Title
Has the time come to revisit our standard measures of disability progression in MS?

Authors & affiliations
Tomas Kalincik, MD, PhD; CORe, Department of Medicine, University of Melbourne, Melbourne, Australia; Melbourne MS Centre, Department of Neurology, Royal Melbourne Hospital, Melbourne, Australia
Maria Pia Sormani, PhD; Department of Health Sciences (DISSAL), Biostatistics Unit, University of Genoa, and Ospedale Policlinico San Martino-IRCCS, Genoa, Italy
Carmen Tur, MD, PhD; Queen Square Multiple Sclerosis Centre, Department of Neuroinflammation, UCL Institute of Neurology, University College London, London, UK; Multiple Sclerosis Centre of Catalonia, Neurology/Neuroimmunology Department, Vall d'Hebron University Hospital, Barcelona, Spain

word count: 979

Deleted: 36

Editorial

For decades progression of disability in MS was assessed by a confirmed increase in Expanded

Disability Status Scale (EDSS). It is the typical primary endpoint in phase 3 clinical trials in

progressive MS, where the lack of effect of most immunotherapies on confirmed disability

progression is usually interpreted as the lack of protection from cumulative damage to the

CNS.

Deleted: Confirmed disability progression

Deleted: Is it possible that confirmed EDSS progression is not optimal for the assessment of disability progression?...

In this issue of *Neurology*, Koch and colleagues[1] report results of a study that explored the

ratios of disability worsening and improvement when detected by EDSS and by two other

commonly assessed tests of neurological performance, timed 25-foot walk (T25FW) and nine-

Commented [CO1]: This is reference 1. All the other references need to be renumbered.

Deleted: (but rarely used as primary outcomes)

hole peg test (9HPT). The study utilised data from two randomised clinical trials in SPMS: IMPACT (interferon β-1a vs. placebo) and ASCEND (natalizumab vs. placebo). It builds on the premise that most of disability accrued during the course of SPMS is irreversible and that improvement in disability mostly represents noise due to measurement error. In both datasets, the rates of improvement of T25FW and 9HPT were substantially lower than their rates of worsening. In contrast, the rates of disability improvement and worsening based on EDSS were very similar. The authors conclude that EDSS is more amenable to measurement error than the two other outcome measures and suggest that T25FW and 9HPT are more suitable measures to quantify disability outcomes in trials of SPMS than EDSS. Koch et all attribute this to their relatively higher inter- and intra-rater reliability and also the fact that T25FW and 9HPT are objective interval scales.

The most important ingredient of a successful trial is its primary outcome. Thus, the present study has considerable implications for the design of future clinical trials, since detecting a treatment effect on an outcome that is subject to large measurement error is difficult.

Measurement error is part of research reality. Despite this fact, its presence has been overlooked in most analytical designs. A *random measurement error* affects all measurements with equal probability and impact, irrespective of the value of the measured outcome and patients' allocation to an intervention. Such error does not influence the mean estimated treatment effect, but it does increase the variability of that estimate, inflating type-II error. On the other hand, a *differential measurement error* influences the measurements of an outcome in patients from different groups and with different values of their measured outcome differently. Therefore, differential measurement error represents a significant problem, as it introduces bias, affecting the estimates of treatment effect and leading to inflation of both type-II and type-I errors. Koch and colleagues highlight the importance of measurement error

associated with EDSS, which could potentially lead to either of the above scenarios. In both scenarios, quantification of measurement error and its incorporation into statistical design of clinical trials will increase the precision of their conclusions.[2]

A designer of a randomised trial should choose the primary outcome measure with the best signal-to-noise ratio. In relapsing-remitting MS, the signal for EDSS progression is large, as the effect of therapies on sudden, stepwise changes in disability is large. However, in MS phenotypes with gradual change in disability, the yield of confirmed EDSS worsening has been underwhelming, as evidenced by a number of clinical trials that showed borderline effects with a large variability. One therefore has to ask: Are we using the right instruments to measure the effect of therapies on disability in progressive MS?

