Minimal clinically important difference of improvement on the Arm Function in Multiple Sclerosis Questionnaire (AMSQ)

Caspar EP van Munster, Levent Kaya, Morgan Obura, Nynke F Kalkers and Bernard MJ Uitdehaag

Abstract

Background: The Arm Function in Multiple Sclerosis Questionnaire (AMSQ) has been developed to assess upper extremity function of patients with multiple sclerosis (MS). A minimal clinically important difference (MCID) value has not been determined yet.

Objective: The objective of this study is to determine an MCID for AMSQ.

Methods: We used the sensitivity- and specificity-based approach with dichotomized global perceived effect as an anchor.

Results: The receiver operating characteristic (ROC) curve yielded an optimal threshold value of 14.5 (sensitivity 0.68 and specificity 0.79). The area under the ROC curve value was 0.77.

Conclusion: We identified an MCID of 15 points for the AMSQ (range 31–186).

Keywords: Multiple sclerosis, upper extremity function, minimal clinically important difference, patient-reported outcome measure

Introduction

Traditionally, the focus of clinical assessment in multiple sclerosis (MS) has been on ambulation. However, other domains are being increasingly assessed in conjunction with ambulation. This includes the assessment of upper extremity function (UEF). Various measures are available of which the 9-hole peg test (9HPT) is considered as the gold standard for manual dexterity.1 Nevertheless, a performance-based measure, such as the 9HPT, does not provide any insight into the patient perspective of UEF. For this purpose, patient-reported outcome measures (PROMs) are valuable tools. To date, only one PROM is available that has been specifically developed to assess UEF in MS patients: the Arm Function in Multiple Sclerosis Questionnaire (AMSQ).2 The AMSQ is a unidimensional 31-item questionnaire with good psychometric properties.3

However, a minimal clinically important difference (MCID) has not been reported yet. An MCID defines the smallest amount of change on a scale that is important or meaningful to a patient.4 Determining an MCID of a PROM is important because a given change on the score generally does not have an obvious clinical importance to the clinician. The objective of this study is to determine an MCID of improvement on the AMSQ.

Methods

Data were derived from patients who have been treated with fampridine. Fampridine increases axonal conduction velocity by selectively blocking potassium channels, which may lead to improvement of various motor functions, including UEF.5 Effects generally occur within 2 weeks of treatment. Therefore, patients treated with fampridine are good subjects to assess change in AMSQ and determine an MCID.

Patients

Patients were recruited in the VU Medical Center in Amsterdam from an outpatient clinic that was specifically organized to assess eligibility for and efficacy of
treatment with fampridine. All patients provided written informed consent prior to inclusion, and the study was approved by the local ethics committee. Patients were considered eligible if they complied with the official treatment label of fampridine. Demographical data and MS characteristics, including an Expanded Disability Status Scale (EDSS), were collected for this study.

AMSQ
Patients were asked to complete the AMSQ before treatment and during a follow-up visit after a minimum of 2 weeks of treatment. The AMSQ consists of 31 items concerning activity limitations due to hand and arm functioning. A patient assigns a number to each item on a 6-point Likert-type scale ranging from “not at all” to “no longer able to.” The sum score ranges from 31 to 186, with a higher score indicating more impairment. Change in the sum score was calculated by subtracting the AMSQ score of the follow-up visit from the baseline value. Consequently, a positive change score indicates an improvement of UEF, and conversely a negative change score indicates worsening of UEF. Questionnaires with more than two missing items were excluded from analysis. If one or two items were missing, the average of the other items was calculated and used as substitutes.

Determining the MCID
We used the sensitivity and specificity anchor-based method to determine an MCID. In short, with an anchor-based approach, the change in PROM score is being compared with change in another measure that is understandable and is considered as an anchor or external criterion. As anchor, we used a global perceived effect (GPE) score that specifically addressed change in UEF and consisted of a 5-point Likert-type scale, including “much deteriorated” (1), “deteriorated” (2), “unchanged” (3), “improved” (4), and “much improved” (5). The GPE was determined by the treating physician on the follow-up visit by asking the patient how much the UEF was changed since the baseline visit. Because we used the sensitivity and specificity approach, the GPE scores were dichotomized into “improved” or “unchanged.” We wanted to address improvement of UEF, so we excluded scores 1, 2, and 5 to minimize the impact of these scores on the MCID value. With this method the GPE is considered the gold standard and the AMSQ as a diagnostic test for which the sensitivity and specificity is discriminated between “improved” and “unchanged.” A receiver operating characteristic (ROC) curve was used to determine the MCID, that is, the AMSQ score that produces the greatest combined sensitivity and specificity (determined with the highest Youden’s index). In addition, the area under the ROC curve (AUROC) was determined. This value represents the probability that scores will correctly discriminate between “improved” and “unchanged” UEF. A value of 0.7 to 0.8 was considered acceptable and 0.8 to 0.9 excellent. The correlation between AMSQ change and GPE was determined using Spearman’s rank-order correlation statistics.

