J Neurol (2013) 260:2505–2510
DOI 10.1007/s00415-013-7004-1

ORIGINAL COMMUNICATION

Validating a novel web-based method to capture disease progression outcomes in multiple sclerosis

Sara Leddy · Shahrzad Hadavi · Andrew McCarren · Gavin Giovannoni · Ruth Dobson

Received: 8 April 2013 / Accepted: 9 June 2013 / Published online: 27 June 2013

Abstract The Expanded Disability Status Scale (EDSS) is the current 'gold standard' for monitoring disease severity in multiple sclerosis (MS). The EDSS is a physician-based assessment. A patient-related surrogate for the EDSS may be useful in remotely capturing information. Eighty-one patients (EDSS range 0–8) having EDSS as part of clinical trials were recruited. All patients carried out the web-based survey with minimal assistance. Full EDSS scores were available for 78 patients. The EDSS scores were compared to those generated by the online survey using analysis of variance, matched pair test, Pearson's coefficient, weighted kappa coefficient, and the intra-class correlation coefficient. The internet-based EDSS scores showed good correlation with the physician-measured assessment (Pearson's coefficient = 0.85). Weighted kappa for full agreement was 0.647. Full agreement was observed in 20 patients who had EDSS scores ranging from 0 to 6; many of those with 100 % agreement had scores of 5.5–6 ($n = 8$).The intra-class coefficient was 0.844 overall for all cases. Internet-based FS and EDSS show good agreement with physician-measured scores. Agreement was better in patients with higher scores. Overall patient satisfaction with the web-based assessment was high. An internet-based assessment tool is likely to prove an invaluable tool in the long-term monitoring in MS.

Keywords Multiple sclerosis · EDSS · Functional system · Disability assessment · Internet

Introduction

The Disability Status Scale (DSS) was first devised by Kurtzke [1] in 1955 to address the lack of a valid method for the measurement of disease progression in patients with multiple sclerosis (MS). The revised form, the Expanded Disability Status Scale (EDSS) was introduced in 1983 [2]. The EDSS is currently the gold standard method for assessing both the extent and progression of disability in patients with multiple sclerosis (MS) [2].

The EDSS is a physician-led examination which assesses eight functional systems (FS) commonly affected by MS; these are cerebellar, pyramidal, sensory, bowel and bladder, visual, brainstem, mental or cognitive function and mobility [1]. The original scale included 11 stages of disease progression, with 0 being a normal neurological examination and 10 being death due to MS; the expansion of the scale included half steps in order to increase sensitivity to changes in disease progression [3]. Grades from 0 to 3.5 reflect impairments of the FS, while midscale scores from 4.0 to 7.0 are the result of difficulties in ambulation, with less emphasis on the FS score. Higher scores reflect more severe disabilities, focussing on the need for assistance and difficulties with communication and feeding [4].

Although widely used by clinicians, the EDSS has been criticised for being physician led, difficult to reproduce and relatively insensitive to change as the disease progresses [4–10]. Several studies have examined the efficacy of patient led assessment of disease progression via telephone and self-report questionnaires. [11, 13] The need for long

Electronic supplementary material The online version of this article (doi:10.1007/s00415-013-7004-1) contains supplementary material, which is available to authorized users.

S. Leddy · S. Hadavi · A. McCarren · G. Giovannoni · R. Dobson (✉)
Barts and the London School of Medicine and Dentistry, Blizard Institute, Queen Mary University of London, 4 Newark Street, London E1 2AT, UK
e-mail: ruth.dobson@qmul.ac.uk

term follow-up in clinical trials means that an internet-based version of the EDSS is likely to provide a valuable method for assessing patients remotely. As widespread internet access increases, patients previously unable to undergo regular assessments for practical reasons can be easily accessed. An internet-based assessment tool is likely to prove an invaluable tool in the long-term monitoring in MS; not least for those patients who have difficulty travelling to see physicians regularly due to the severity of their disease.

