Chapter 4

Measurement of treatment effect after IV Methylprednisolone treatment

Chapter 4.1

Differential treatment effect on measures of neurologic exam, functional impairment and patient self-report in MS

Hoogervorst ELJ, MD, Kalkers NF, MD, van Winsen LML, MD, Uitdehaag BMJ, MD, Polman CH, MD

Multiple Sclerosis 2002 Oct; 7 (5):335-9

Abstract

To determine the relative sensitivity of the expanded disability status scale (EDSS), the newly developed MS functional composite (MSFC) and the Guy's Neurological Disability Scale (GNDS) to changes in the neurological condition of MS patients induced by treatment with IVMP. Sixty MS patients were treated withintravenous methylprednisolone (IVMP). On the first day of treatment patients were trained for the three domains of the MSFC; on the second day baseline data were obtained for all measurements. Follow-up data were obtained 6-8 weeks after IVMP treatment.

Significant changes were found for both EDSS and GNDS. Remarkably, the improvements on the GNDS were mainly due to changes in the subcategories cognition, speech, fatigue and 'others'. No significant change was found for the MSFC. Forty-seven patients reported a subjective improvement in their condition. Twenty one patients showed a significant improvement in the EDSS, 28 patients showed a significant improvement in the GNDS and a very small number of significant changes were found on the MSFC (actual number depending on the definition of the reference population).

The observations in this study show that the relative sensitivity to change in acute or sub acute deterioration in MS patients is low for the MSFC and high for the GNDS. It is obvious from our study that a treatment can have a differential effect on measurements of functional impairment, rating of neurologic exam and patient self-report.

Introduction

In clinical trials of MS, the EDSS is traditionally being used as the primary outcome measure for neurological impairment and disability (23). The EDSS is based on neurologist rating of neurological exam depending on eight functional systems (visual, brainstem, pyramidal, sensory, cerebellar, bowel/bladder, cerebral and 'others') and the patient's ability to walk. Clear limitations have been identified when using the EDSS. The major problems are related to standardization, sensitivity, reliability, interrater variability and the scale being heavily biased to locomotor function (141-144).

Recently, the National MS Society (NMSS) Clinical Outcomes Assessment Task Force developed the MS functional composite (MSFC) to address the poor reliability and insensitivity to change of the available MS clinical rating scales (24, 145, 146). The MSFC assesses functional impairment and covers three domains: arm/hand function (9-hole Peg test) (9-HPT)), cognition (3" version of the paced auditory serial addition test (PASAT)) and leg function/ ambulation (25-foot timed walk test (T25FW)). The combination of these three domains comprises a composite score expressed as standard deviations from a reference population. The composite score thus generated showed high reliability, clear relationship with disease duration and excellent intrarater and interrater reliability (24, 147, 148).

The Guy's Neurological Disability Scale (GNDS) has recently been introduced as a new measure, based on patient self report (149). It is patient orientated, multidimensional, and not biased towards any particular disability (149, 150). The GNDS is a questionnaire, directed to assess the disability in the previous month, driven by patient interview and it can be applied by any health care personnel. It is therefore extremely practical and capable of incorporating patients' views of their disability in a structured manner (150, 151). The GNDS is divided into twelve separate subcategories (cognition, mood, vision, speech, swallowing, upper limb function, lower limb function, bladder function, bowel function, sexual function, fatigue and 'others'). The GNDS was devised to be a simple, user-friendly clinical disability scale capable of embracing the whole range of disabilities which can be encountered in the course of MS.

In a previous study we compared the EDSS, MSFC and GNDS as measures of neurologist rating of neurological exam abnormalities, observations of functional impairment and patient self report (150). In this study good correlations are shown between GNDS and both EDSS and MSFC, mainly due to the importance of spinal cord related neurological functions. However, a very marked discrepancy was found for the assessment of cognition between objective measurements (PASAT3") and subjective complaints (subcategory cognition of the GNDS) [12]. This study also showed that the GNDS measures new dimensions of disability and incorporates patients' perspectives; the subcategories fatigue and sexual function are not represented in the EDSS while fatigue seems to be a determinant for EDSS scores. Likewise the GNDS contains subjective subcategories which are not represented in the MSFC, like swallowing, bowel-, bladder and sexual function and fatigue.