Deleted: the

Deleted: on of

In addition to the correct choice of disability outcomes in trials, we need to understand their properties better. This study is an example of an analysis that explores the signal-to-noise ratio of the most widely used primary outcome - confirmed progression of EDSS. Another very important decision concerns the choice of thresholds for progression and improvement on the various scales. This study did not explore the importance of different thresholds for categorising progression of disability. 9HPT and T25FW traditionally use a 20% threshold for worsening. However, a relative threshold would disadvantage patients with good baseline performance (with short time of 9HPT and T25FW) and inflate measurement error in this group. Interestingly, a comparison between, for instance, screening and baseline 9HPT and T25FW would allow direct estimation of their measurement errors, which would then inform definitions of disability progression and improvement for different baseline values.

In this study, the authors assume that improvement in disability measures in SPMS is due to measurement error. Even though this assumption may be valid in most cases, especially if

improvement is only detected by one of the scales, true recovery of neurological function in progressive MS is possible. An alternative interpretation of the presented results could be that recovery of neurological function is more common in SPMS than what we had previously thought, and that EDSS is more sensitive to its detection than the other two measures. This study does not provide a definitive answer to this question, as the 'ground truth' remains unknown.

A recovery may be driven by different biological mechanisms.[3] First, *spontaneous remyelination*, present in relapsing as well as progressive MS, can restore the function and metabolic support of demyelinated axons.[4] Second, *acute localised inflammation* in the CNS, which is common even in progressive MS forms,[5, 6] represents a therapeutic target for immunotherapies[7] and its resolution can slow progression and even lead to clinical improvement. Third, functional reorganisation of the damaged nervous circuitry *(neuroplasticity)* has been linked to clinical recovery in all MS phenotypes.[3, 8]

The study by Koch and colleagues brings to our attention the error in measurement of disability outcomes through demonstrating an incongruence among three commonly used disability measures. At present, most clinical trials in progressive MS use confirmed change in EDSS as their primary or key secondary outcomes.[9] However, as the authors elegantly show, other, more reliable clinical outcomes are needed. As we are revisiting our biological hypotheses for treatment of progressive MS, perhaps the time has come that we should also revisit the instruments that we use to examine their efficacy.

References
1.   Koch MW, Mostert J et al., Reliability of outcome measures in clinical trials in secondary progressive multiple sclerosis, Neurology 2020 XXXX.

2.   Nab L, Groenwold RHH, Welsing PMJ, van Smeden M. Measurement error in continuous endpoints in randomised trials: Problems and solutions. Stat Med 2019;38:5182-5196.

3.   Tomassini V, Matthews PM, Thompson AJ, et al. Neuroplasticity and functional recovery in multiple sclerosis. Nat Rev Neurol 2012;8:635-646.

4.   Lubetzki C, Zalc B, Williams A, Stadelmann C, Stankoff B. Remyelination in multiple sclerosis: from basic science to clinical translation. Lancet Neurol 2020;19:678-688.

5.   Kapoor R, Ho PR, Campbell N, et al. Effect of natalizumab on disease progression in secondary progressive multiple sclerosis (ASCEND): a phase 3, randomised, double-blind, placebo-controlled trial with an open-label extension. Lancet Neurol 2018;17:405-415.

6.   Montalban X, Hauser SL, Kappos L, et al. Ocrelizumab versus Placebo in Primary Progressive Multiple Sclerosis. N Engl J Med 2017;376:209-220.

7.   Lizak N, Malpas CB, Sharmin S, et al. Association of continued immunotherapy with slower disability worsening in active secondary-progressive multiple sclerosis. JAMA Neurol in press.

8.   Ontaneda D, Thompson AJ, Fox RJ, Cohen JA. Progressive multiple sclerosis: prospects for disease therapy, repair, and restoration of function. Lancet 2017;389:1357-1366.

9.   Tur C, Moccia M, Barkhof F, et al. Assessing treatment outcomes in multiple sclerosis trials and in the clinical setting. Nat Rev Neurol 2018;14:75-93.