Results
Data from 223 patients were analyzed. The mean age was 51.3 years (standard deviation 10.5), with 57.4% females. Most patients had a progressive disease type (56.5%). The median (interquartile range) for disease duration was 11.4 years (4.4–16.6) and for EDSS was 6.0 (4.0–6.5). The correlation coefficient between AMSQ and GPE was 0.37 ($p < 0.001$). The AMSQ thresholds from the ROC curve with corresponding sensitivity and specificity are displayed in Table 1. A threshold value of 14.5 yielded the highest sensitivity (0.68) and specificity (0.79). The ROC curve is shown in Figure 1. The AUROC value was 0.77. The SEM was 13.0.

Discussion
We found an MCID value of 14.5 with a sensitivity of 0.68, a specificity of 0.79, and an acceptable AUROC value. Since the AMSQ has no decimals, we rounded the threshold to 15 points.

There is no consensus about the preferred threshold value that determines sensitivity and specificity. Mostly, the threshold is chosen that jointly maximizes sensitivity and specificity in order to have the lowest overall misclassification. We used this rationale to determine the threshold.

There is a certain degree of uncertainty in our findings. This is reflected in a weak, albeit sufficient, correlation between AMSQ and GPE, a moderate sensitivity and specificity and rather similar threshold values around the MCID value. Therefore, our findings will have to be confirmed in future studies.
This is the first study to determine an MCID value for the AMSQ. The strength of our study lies in the large sample size and the normal distribution of AMSQ results. However, our study also has some limitations. First, we used only one method to determine an MCID, while there are others available. Second, we used a subjective anchor that is prone to recall bias in which case a patient may have recalled answers given at baseline that have subsequently influenced completion of the questionnaire at follow-up. Therefore, an MCID should also be assessed with an objective anchor for UEF, such as the 9HPT. Finally, a placebo effect might have influenced our findings. This accounts particularly for the GPE, since a patient may have experienced improvement of UEF, while no improvement has been noticed on the ability to perform a task of UEF (as assessed with AMSQ). Inclusion of other measures that assess capacity of UEF objectively, such as the 9HPT, allows more certainty and magnitude of true improvement of UEF. Furthermore, additional objective measures contribute to a more detailed description of UEF of a group of patients.
In conclusion, our MCID estimate for AMSQ is 15 points based on a sensitivity and specificity anchor-based method. Future studies should investigate reproducibility of this finding with similar and other methods, in a cohort with extensive assessment of different domains of UEF.

Declaration of Conflicting Interests
The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: C.E.P.v.M. has received travel support from Novartis Pharma AG, Sanofi Genzyme, and Teva Pharmaceuticals and honoraria for lecturing and consulting from Biogen-Idec and Merck Serono. L.K. and M.O. have no conflicts of interest. N.F.K. has received speaker and consultancy fees from Biogen, Teva, Genzyme, Roche, and Novartis. B.M.J.U. has received consultancy fees from Biogen-Idec, Genzyme, Merck Serono, Novartis, Roche, and Teva.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD
Caspar EP van Munster  https://orcid.org/0000-0001-8701-5413

References
1. van Munster CE and Uitdehaag BM. Outcome measures in clinical trials for multiple sclerosis. CNS Drugs 2017; 31(3): 217–236.
2. Mokkink LB, Knol DL, van der Linden FH, et al. The Arm Function in Multiple Sclerosis Questionnaire (AMSQ): development and validation of a new tool using IRT methods. Disabil Rehabil. Epub ahead of print 26 July 2015. DOI: 10.3109/09638288.2015.1027005.
3. van Leeuwen LM, Mokkink LB, Kamm CP, et al. Measurement properties of the Arm Function in Multiple Sclerosis Questionnaire (AMSQ): a study based on Classical Test Theory. Disabil Rehabil 2017; 39(20): 2097–2104.
4. McGlothlin AE and Lewis RJ. Minimal clinically important difference: defining what really matters to patients. JAMA 2014; 312(13): 1342–1343.
5. Savin Z, Lejbkowicz I, Glass-Marmor L, et al. Effect of Fampridine-PR (prolonged released 4-aminopyridine) on the manual functions of patients with multiple sclerosis. J Neurol Sci 2016; 360: 102–109.
6. https://www.ema.europa.eu/documents/product-information/fampyra-epar-product-information_en.pdf
7. Copay AG, Subach BR, Glassman SD, et al. Understanding the minimum clinically important difference: a review of concepts and methods. Spine J 2007; 7(5): 541–546.
8. de Vet HC, Terluin B, Knol DL, et al. Three ways to quantify uncertainty in individually applied “minimally important change” values. J Clin Epidemiol 2010; 63(1): 37–45.
9. Deyo RA and Centor RM. Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. J Chronic Di. 1986; 39(11): 897–906.
10. Hosmer DW. Applied logistic regression. New York: John Wiley & Sons, 2000.
11. Stratford PW and Riddle DL. When minimal detectable change exceeds a diagnostic test-based threshold change value for an outcome measure: resolving the conflict. Phys Ther 2012; 92(10): 1338–1347.