Because the GNDS can offer a valuable way to document disease impact in MS (150), and

because information on the sensitivity to change for the MSFC and GNDS is lacking to date, we decided to investigate the relative sensitivity of the MSFC and GNDS compared to the EDSS, by studying the behaviour of these scales when documenting the change in impairment/disability induced in MS patients by treatment with high-dose IVMP.

Methods

Patients

Sixty patients, aged 19 to 73, with clinically definite MS (18), who were treated with IVMP (1000 mg once daily for 3 consecutive days or 500 mg once daily for 5 consecutive days) were scored for EDSS, MSFC and GNDS at specified predefined points in time before, during and after this treatment intervention. The indication for IVMP treatment was an acute relapse in 27 RRMS (102) patients or sub acute deterioration in the absence of relapses in 33 patients (8 patients being primary progressive (PP) and 25 secondary progressive (SP) (102). The time from onset of an acute relapse or sub acute deterioration to start IVMP treatment varied from 1 to 4 weeks and there was no spontaneous improvement before treatment with IVMP was started.

Test procedures

EDSS- and MSFC-data were collected under carefully standardized conditions by well-trained examiners (147, 148, 150). Assessment of the MSFC took place on 4 sessions, these being the first day of treatment (twice), the second day of treatment and six to eight weeks after treatment. The assessment of impairment and disability using EDSS and GNDS took place on the same day as the third and fourth session of the MSFC assessment. The assessments took place in the same order; first the MSFC components, followed by the GNDS and EDSS. Data from both sessions on the first day (9-HPT, PASAT 3" and T25FW were used for training purposes only (148), whereas data from the third and fourth sessions were used as parameters to evaluate sensitivity to change. The 9-HPT was performed twice for each hand, the mean of the best time for each hand (in seconds) being analyzed. The PASAT 3" was performed once (score in total numbers correct). The T25FW was performed twice, and the best time (in seconds) was used for analysis. Inability to perform a test, due to MS related symptoms, was scored with the maximum time allowed for the 9-HPT (300 sec) and T25FW (180 sec), and with the worst score for the PASAT3" (0). Each subcategory of the GNDS was scored separately ranging from 0 (normal) to 5 (maximum help required) (149). All patients were also asked whether they had experienced a subjective change in their condition after IVMP treatment.

Analysis

For creating the MSFC-score, Z-scores were created for the 9-HPT, PASAT 3" and T25FW. Z-scores were obtained by subtracting the mean of the total population of the third assessment session, from the test result of each individual patient and dividing the difference by the

standard deviation of the total population (internal reference population). In addition Z-scores were calculated using the mean and standard deviation of an external reference population consisting of a wide range of MS patients (147) (external reference population); Z-scores were also calculated for the distinct phenotypes using means and standard deviations of these subgroups of the external reference population. In order to achieve that for all three tests deterioration pointed in the same direction(152), the 9-HPT was transformed to its inverse and the Z-score of the T25FW was multiplied by -1. The composite score then was calculated by adding the three Z-scores:

$$MSFC = (Z_{[1/(9\text{-HPT}),\ average]} - Z_{[T25FW]} + Z_{[PASAT\ 3'']})/3.$$

The MSFC-score becomes higher when patients have better scores and vice versa, the MSFC-score becomes lower when patients have worse scores on the MSFC components compared to the reference population. For creating the GNDS-score the sum score of the twelve separate subcategories was used.

Results were analyzed in several ways; comparisons between EDSS, MSFC, GNDS and the components of the MSFC and GNDS before and after treatment were made, effect sizes were calculated. Since this analysis incorrectly assumes that the EDSS is a continuous variable, we also compared the total number of patients responding to treatment for each scoring system. A significant change of the EDSS was defined as a change of 1.0 point or more at EDSS levels < 5.5 or 0.5 point or more at EDSS levels ≥ 5.5 (153). A significant change of the MSFC was defined as a change of at least 0.5 standard deviation, a criterion used in previous studies and very similar to that applied in a currently ongoing clinical trial (154, 155). Based on Sharrack and Hughes (151) we defined a significant change of the GNDS as a change of three or more points in the GNDS sum score. In their study a change of 3 or more points had an interrater agreement of 100% and an intrarater agreement of 92%.

Statistics

Comparisons between EDSS, GNDS and GNDS subcategories before and after treatment were made using Wilcoxon signed-rank test. For the MSFC and its components significant changes of the mean from baseline were determined by Student's *t*-test.