Materials and methods

Study design

This study was an early stage, proof-of concept study. Patients who were having their EDSS measured at the Royal London Hospital, as part of clinical research protocols, were recruited. Patients were required to have clinically definite MS, be able to read English and be aged over 18. Patients were first seen by their assessing physician for EDSS; they then completed the online assessment. Demographic details including gender, age, ethnicity, year of first MS symptoms, year of diagnosis, disease modifying therapy where known (some patients were taking part in clinical trials), previous EDSS and number of relapses were also recorded. This study had ethical approval from the North London REC 2 (ref 10/H0724/27).

Both the physician assessment and online assessment were completed during the same visit. The assessing physician was aware that the web-EDSS study was being performed, but was not informed of the scores generated by the online assessment. The patient was also unaware of the scores generated by their completion of the assessment.

Description of online assessment

The web based EDSS calculator is based on the telephone-based assessment developed by Lechner-Scott et al. [11], and the interface was generated using Survey Monkey. Each patient was given a unique username during the consent process, and all questionnaires were filled in using this. A list of the questions included within the questionnaire is given in supplementary appendix 1. The questions were designed to mirror the FS scores within the EDSS, with care taken to ensure the language was suitable for a lay audience. An FS score was generated for each system using the responses gathered, and from this the overall EDSS was calculated.

As not all questions within the EDSS are relevant to all patients, and indeed some patients with early MS may find questions regarding higher EDSS steps distressing, the questionnaire ensured that not all patients were asked all questions. Care was taken to ensure that patients were not asked similar or identical questions repeatedly. This was achieved through the use of logic built into the questionnaire. This meant that each patient experienced an individualised questionnaire, and additionally ensured that patients were not asked large numbers of questions, reducing the risk of questionnaire fatigue. Access to a copy of the online questionnaire will be made available from the authors on request.

The final data gathered was in relation to the patients' experience of the online calculator and their opinions regarding the need for such a service. Patients were encouraged to leave suggestions about how the calculator could be improved and their opinions about the use of a website which may allow patients to monitor their disease progression.

Statistical analysis

Statistical analysis was performed using JMP, Graphpad (Prism 6) and PASW v18 (SPSS).

This study aimed to compare a web-based EDSS assessment with the physician led EDSS. The non-parametric Kruskal-Wallis test was used and a one-way analysis of variance (ANOVA) was performed. The Bartlett test is used to test whether the variances of the means are equal or similar. The ANOVA generally assumes that the variances are equal; hence, the Bartlett test is used to test this assumption. [15] The EDSS is an ordinal variable, thus, to measure agreement between the web-based and physician EDSS and functional system scores, the kappa coefficient and intra-class correlation coefficient were used. [10] The kappa coefficient was used to examine the degree of inter-observer agreement, and the intra-class correlation co-efficient used to demonstrate overall agreement between the two methods of assessment.

Results

Eighty-one patients were enrolled; 29 with primary progressive MS, 5 with relapsing progressive, and 47 with relapsing-remitting MS. The EDSS at the time of data collection ranged from 0 to 8. The group consisted of 49 women and 39 men; the average age was 41.5 years (range 18–68) (Table 1). EDSS from previous visit scores ranged from 0 to 8.5 (Fig. 1). The mean duration of disease was 6.6 years (range <1–29 years). All patients carried out the web-based survey with minimal assistance. Full EDSS scores were available for 78 patients.

