

## **The measure tells the tale**

### *Clinical outcome measures in Multiple Sclerosis*

*Hanneke E. Hulst<sup>1</sup>, Alan J. Thompson<sup>2</sup>, Jeroen J. G. Geurts<sup>1</sup>*

<sup>1</sup> Department of Anatomy and Neurosciences, VU University Medical Center, MS center Amsterdam, Amsterdam Neuroscience, Amsterdam, The Netherlands

<sup>2</sup> UCL, London, United Kingdom

Clinical trials and research projects depend on valid and reliable outcome measures.<sup>1</sup> Here, there are several important questions, such as: is the measure validated for use in an MS population? Does the measure reflect the function that we are interested in? Is it reliable over time and between different raters? Moreover, researchers are increasingly interested in so-called “clinically meaningful” improvement on these measures, reflective of the (perceived) day-to-day functioning of patients. In other words, we are in great need for solid, robust and clinically meaningful measures in order to ‘tell the tale’ in clinical and therapeutic practice.

While a (justified) critical attitude exists when it comes to novel measures materializing in the field of MS, we are relatively comfortable using measurements that have been around for a longer time. However, the question is whether these measures are accepted in the field simply because they have stood the test of time or because they have outstanding psychometric properties and are the best within their category?

This question formed the incentive for the MS Outcome Assessment Consortium (MSOAC)<sup>1-4</sup> to write four in-depth reviews on some of the most commonly and widely

used outcome measures in MS research: the Symbol Digit Modalities Test (SDMT) for measuring information processing speed,<sup>2</sup> the 9-hole peg test (9-HPT) for measuring manual dexterity performance,<sup>3</sup> the Timed-25-Foot Walk (T25FW) as a metric of ambulation,<sup>4</sup> and the Sloan low-contrast letter acuity (LCLA) test for measuring visual performance.<sup>5</sup>

The overall aim of MSOAC is to adopt a clinical outcome assessment tool for clinical trials to better capture MS-related disability. The four companion papers in the current issue of *Multiple Sclerosis Journal* consistently and comprehensively describe the psychometric characteristics of the different outcome measures, their clinical relevance and they touch upon clinically meaningful change.

A great advantage of the outcome measures discussed, is that they are fast, inexpensive, and relatively easy to administer. From that point of view, they are ideal candidates for use in clinical and scientific settings. The SDMT, which is already incorporated in several internationally renowned cognitive test batteries<sup>6-8</sup>, has an excellent ability to differentiate patients from healthy participants, a high test-retest reliability and sensitivity to detect changes over time using alternating but equivalent versions of the test. A score of 3 to 4 points difference on the SDMT can predict employment versus unemployment status, which can certainly count as a clinically meaningful change.

The 9-HPT is currently regarded as the gold standard for manual dexterity in patients with MS due to its outstanding psychometric properties and serves as a reference point for other upper limb outcome measures. With caution, a 20% decrease in 9-HPT score is associated with clinically meaningful worsening, but this should be further explored in future studies. The latter holds true for the T25FW test, as well: 20% change in test score is associated with a significant worsening on self-reported

physical health. An important note is that the T25FW is especially reliable and valid in a subset of MS patients (EDSS <6.5).

Practical constraints such as fading of letters, backlighting, learning effects and the inability to recognize letters, make the LCLA is more difficult to work with. Nevertheless, LCLA is better at differentiating patients from healthy participants than high-contrast visual acuity tests and once the circumstances are ideal, the LCLA has a high reliability and validity. A seven point difference on the LCLA was suggested to be clinically meaningful by the MSOAC team.

From the current MSOAC work by Benedict, Motl, Feys, Balcer and colleagues, it becomes clear that all four outcome measures (SDMT, 9-HPT, T25FW, LCLA) show firm reliability and validity for patients with MS. An encouraging outcome. Yet, it does not necessarily mean that 'the measure can tell the tale'. As the authors mention, it is important to realize that one single measure of a specific function might not always be sufficient to explain every aspect of this particular function. For example, the T25FW test is described as 'the best characterized objective measure of walking disability', however it has a ceiling effect for patients with an EDSS of >6.5 and says something about speed of walking but not per se about the quality of gait.<sup>9</sup> This means that there will always be a need for additional outcome measures testing related aspects of the same construct with at least similar psychometric properties, in order to give the best reflection of the patient's functioning.

We are able to tell the 'statistical tale'. But there is more to pursue: in this day and age, where personalized medicine and patient-tailored care are central concepts, we need to identify changes on these outcome measures that have a clinical impact on our patients' lives (e.g. what does a significant two-point improvement on the SDMT mean on a day-to-day basis?) For all four studied outcome measures, the first steps

towards defining clinically meaningful change have been taken. These steps need to be validated and discussed by experts in the field, but the foundations have been laid to enable the measures tell a broader tale.

## References

1. Cohen JA, Reingold SC, Polman CH, Wolinsky JS; International Advisory Committee on Clinical Trials in Multiple Sclerosis. Disability outcome measures in multiple sclerosis clinical trials: current status and future prospects. *Lancet Neurol*. 2012 May;11(5):467-76.
2. Benedict RH, Deluca J, Phillips G, LaRocca N, Hudson L, Rudick R. Validity of the Symbol Digit Modalities test as a cognition performance outcome measure for multiple sclerosis. *MSJ*.
3. Feys, P, Lamers I, Francis, G, Benedict RH, Phillips G, LaRocca N, Hudson L, Rudick R. The Nine-Hole Peg Test as a manual dexterity performance measure for multiple sclerosis. *MSJ*.
4. Benedict RH, Deluca J, Phillips G, LaRocca N, Hudson L, Rudick R. Validity of the Timed 25-Foot Walk as an ambulatory performance outcome measure for multiple sclerosis. *MSJ*.
5. Balcer L, Raynowska J, Nolan R, Galetta S, Kapoor R, Benedict RH, Phillips G, LaRocca N, Hudson L, Rudick R. Validity of low-contrast letter acuity as a visual performance outcome measure for multiple sclerosis. *MSJ*.
6. Rao SM. A manual for the Brief Repeatable Battery of Neuropsychological Tests in multiple sclerosis. Medical College of Wisconsin: Milwaukee, WI. 1990.

7. Benedict RH, Cookfair D, Gavett R, Gunther M, Munschauer F, Garg N, Weinstock-Guttman B. Validity of the minimal assessment of cognitive function in multiple sclerosis (MACFIMS). *J Int Neuropsychol Soc.* 2006 Jul;12(4):549-58.
8. Langdon DW, Amato MP, Boringa J, Brochet B, Foley F, Fredrikson S, Hämäläinen P, Hartung HP, Krupp L, Penner IK, Reder AT, Benedict RH. Recommendations for a Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS). *Mult Scler.* 2012 Jun;18(6):891-8.
9. Kempen JC, Doorenbosch CA, Knol DL, de Groot V, Beckerman H. Newly Identified Gait Patterns in Patients With Multiple Sclerosis May Be Related to Push-off Quality. *Phys Ther.* 2016 Nov;96(11):1744-1752.