Effect sizes were calculated by dividing the mean change score (baseline minus post-treatment mean score) by the baseline standard deviation (156), a larger effect size indicating a greater responsiveness.

Results

Patient characteristics and median scores on the EDSS, GNDS, MSFC and its components are summarized in Table 1.

Table 1. Patient characteristics (expressed as median (interquartile range)).

Characteristics	Total	
Total	60	
M	27	
F	33	
Age in years	42.5 (36.5-48.2)	
	Baseline	After treatment
EDSS*	6.0 (4.0-6.5)	5.0 (3.0-6.5)
GNDS†	17 (2-23)	14 (9-22)
MSFC‡ internal reference population		
external reference population	0.23 (-0.39-0.58)	
0.12 (-0.64-0.55)	0.30 (-0.54-0.53)	
0.24 (-0.67-0.52)		
9-HPT§, sec	25.0 (20.5-34.0)	25.1 (20.3-44.5)
PASAT‖, numbers correct	50 (37-57)	50 (38-56)
T25FW¶, sec	7.4 (4.9-12.6)	6.2 (4.7-11.4)

*EDSS = Expanded Disability Status Scale; †GNDS = Guy's Neurological Disability Scale; ‡MSFC = Multiple Sclerosis Functional Composite; §9-HPT = 9-hole peg test; ‖PASAT = paced auditory serial addition test; ¶T25FW = timed walk test

In the total population median EDSS score at baseline was 6.0 (Interquartile Range (IQR) 4.0-6.5) and after treatment 5.0 (IQR 3.0-6.5) ($p<0.001$), effect size 0.28. For the different subtypes median EDSS score at baseline in RRMS patients was 4.0 (IQR 3.5-6.0) and at follow-up 3.5 (2.0-5.0) for patients with sub acute deterioration median EDSS at baseline was 6.5 (IQR 4.8-7.5) and at follow-up 6.0 (IQR 4.5-7.5). Another significant improvement in the total population was found for the GNDS, median GNDS score at baseline was 17.0 (IQR 12.0-23.0) and after treatment 14.0 (IQR 9.0-22.0) with an effect size of 0.25. No significant change was found for the MSFC before and after treatment intervention, median MSFC-score (external reference population used) at baseline was 0.12 (IQR -0.64-0.55) and after treatment 0.24 (IQR -0.67-0.52), effect size 0.01. Results for the 9-HPT and PASAT3" before and after IVMP treatment were almost the same and T25FW improved significantly (95% Confidence Interval: 0.12-0.95).

Table 2 shows the number of patients having significant changes on EDSS, MSFC and GNDS. For the EDSS there was a significant improvement in 21 patients (10 progressive and 11 RR); three patients showed a significant worsening on the EDSS (2 progressive and 1 RR). Significant improvement in GNDS-score was found in 28 patients (15 progressive and 13 RR) and worsening in 7 patients (3 progressive and 4 RR). Remarkably the significant improvements on the GNDS were induced by changes in the subcategories cognition (17 patients, $p<0.02$), speech (16 patients, $p<0.01$), fatigue (22 patients, $p<0.01$) and 'others' (19 patients, $p<0.02$) mainly regarding pain or spasm. The MSFC showed no significant improvement and 1 significant worsening (RR patient) when either the internal or external reference population was used for analysis. When calculating MSFC-scores for the different phenotypes using

Table 2. Number of patients showing significant change in EDSS, GNDS and MSFC

			Changes in the EDSS*		
			Improvement	No change	Worsening
		Total Number	21	36	3
GNDS†	Improvement	28	9	19	0
	No change	25	9	14	2
	Worsening	7	3	3	1
MSFC‡	Improvement	0	0	0	0
	No change	59	21	36	2
	Worsening	1	0	0	1

*EDSS = Expanded Disability Status Scale; †GNDS = Guy's Neurological Disability Scale; ‡MSFC = Multiple Sclerosis Functional Composite.

Table 3. Number of patients with subjective improvement in their condition (n=47) and significant change in EDSS, MSFC and GNDS.

	EDSS*	MSFC†	GNDS‡
Improvement	15	0	27
No Change	32	47	17
Worsening	0	0	3

*EDSS = Expanded Disability Status Scale; †MSFC = Multiple Sclerosis Functional Composite; ‡GNDS = Guy's Neurological Disability Scale.

means and standard deviations of those phenotypes in the external reference population, there were 3 significant improvements and 2 significant worsenings (all RR patients). Ten (progressive) patients were unable to perform the 9-HPT, of which 2 patients were unable to perform the test for both hands and 3 patients improved after IVMP, 11 (progressive) patients were unable to perform the T25FW before as well as after treatment and no patients were unable to perform the PASAT.