The web-EDSS score showed good agreement with the physician (or actual) EDSS (Fig. 2). The Pearson's

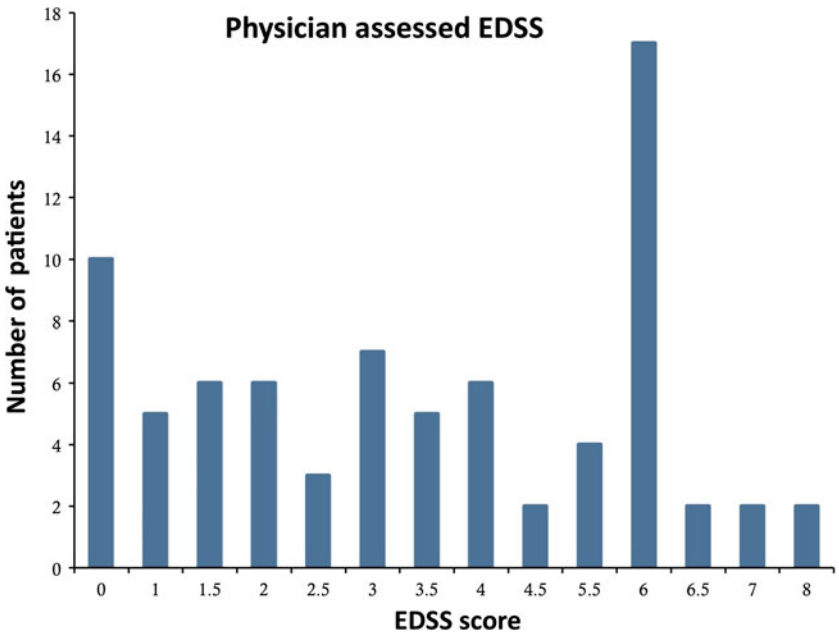

Fig. 1 Distribution of previous (*baseline*) EDSS

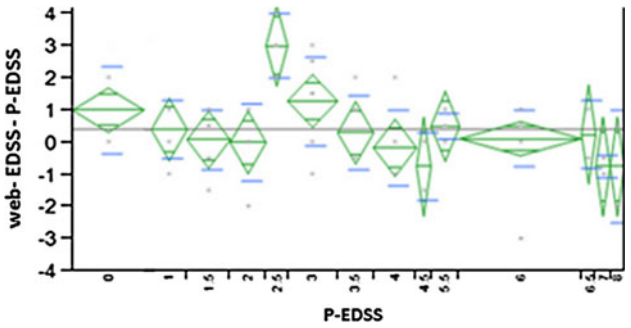

Fig. 2 The actual EDSS plotted against the difference between the means of physician or actual EDSS (P-EDSS) and web-based EDSS (W-EDSS). The midpoint of the diamonds is the mean difference between the two EDSS scores, the *upper* and *lower lines* within the diamonds are the 95 % confidence interval. The width of the diamond indicates the sample size, the *dots* the actual values. The *horizontal line* at 0.46 indicates the mean difference between the two scores. The graph indicates the greater variation at lower EDSS scores, with greater agreement at scores >5

correlation coefficient was 0.85. Figure 2 depicts the relationship between the actual (or physician-led) EDSS and the differences between the means of the actual and web-based method. One-way ANOVA showed EDSS and web assessment agreement was best at scores <2.5 and >4.5, with more than 50 % patients showing a difference of 0.5–1.

There was a significant difference in the mean results between the score categories when using ANOVA. However, when a match-paired test (Kruskal-Wallis) was applied, the result was less conclusive, with a Chi square of 0.068 indicating no significant difference between the means. Additionally, the Bartlett test demonstrated no significant difference between the variations of the scoring groups. The mean of the difference in EDSS scores was 0.46, and when the web-EDSS score was adjusted by 0.4 there was no significant difference between the means.

Weighted kappa for full agreement was 0.647. Full agreement was observed in 20 patients who had EDSS scores ranging from 0 to 6, most of those with 100 % agreement had scores of 5.5–6 ($n = 8$). EDSS scores within ± 0.5 ($n = 19$) demonstrated a weighted kappa of 0.869 representing almost perfect agreement. However, EDSS scores within ± 1.0 ($n = 15$) had a weighted kappa of 0.09, showing poor agreement. Six patients had scores with a difference of ± 1.5, seven a difference of ± 2, five a difference of ± 3 and one a difference of 4. The intra-class coefficient was 0.844 overall for all cases. It was 0.50 for patients with an EDSS <4.0 and 0.52 for EDSS >4.0. This is consistent with the findings shown in Fig. 2.

