Chapter 2.2

Progression on the Multiple Sclerosis Functional Composite in multiple sclerosis: what is the optimal cut-off for the three components?

LVAE Bosma, JJ Kragt, L Brieva, Z Khaleeli, X Montalban,
CH Polman, AJ Thompson, M Tintoré, BMJ Uitdehaag

ABSTRACT

Background: For the Timed 25-Foot Walk (T25FW) and 9-Hole Peg Test (9HPT), components of the Multiple Sclerosis Functional Composite (MSFC), cut-off points of 20% change have previously been defined as meaningful endpoints of functional decline. Recently, however, a 15% change of MSFC components was introduced.

Objective: To determine optimal cut-offs for all MSFC components to indicate clinical disease progression in a primary progressive (PP) MS population.

Methods: T25FW, 9HPT and Paced Auditory Serial Addition Test (PASAT) were performed in 161 patients with PPMS with a two year interval. Absolute and relative differences in test scores were calculated. For each cut-off point of relative change, proportions of patients who progressed (deterioration beyond cut-off value) and improved (improvement beyond cut-off value) were calculated. Further, we calculated the ratio of 'improved' versus 'progressed' patients. Line graphs were created indicating: percentage progressed patients, percentage improved patients, ratio of improved versus progressed patients. The optimal cut-off was determined by searching the cut-off point with the lowest ratio of improved versus progressed patients, while at the same time capturing a substantial amount of progression.

Results: For both T25FW and 9HPT, the ratio between patients that improved and worsened clearly decreased between the cut-offs of 15% and 20%. For the PASAT, the ratio between patients improved and worsened was persistently poor.

Conclusions: A cut-off of 20% for both T25FW and 9HPT has a better signal-to-noise ratio than lower values (e.g. 15%) and is therefore preferable for assessment of disease progression. No satisfactory cut-off point for the PASAT could be determined.

INTRODUCTION

Although the effectiveness of therapies for the relapsing-remitting phase of multiple sclerosis (MS) is improving,[1-9] the need for therapies targeting disease progression remains. There is therefore an increasing requirement for clinical scales sensitive to gradual disease progression. Ordinal scales used to rate impairment, such as the Expanded Disability Status Scale (EDSS), have relatively poor reliability and responsiveness.[10] Consequently, quantitative functional measures like the Multiple Sclerosis Functional Composite (MSFC) have been developed.[11] The MSFC includes three quantitative, continuous tests that evaluate ambulation (the Timed 25-Foot Walk, T25FW), arm dexterity (the 9-Hole Peg Test, 9HPT) and cognition (the Paced Auditory Serial Addition Test, PASAT). Because these continuous scales are able to measure small differences in function, they are more sensitive to change.

However, because of difficulty with the clinical interpretation of differences observed on these scales, they are not always used as continuous scales, but rather in a dichotomized way by defining meaningful endpoints of functional decline. In previous studies, a cut-off point of 20% change has been defined for T25FW and 9HPT, as this was considered to indicate a true change in function,[12] is associated with a change in patient perceived disability and is therefore clinically meaningful.[13,14] For the third MSFC-component, the PASAT, no cut-off points have been defined.

In a recent study in a trial cohort of patients with relapsing-remitting MS (RRMS),[15] a 15% change of the MSFC components was introduced. This raises questions about the optimal cut-off (the optimal percentage change) to be applied when defining progression on the individual MSFC components, specifically: Which cut-off point is sensitive enough to measure progression without incorporating too much noise? To answer this question, we decided to examine the optimal cut-off point for all three MSFC components. A cohort of patients with primary progressive MS (PPMS) is ideally suited to such a study as there is no confounding influence of relapses and remissions and substantial improvement over time is unlikely to occur.