Forty seven patients reported subjective improvement in their condition after IVMP treatment (26 progressive and 21 RR patients) (table 3). Twenty seven of the 28 patients with significant improvement on the GNDS reported subjective improvements in their condition. The patient who improved significant on the GNDS and didn't report a subjective improvement in his condition also showed a significant improvement on the EDSS. Of the 7 patients who showed significant worsening on the GNDS, 3 patients reported subjective improvement in their condition and 4 patients reported a subjective worsening. Of the 21 patients who showed significant improvement on the EDSS, 15 patients (71%) reported subjective improvement in their condition. Eight patients reported subjective improvement and showed significant improvements on both the EDSS and GNDS.

Discussion

In this study we evaluated the relative sensitivity to change of the MSFC and GNDS compared to the EDSS in MS patients treated with IVMP. The observations of this study show that the relative sensitivity to change is low for the MSFC and high for the GNDS.

Our data do not confirm previous suggestions that the MSFC is more sensitive to change than the EDSS. Of course this can be partially explained by subjective changes not being reflected in the MSFC, but it is very striking that there was no improvement at all in the mean MSFC score even though there were marked neurological improvements in quite a number of patients. Ambulation improved in quite a number of patients, as is reflected in both an improvement of some EDSS scores of 4.0 and higher and a significant improvement of the T25FW. However this significant improvement in the T25FW was not reflected in an improvement in the overall MSFC due to the fact that there were essentially no changes in the 9-HPT and PASAT3". Part of the lack of responsiveness of the MSFC can also be explained by a pronounced ceiling effect of the tests applied; of our patients 17% was unable to perform the 9-HPT and 18% was unable to perform the T25FW.

Our data also re-emphasize the importance of the selection of an appropriate reference population for the MSFC. Whereas no significant improvement was seen in the total population (using the internal or a large external population as a reference) some significant improvements were seen in specific -more homogeneous- subpopulations (where the same specific subpopulations of the external population were used as a reference).

At present we are unable to identify major methodological issues which might have caused the lack of responsiveness of the MSFC, since proper training took place before baseline data were obtained and because all assessments took place under carefully standardized conditions utilizing only three well-trained examiners.

We have identified two possible explanations for our observation that the GNDS seems to be a sensitive outcome measure in this study. The first point being that IVMP treatment may induce a more pronounced subjective than objective change in the patient's condition, and the second being that so far there is no generally accepted definition of a significant change for the GNDS.

The suggestion that corticosteroids might especially have an effect on subjective features of well being is certainly supported by the data collected in this study. After IVMP treatment 78% (47/60) of the patients reported a subjective improvement, 47% (28/60) showed a significant improvement on the GNDS whereas 35% (21/60) of the patients showed a significant improvement on the EDSS and none of the patients showed an improvement on the MSFC. Changes on the individual subcategories of the GNDS were especially reported for rather subjective measures as cognition, mood, fatigue and 'others' (reflecting pain or spasm). Whereas in our previous cross-sectional study (150) we already noted the poor correlation between the subjective subcategory cognition of the GNDS and the measurement of the

PASAT, we here show that reported changes in cognition are not reflected in changed scores on the PASAT.

It is possible that the chosen three point change on the GNDS sum score is too small a difference to be really significant. Our criteria for significant change of the GNDS were based on the first validation paper of the scale (149); the fact that IVMP not only induced a high number of significant improvements (28 patients) but also a high number of significant worsenings (7 patients), certainly may suggest that these criteria for significant change are too moderate.

Even though this was not the primary purpose of this study, we found that quite some significant improvements where found in patients in the progressive phase of the disease, where steroid treatment has not convincingly been proven to be effective.

In conclusion, this comparative study in patients being treated with IVMP shows that the relative sensitivity to change is low for the MSFC and high for the GNDS.

Even though a number of variables that might affect the sensitivity to change for the respective scales have been identified, we suggest that further studies will be performed to study their behaviour under different conditions. It is obvious from our study that a treatment can have a differential effect on measures of functional impairment, rating of neurological exam and patient self-report.