Functional system (FS) comparison was available for 52 patients. Weighted kappa values for visual, brainstem, pyramidal, cerebellar, sensory, bowel/bladder, and cerebral were 0.26, 0.238, 0.621, 0.532, 0.443, 0.586 and 0.526, respectively. The intra-class coefficient showed the best correlation in the pyramidal and bowel/bladder systems with scores of 0.78 and 0.75, respectively. Poor correlation was found for Brainstem and Visual FS (ICC 0.19 and 0.25) (Table 2).

Taking mental impairment into account, there were 35 patients with a cerebral/mental FS equal to 0, 5 patients with a score of 1, and 12 with a score of 2, as rated by their assessing physician. The ICC for patients with a score of 0 was 0.805, for those with a score of 1, 0.902 and, for those with significant mental impairment, the ICC was found to be 0.696.

Table 1 Details of the patient group on whom full demographic data were available ($n = 62$)

Demographic	Number of patients
Gender (M:F)	32:49
Age (mean; range; SD)	41.5, 24–59, 10.49
Duration of disease (since diagnosis) (mean; SD)	6.6 years; 5.6 years
Type of MS (PPMS:RPMS:RRMS)	29:5:47
EDSS score (physician measured) (mean; range)	3.49; 2.31
Level of education (GCSE:A-level:University:post grad:other)	14:8:33:3:4 (62)
Marital Status (single:partner:married:separated:divorced)	20:5:30:2:5 (62)
Ethnicity (White-British:White-Irish:White-Polish:White-Swedish:Indian:Afro-Caribbean:British-Asian:White-other:White-Black Caribbean)	50:2:2:1:2:1:1:1:1 (62)
Work status (full-time employment:part-time:unemployed:self-employed:student	40:1:19:1:1 (62)
Hours worked (mean; SD)	25.4; 18.9
Number of relapses 2 years prior to diagnosis (N/A:>5:4:3:2:1:0)	20:4:5:15:12:3: 3 (62)
Number of relapses in past 12 months (N/A:>5:4:3:2:1:0)	18: 0: 1: 1: 0: 6: 36 (62)

Feedback from patients regarding the web-EDSS was good. One hundred percent of the patients who started the questionnaire completed it. All patients were able to complete the questionnaire independently. The majority of patients reported that the time taken to complete the questionnaire was "just right", with the next most popular answer "too short", indicating that patients did not feel over-burdened by the questionnaire.

Discussion and conclusions

This study shows that the web-EDSS demonstrates good agreement with the physician-measured EDSS in several aspects of the assessment. To enable the results of this study to be compared with others done on this subject, the kappa coefficient and ICC were used. Exact agreement was seen in 25 % of the patient group (Table 2).

The weighted kappa for overall agreement was 0.647, which is considered to be good and is significantly better than the values previously seen in the study of telephone-based EDSS by Lechner-Scott et al. [11] (0.48) and in a study using self-reported EDSS via questionnaire, carried out by Cheung et al. [13] (0.43). When scores within ±0.5 EDSS steps were interpreted as being equivalent, the weighted Kappa increased to 0.869, demonstrating almost perfect agreement. It is important to consider these figures in relation to agreement observed when the EDSS is carried out by two different physicians. Amato et al. [4] investigated the agreement between two physicians carrying out EDSS assessment on patients with an EDSS range of 1.0–8.5 (i.e., a patient group comparable to ours) and found the agreement to be 0.5. When half a point variation was included the agreement was found to be 0.75. In this study the physician-measured EDSS on average gives a lower score, however, when a weighting factor of 0.4 is removed there is a less significant difference between the two methods of assessment. It is likely that data from a larger group of patients is necessary to draw more definitive conclusions on the web-EDSS as an exact comparison to the physician measured EDSS.