METHODS

Patients and test procedures

Data of patients with PPMS[16] from three European MS centers were retrospectively selected for analysis, as described previously.[17] The main selection criteria were: PPMS,[18] baseline EDSS ranging from 2.0 to 6.5 and data of at least three assessments of EDSS and MSFC: a

Chapter 2.2

baseline assessment followed by a first follow-up visit (FU1) after one year (≥ 260 days) and a second follow-up visit (FU2) after two years. Neither age nor gender were selection criteria. All assessments were obtained as part of routine outpatient care during regular visits. Both EDSS[19] and MSFC[11] were performed in the same visit under standardized conditions. The three tests of the MSFC were practised at least once before baseline assessments were completed. If patients were unable to perform a test due to MS-related symptoms, the maximum allowed time for this test was assigned.

Analyses and statistics

We calculated both absolute and relative (compared to baseline) differences in the scores of the T25FW, 9HPT (mean of both hands) and PASAT over two years, because our previous study showed that in this PPMS population over two years, rates of clinically meaningful worsening were higher whereas improvement rates were lower compared with 1-year data.[17]

In order to determine the optimal cut-off (percentage change) on the MSFC-components, we studied the effect of altering the cut-off on the proportions of patients progressing (deterioration beyond cut-off value) and improving (improvement beyond cut-off value). For each cut-off point of relative change, starting from 0% to 50%, the proportions of patients who progressed and improved beyond that cut-off point were calculated for all three outcome measures. Line graphs for each outcome measure were created (Figures 2.2.1-2.2.3) with the cut-off value on the x-axis and including lines expressing the percentage progressed patients (blue line; diamonds) and the percentage improved patients (pink line; squares).

Because real improvement over two years is unlikely to occur in PPMS, improvements on a scale were considered as 'false' (noise) in this study and should thus be minimal. To further investigate this noise, and to see how the relationship between noise and signal changes when altering the cut-off, we also calculated the ratio of improved versus progressed patients for each cut-off point. This was expressed as a percentage:

$$[\% \text{ of patients improved} / \% \text{ of patients progressed}] * 100\%$$

A third line was included in the line graphs for each outcome measure, expressing this ratio of improved versus progressed patients (brown line; dots). The line graphs show that in general, when the cut-off increases, unfavourably the signal decreases (you lose sensitivity to detect worsening of function) while the advantage is that the noise decreases as well. This means that to determine the optimal cut-off is always a trade-off between sensitivity to progression and excessive noise, by which the line representing the ratio is meant as an aid. The optimal cut-

off point was determined by exploring the lines in search of the cut-off point with the lowest ratio of improved versus progressed patients, while at the same time capturing a substantial amount of progression.

All statistical analyses were performed using the Statistical Package for Social Sciences (SPSS) version 14.0 (SPSS, Inc., Chicago, IL, USA) and Microsoft Office Excel version 2003.

RESULTS

Timed 25-Foot Walk

The results for the T25FW show that, when looking at the cut-off of 20% for example (x-axis), 45% of all patients showed a worsening (blue line) and 11% showed an improvement (pink line) of more than 20% on the T25FW (**Figure 2.2.1**). The ratio between patients improved and patients worsened is (11/45*100=) 24% for the cut-off of 20%. When looking at the cut-off of 15%, 53% of patients showed a worsening and 17% showed an improvement of more than 15%. The ratio between patients improved and patients worsened is (17/53*100=) 33% for the cut-off of 15%.

