between the responsiveness of the two scales. For example, a recent study compared the change in 1 year in MSFC and EDSS scores; while the cross-sectional correlations at baseline and follow-up were moderate ($r = 0.72$), the changes in the scores only correlated weakly ($r = 0.2$).⁷ This may be because of different dynamics in the change scores. Change in the EDSS relates mainly to change in the timed walk test of the MSFC whereas the change in the MSFC relates mainly to change in the Paced Auditory Serial Addition Test and the Nine-Hole Peg Test. Similarly, a comparison of Guy's Neurological Disability Scale and the EDSS found only a weak change score relationship over 1 year ($r = 0.19$),⁶ and when the MSFC was examined with the GDNS over 2 years, there was no longitudinal relationship.⁸

A number of factors underlie the lack of longitudinal correlation between physician and patient assessments of change, including: (a) different perceptions of the patient and physician; (b) differing scale performance (change in sphincter symptoms has no effect on the upper range of the EDSS but could have a marked effect on a QoL measure); and (c) the response shift phenomenon. Response shift refers to a change in the meaning of the patient's self-evaluation of their target construct (MSIS-29 physical score) as a result of a change in the internal standards of measurement (scale recalibration), a change in the patient's values or a redefinition of the target construct (reconceptualisation).⁹ These various factors may lead to patients reassessing their perceived limitations of everyday life and resetting goals so that they may consider that the impact of their MS is less marked than they thought formerly. In this study it was noted that stable patients with significant disability (EDSS 5.5–8.0) scored significantly better at the last assessment than the first (effect size was -0.57); we would consider this an example of response shift. Response shift is an intrinsic, unavoidable and undesirable aspect of self-report measures which may limit the validity of such measures over time.

Apart from response shift, what are the other limitations of the MSIS-29 physical? Clearly, in comparison with the more disabled patient, the MSIS-29 physical in the EDSS range 0–5 appears to be less sensitive to change with smaller effect sizes, weaker correlation of change scores and a less discriminatory ROC area under the curve, resulting in a less sensitive, specific and predictive change score. This is unfortunate as it is often these less disabled, relapsing remitting patients that are recruited into clinical trials for which responsive patient report measures are needed. The weaker relationship between the two

scales in patients with low EDSS scores relates to an intrinsic defect in the construction of the EDSS. In the range 0–4.0, assessment of the EDSS is based mainly on impairments found on neurological examination and these impairments probably have little effect on the patient's perception of disease impact. For example, to move from EDSS 0 to 1.5 only requires minor sensory symptoms/signs and an extensor plantar response. In this low EDSS range, a number of patients show a floor effect in their MSIS-29 physical.

In conclusion, in this study we have established a minimally important change score for the MSIS-29 physical and demonstrated that, if validated by other groups, this patient report scale has excellent clinimetric properties which make it suitable for use in phase three studies as a secondary end point.

Authors' affiliations

Lisa Costelloe, Killian O'Rourke, Hugh Kearney, Christopher McGuigan, Lisa Gribbin, Marguerite Duggan, Niall Tubridy, Michael Hutchinson, Department of Neurology, St Vincent's University Hospital, Dublin, Ireland
Leslie Daly, School of Public Health Medicine and Population Science, University College, Dublin, Ireland

Competing interests: None.

REFERENCES

- 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.
- Sharrack B, Hughes RA. The Guy's Neurological Disability Scale (GNDS): a new disability measure for multiple sclerosis. *Mult Scler* 1999;**5**:223–33.
- Hobart J, Lamping D, Fitzpatrick R, et al. The Multiple Sclerosis Impact Scale (MSIS-29): a new patient-based outcome measure. *Brain* 2001;**124**:962–73.
- Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 1983;**33**:1444–52.
- Cutter GR, Baier ML, Rudick RA, et al. Development of a multiple sclerosis functional composite as a clinical trial outcome measure. *Brain* 1999;**122**:871–82.
- Hoogervorst EL, Eikelenboom MJ, Uitdehaag BMJ, et al. One year changes in disability in multiple sclerosis: neurological examination compared with patients self report. *J Neurol Neurosurg Psychiatry* 2003;**74**:439–42.
- Hoogervorst EL, Kalkers NF, Uitdehaag BM, et al. A study validating changes in the multiple sclerosis functional composite. *Arch Neurol* 2002;**59**:113–16.
- Hoogervorst EL, Kalkers NF, Cutter GR, et al. The patient's perception of a (reliable) change in the Multiple Sclerosis Functional Composite. *Mult Scler* 2004;**10**:55–60.
- Schwartz CE, Sprangers MAG. Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Soc Sci Med* 1999;**48**:1531–48.
- Riazi A, Hobart JC, Lamping DL, et al. The Multiple Sclerosis Impact Scale (MSIS-29): reliability and validity in hospital based samples. *J Neurol Neurosurg Psychiatry* 2002;**73**:701–4.
- Hobart JC, Riazi A, Lamping DL, et al. How responsive is the Multiple Sclerosis Impact Scale (MSIS-29)? A comparison with some other self report scales. *J Neurol Neurosurg Psychiatry* 2005;**76**:1539–43.
- McGuigan C, Hutchinson M. The multiple sclerosis impact scale (MSIS-29) is a reliable and sensitive measure. *J Neurol Neurosurg Psychiatry* 2004;**75**:266–9.
- van der Linden FA, Kragt JJ, Klein M, et al. Psychometric evaluation of the multiple sclerosis impact scale (MSIS-29) for proxy use. *J Neurol Neurosurg Psychiatry* 2005;**76**:1677–81.
- Hoogervorst EL, Zwemmer JN, Jelles B, et al. Multiple Sclerosis Impact Scale (MSIS-29): relation to established measures of impairment and disability. *Mult Scler* 2004;**10**:569–74.
- Kazis LE, Anderson JJ, Meenan R. Effect sizes for interpreting changes in health status. *Med Care* 1989;**27**(Suppl 3):S178–89.
- Cohen J. A power primer. *Psychol Bull* 1992;**112**:155–9.