Weighted kappa was lowest in the visual and brainstem FS (0.26, 0.238) (Table 2). The poor agreement in visual FS has been noted by Lechner-Scott et al. and by Cheung et al. [13]. The authors suggest that the visual part of the assessment can only be carried out accurately by a physician, due to the nature of the examination. Greatest agreement was seen in the pyramidal and bowel/bladder functional systems (0.62, 0.58). The moderate agreement seen with the bowel/bladder systems is also observed by Lechner-Scott et al. (0.42) and Cheung et al. (0.57).The telephone-based study also demonstrated high levels of agreement in the pyramidal FS (0.54) but unlike the web-EDSS, the brainstem FS shows high levels of agreement (0.59). The self-report study shows less agreement in the pyramidal FS (0.34) and brainstem FS (0.31) [12].

It is likely that high levels of agreement are seen in the bowel/bladder functional system due to the fact that this component of the physician-rated EDSS is based on directly asking the patient. Considering this, perhaps higher levels of agreement should be expected. Interestingly more than 90 % of those patients whose computer scores were not equal to that assigned by their physician in the bowel/bladder functional systems had higher scores from the web-based assessment, possibly indicating that they perceived their disability to be greater in these areas than the physician. A web-based assessment may be particularly useful for patients who find it difficult to discuss problems relating to bowel, bladder and mental function. The cerebral FS in the physician-based EDSS assessment is also generated by questions directed to the patient by the physician, again indicating that one would expect higher levels of agreement. In this study the agreement was greater than that seen in that of the telephone and self-reported methods of EDSS assessment (0.526, 0.38, and 0.31, respectively).

The ICC was found to be 0.844 for the entire study group, which indicates almost perfect agreement. A slight increase in agreement was seen in patients with EDSS scores >4.5; ($n = 45$) compared to those with scores <4.5

Table 2 Kappa values, correlation and Intra-class coefficient (ICC) of functional system scores

Functional system	Weighted kappa	Standard error	Weighted kappa comment	Intra-class correlation coefficient	Correlation
Visual	0.26	0.125	Poor	0.2527	0.22149
Brainstem	0.238	0.084	Poor	0.1965	0.2779
Pyramidal	0.621	0.081	Good	0.7783	0.81624
Cerebellar	0.532	0.073	Moderate	0.6724	0.76799
Sensory	0.443	0.083	Moderate	0.5894	0.59634
Bowel/bladder	0.586	0.07	Moderate	0.7246	0.7711
Mental	0.526	0.094	Moderate	0.6088	0.67833

($n = 33$); other studies have also noted a more significant agreement in patients with greater disability. The increased agreement in those with higher scores may be due to the level of awareness patients with greater disability have about their disease, or the more clear-cut boundaries between EDSS steps.

As described by Amato et al. [4], the use of 'self-reported' or patient-based methods of EDSS assessment such as this web-based EDSS assessment tool appears to be most useful for patients who demonstrate greater levels of disability due to MS, rather than the relatively mild neurological impairment seen in the early stages of disease. It might also be a useful tool for stratifying patients according to their perceived level of disability, allowing clinicians to educate patients and to manage symptomatic aspects of disease in association with the patients understanding of their disability. A study by Van der Linden et al. [14] investigated the use of patient proxy and self-assessment with the MSIS-29. It found that on average the patient's carer or proxy viewed the patient's disability as more severe than the patient themselves did. However, the overall level of agreement was good and this may be an area which could be investigated with the web-based EDSS. It could prove useful for patients with a very severe level of disability.