Obviously, both the percentage of patients progressing (upper line) and the percentage of patients improving (lowest line) steadily decrease with increase of the cut-off value (i.e. with
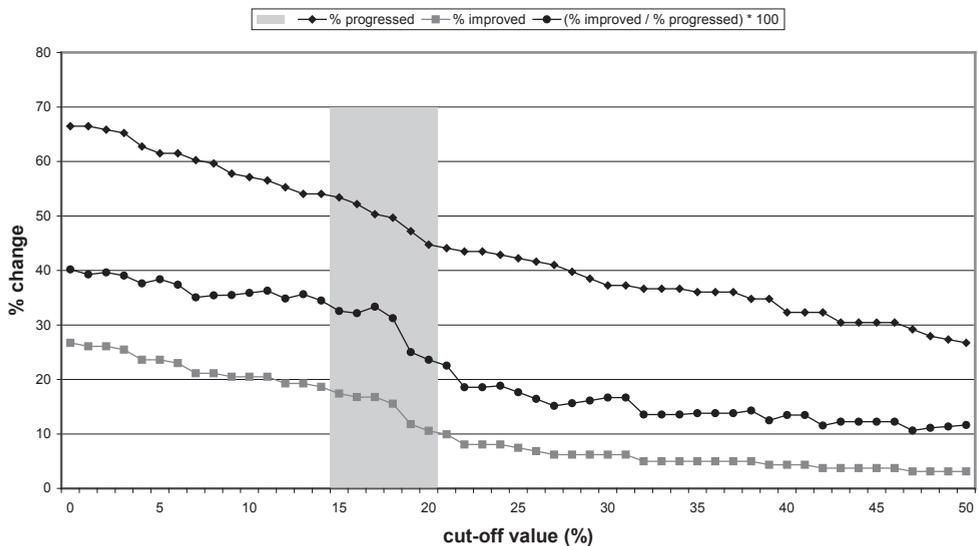

Figure 2.2.1   T25FW change in 2 yrs.

increasing value on the x-axis). When looking at the middle line representing the ratio between improvement and progression the steepest decline is apparently between the cut-off values 17% and 22%.

9-Hole Peg Test

Results for the 9HPT (mean of both hands) are shown in **Figure 2.2.2**. For example, looking at the cut-off of 20%, one can see that 16% of all patients showed a worsening, while 3% of all patients showed improvement of more than 20%. The ratio between patients improved and patients worsened is (3/16*100=) 20% for the cut-off of 20%.

The ratio between patients improved and progressed on the 9HPT decreases most prominently between the cut-off values 10% and 20%. Comparing Figures 2.2.1 and 2.2.2 it is obvious that the percentage of patients worsening on the 9HPT is much lower than that worsening on the T25FW and that for the 9HPT the difference between patients worsening and improving is much smaller than for the T25FW.

Paced Auditory Serial Addition Test

Results for the PASAT are shown in **Figure 2.2.3**. For example, again looking at the cut-off of 20%, one can see that 10% of all patients showed a worsening, while 24% showed improvement.
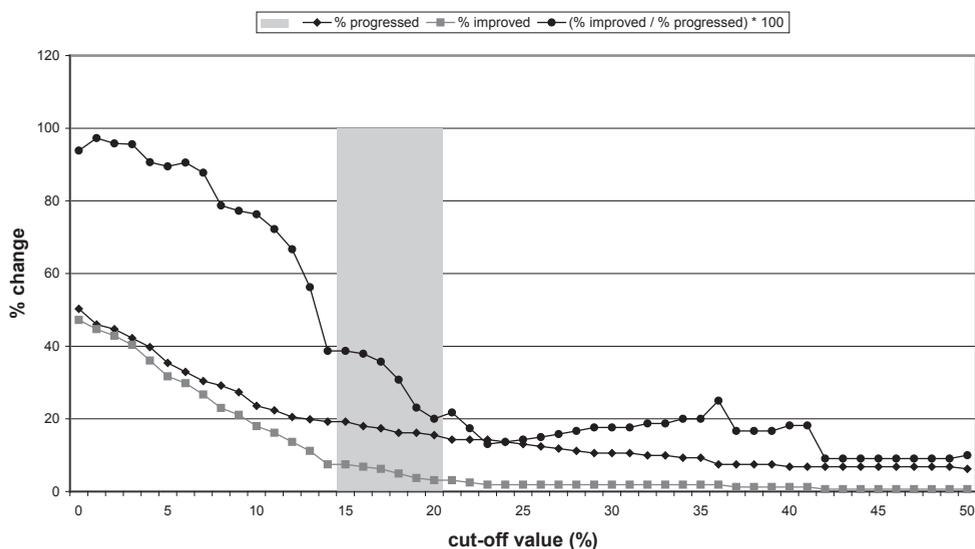

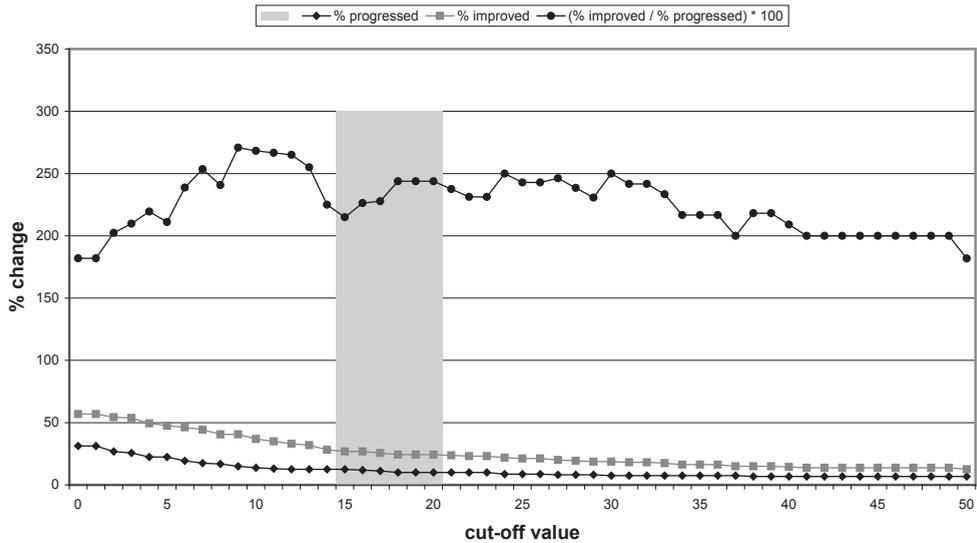**Figure 2.2.2   9HPT (mean of both hands) change in 2 yrs.**

Figure 2.2.3   PASAT change in 2 yrs.

The ratio between patients improved and patients worsened is (24/10*100=) 244% for the cut-off of 20%.

Strikingly, there is more improvement than progression on the PASAT irrespective of the choice of the cut-off value, resulting in the ratio of improved versus progressed patients being constantly very high.

CONCLUSIONS

In this study in a clinical population of PPMS-patients we searched for the optimal cut-off point for the quantitative components of the MSFC to measure disease progression using a different approach than previously applied. For the T25FW and 9HPT, the clinical rationale for using a 20% change has been demonstrated.[12-14] More recently the 15% cut-off was presented in a study based on a trial cohort of RRMS patients[15] and therefore, when comparing different cut-offs, we used these values as our main reference. However, we realize that there could well be differences between different forms of the disease and between a trial cohort and a clinical population.

Concerning the T25FW, there is a constant and substantial difference score between progressed and improved patients. The curve indicating the ratio between improvement and progression, in other words indicating the amount of noise being measured on the scale, shows quite a steep

decline between the cut-off points of 15% and 20%. From this curve we can conclude that, regarding the reliability of the scale, a cut-off point of 20% yields a profit, because the 'noise-to-signal' ratio is much more advantageous than at lower cut-off values. From that perspective a cut-off point of 25% would be even more advantageous, but between 20% and 25% the gain is less, and the obvious trade-off is a loss of sensitivity to detect worsening of function. Thus our approach confirms the previously suggested cut-off of 20% for the T25FW.