As this is a pilot study of the web-based assessment tool, clearly further work is required in a larger sample size to validate the calculator. In addition its sensitivity to change has not been explored. The EDSS as a form of disease measurement has been criticised for its poor ability in this regard; whether patient-based self-assessment might alter this remains to be seen. Another potential advantage of the web EDSS is the opportunity to expand the sections of the scale that are less responsive to change, in particular in the range from 5.5 to 7.0. We are currently exploring this opportunity in addition to validating the calculator in a larger patient cohort.

Acknowledgments RD is funded by an Association of British Neurologists/MS Society of Great Britain Clinical Research Fellowship. GG receives grant support from the MRC, National MS Society, MS Society of Great Britain and Northern Ireland, AIMS2-CURE and the Roan Charitable Trust. This study received no specific funding.

Conflicts of interest SL, SH, AM and RD have no conflicts of interest to declare. GG has received research grant support from Bayer-Schering Healthcare, Biogen-Idec, GW Pharma, Merck Serono, Merz, Novartis, Teva and Sanofi-Aventis. GG has received personal compensation for participating on Advisory Boards in relation to clinical trial design, trial steering committees and data and safety monitoring committees from: Bayer-Schering Healthcare, Biogen-Idec, Eisai, Elan, Fiveprime, Genzyme, Genentech, GSK, Ironwood, Merck-Serono, Novartis, Pfizer, Roche, Sanofi-Aventis, Synthon BV, Teva, UCB Pharma and Vertex Pharmaceuticals.

Ethical standard This study had ethical approval from the North London REC 2 (ref 10/H0724/27).

References

1. Kurtzke JF (1995) A new scale for evaluating disability in multiple sclerosis. Neurology 5:580–583
2. Kurtzke JF (1983) Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 33(11):144–152
3. Kurtze JF (1965) Further notes on disability evaluation in multiple sclerosis, with scale modifications. Neurology 15:654–661
4. Amato MP, Ponziani G (1999) Quantification of impairment in MS: discussion of the scales in use. Mult Scler 5:216–219
5. Willoughby EW, Paty DW (1988) Scales for rating impairment in multiple sclerosis: a critique. Neurology 38:1793–1798
6. Noseworthy JH, Vandervoort MK, Hopkins M, Ebers GC (1989) A referendum on clinical trial research in multiple sclerosis: the opinion of the participants at the Jekyll Island workshop. Neurology 39:977–981
7. Weiner HL, Paty DW (1989) Diagnostic and therapeutic trials in multiple sclerosis: a new look. Summary of Jekyll Island workshop. Neurology 39:972–976
8. Sharrack B, Hughes RA, Soudain S, Dunn G (1999) The psychometric properties of clinical rating scales used in multiple sclerosis. Brain 122:141–159
9. Hobart J, Freeman J, Thompson A (2000) Kurtzke scales revisited: the application of psychometric methods to clinical intuition. Brain 123:1027–1040

10. Portney LG, Watkins MP (2000) Foundations of clinical research applications to practice. Prentice Hall Inc., New Jersey, ISBN 0-8385-2695-0, p 560–567
11. Lechner-Scott J et al (2003) Can the Expanded Disability Status Scale be assessed by telephone? Mult Scler 9(2):154–159
12. Bowen J, Gibbons L, Gianas H, Kraft GH (2001) Self-administered Expanded Disability Status Scale with functional system scores correlates well with a physician-administered test. Mult Scler 7:201–206
13. Cheng EM, Hays RD, Myers LW, Ellison GW, Beckstrand M, Vickrey BG (2001) Factors related to agreement between self-reported and conventional Expanded Disability Status Scale (EDSS) scores. Mult Scler 7:405–410
14. Van der Linden FAH, Kragt JJ, Van Bon M, Klein M, Thompson AJ, Van der Ploeg HM, Polman CH, Uirdehaag BMJ (2008) Longitudinal proxy measurements in multiple sclerosis: patient-proxy agreement on the impact of MS on daily life over a period of 2 years. BMC Neurol 8:2
15. Snedecor GW, Cochran WG (1989) Statistical methods. Iowa State University Press