Regarding the 9HPT, the progression is less than on the T25FW. There is relatively more improvement, indicating a greater amount of noise on this scale. This could be due to the practice effect on the MSFC, which is known to be present on the 9HPT and PASAT, in contrast to the T25FW.[20,21] The curve indicating the ratio between improvement and progression on the 9HPT shows a steep decline until the cut-off of about 20%, indicating an advantageous 'noise-to-signal' ratio for a cut-off of 20% compared with lower values. The gain between the cut-offs of 15% and 20% is clearly demonstrated by the steepness of the curve. Beyond the cut-off of 20% there is no clear gain. Therefore, when looking at the mean change of both hands on the 9HPT, our approach leads us to a cut-off of 20% as well.

Concerning the PASAT, based on the data of this study, we are not able to indicate a reasonable cut-off point. Based on these data, we can only conclude that the PASAT is not very suitable to use as a scale to identify disease progression in a PPMS population. More research is warranted with this test or other cognitive tests in (primary) progressive MS patients.

We would like to comment on the improvements found on the quantitative functional measures in this study. It is known that the PASAT has the most evident practice effects of the three MSFC components, manifesting as improved performance on the first test sessions that are accomplished.[20,21] Because some previous studies involving the MSFC components have shown that the practice effects were present in the first three test sessions and stabilized by the fourth session,[20,22] it could be possible that – part of – the improvement found in this study on the PASAT and, to a lesser extent the 9HPT, is due to insufficient pre-baseline testing. Concerns have been raised about using the PASAT as the cognitive component of the MSFC, and research has been done to replace the PASAT by other tests measuring cognition (i.e. the Symbol Digit Modalities test).[23]

Finally, we would like to address some other possible limitations of this study. First, the reported MSFC changes were not confirmed by repeated measurements after three or six months. Second, our study population was a clinical population and not a trial cohort. Third, the data were retrospectively selected for analysis. Although most patients were seen during regular, scheduled visits to their neurologists, it might be possible that some came to the hospital because

of worsening, so there might be some selection bias favouring more progressive patients. Finally, a population of PPMS patients might not be the most appropriate to examine in order to determine the optimal cut-off for the PASAT, as cognitive decline in PPMS is less distinct than for example motor dysfunction. Previous studies on cognitive impairment and decline in PPMS showed that over 2 years no clear worsening could be seen on the PASAT and that cognitive decline might be more evident in secondary progressive (SP) MS than in PPMS.[24-26]

Taken together, using a different approach we confirm the previously proposed cut-off of 20% for both the T25FW and the 9HPT. Based on the data of this study, we feel that for both these MSFC-components a cut-off of 20% has a more advantageous 'noise-to-signal' ratio than a cut-off of 15%. Concerning the PASAT, we were not able to indicate a reasonable cut-off point and more research is needed.

<div style="text-align: right">Chapter 2.2</div>

REFERENCES

1. Jacobs LD, Beck RW, Simon JH, et al. Intramuscular interferon beta-1a therapy initiated during a first demyelinating event in multiple sclerosis. CHAMPS Study Group. N Engl J Med 2000; 343: 898-904.

2. Comi G, Filippi M, Barkhof F, et al. Effect of early interferon treatment on conversion to definite multiple sclerosis: a randomised study. Lancet 2001; 357: 1576-1582.

3. Kappos L, Polman CH, Freedman MS, et al. Treatment with interferon beta-1b delays conversion to clinically definite and McDonald MS in patients with clinically isolated syndromes. Neurology 2006; 67: 1242-1249.

4. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. I. Clinical results of a multicenter, randomized, double-blind, placebo-controlled trial. The IFNB Multiple Sclerosis Study Group. Neurology 1993; 43: 655-661.

5. Randomised double-blind placebo-controlled study of interferon beta-1a in relapsing/remitting multiple sclerosis. PRISMS (Prevention of Relapses and Disability by Interferon beta-1a Subcutaneously in Multiple Sclerosis) Study Group. Lancet 1998; 352: 1498-1504.

6. Johnson KP, Brooks BR, Cohen JA, et al. Copolymer 1 reduces relapse rate and improves disability in relapsing-remitting multiple sclerosis: results of a phase III multicenter, double-blind placebo-controlled trial. The Copolymer 1 Multiple Sclerosis Study Group. Neurology 1995; 45: 1268-1276.

7. Comi G, Filippi M, Wolinsky JS. European/Canadian multicenter, double-blind, randomized, placebo-controlled study of the effects of glatiramer acetate on magnetic resonance imaging--measured disease activity and burden in patients with relapsing multiple sclerosis. European/Canadian Glatiramer Acetate Study Group. Ann Neurol 2001; 49: 290-297.

8. Polman CH, O'Connor PW, Havrdova E, et al. A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. N Engl J Med 2006; 354: 899-910.

9. Rudick RA, Stuart WH, Calabresi PA, et al. Natalizumab plus interferon beta-1a for relapsing multiple sclerosis. N Engl J Med 2006; 354: 911-923.

10. Hobart J, Freeman J, Thompson A. Kurtzke scales revisited: the application of psychometric methods to clinical intuition. Brain 2000; 123: 1027-1040.

11. Cutter GR, Baier ML, Rudick RA, et al. Development of a multiple sclerosis functional composite as a clinical trial outcome measure. Brain 1999; 122: 871-882.

12. Schwid SR, Goodman AD, McDermott MP, Bever CF, Cook SD. Quantitative functional measures in MS: what is a reliable change? Neurology 2002; 58: 1294-1296.

13. Hoogervorst EL, Kalkers NF, Cutter GR, Uitdehaag BM, Polman CH. The patient's perception of a (reliable) change in the Multiple Sclerosis Functional Composite. Mult Scler 2004; 10: 55-60.

14. Kragt JJ, van der Linden FA, Nielsen JM, Uitdehaag BM, Polman CH. Clinical impact of 20% worsening on Timed 25-foot Walk and 9-hole Peg Test in multiple sclerosis. Mult Scler 2006; 12: 594-598.

15. Rudick R, Polman C, Cohen J, et al. Assessing disability progression with the Multiple Sclerosis Functional Composite. Mult Scler 2009; 15: 984-997.

16. Poser CM, Paty DW, Scheinberg L, et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. Ann Neurol 1983; 13: 227-231.

17. Bosma L, Kragt J, Brieva L, et al. The search for responsive clinical endpoints in primary progressive multiple sclerosis. Mult Scler 2009; 15: 715-720.

18. Lublin FD, Reingold SC. Defining the clinical course of multiple sclerosis: results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis. Neurology 1996; 46: 907-911.

19. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983; 33: 1444-1452.

20. Cohen JA, Cutter GR, Fischer JS, et al. Use of the multiple sclerosis functional composite as an outcome measure in a phase 3 clinical trial. Arch Neurol 2001; 58: 961-967.

21. Solari A, Radice D, Manneschi L, Motti L, Montanari, E. The multiple sclerosis functional composite: different practice effects in the three test components. J Neurol Sci 2005; 228: 71-74.

22. Cohen JA, Fischer JS, Bolibrush DM, et al. Intrarater and interrater reliability of the MS functional composite outcome measure. Neurology 2000; 54: 802-806.

23. Drake AS, Weinstock-Guttman B, Morrow SA, et al. Psychometrics and normative data for the Multiple Sclerosis Functional Composite: replacing the PASAT with the Symbol Digit Modalities Test. Mult Scler 2010; 16: 228-237.

24. Huijbregts SC, Kalkers NF, de Sonneville LM, de Groot V, Polman CH. Cognitive impairment and decline in different MS subtypes. J Neurol Sci 2006; 245: 187-194.

25. Bergendal G, Fredrikson S, Almkvist O. Selective decline in information processing in subgroups of multiple sclerosis: an 8-year longitudinal study. Eur Neurol 2007; 57: 193-202.

26. Camp SJ, Stevenson VL, Thompson AJ, et al. A longitudinal study of cognition in primary progressive multiple sclerosis. Brain 2005; 128: 2891-2898.