EDSS progression assessment heterogeneity in MS according to geographical areas

F. Bovis[1], A. Signori[1], L. Carmisciano[1], I. Maietta[1], JR. Steinerman[2], T. Li [2], AP. Tansy[2], and MP Sormani[1,3]

[1]Department of Health Sciences (DISSAL), Section of Biostatistics, University of Genova, Genova, Italy, [2]Teva Pharmaceuticals, Malvern, PA, USA, [3]Ospedale Policlinico San Martino IRCCS, Genova, Italy

Corresponding Author: Prof Maria Pia Sormani, Department of Health Sciences (DISSAL), Via Pastore 1, Genova, Italy, Telephone number: +39 010-3538473, fax number: +39 010-3538441, e-mail address: mariapia.sormani@unige.it

Number of characters in the title: 79

Number of characters in the running head: 40

Number of words in the Abstract, Introduction, Discussion and the body of the manuscript (not including abstract or references, figure legends, etc.): 100+1659

Number of figures, color figures, and tables: 2 Figures and 1 Table

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ana.25323

Abstract

Using placebo data from 3 randomized MS trials with uniform inclusion criteria, we investigated heterogeneity of EDSS progression by geographical areas. Our analysis revealed a significantly lower EDSS progression in Eastern European countries (10.8%) compared with Europe (13.1%) or USA/Canada (21.4%,$p<0.001$); EDSS improvement behaved the same way . This heterogeneity is not explained by differences of baseline variables. No differences were detected on more easily quantifiable measures, the T25FW or the MSFC. At a time when disease progression represents the target for future interventions in MS, establishment of more quantitative and objective outcomes remains a key priority of MS research.

Introduction

Confirmed expanded disability status scale (EDSS)[1-2] progression, defined as an increase in EDSS confirmed in the subsequent months (typically 3 or 6), has been the endpoint used to assess the effect of treatments on disability accumulation in all MS phase III clinical trials, both in relapsing-remitting (RR) and in progressive MS. In trials assessing highly effective drugs, the concept of confirmed EDSS improvement (defined as a decrease in EDSS confirmed in the subsequent months) was recently introduced[3]. Current clinical trials reported an unexpectedly low rate of EDSS progression events. In the CONCERTO study, for example, the EDSS progression rate in placebo patients was around 11% over 2 years[4]; in the pooled RADIANCE and SUNBEAM studies the 15-months rate of events in the IFNbeta-1a arms was 4.2%.[5-6]

A recent meta-analysis showed that EDSS disease progression in placebo arms of clinical trials exhibits similar decline over time as the annualized relapse rate[7]. It is therefore recognized that modern trials include patients with less active disease than past trials, and that more rigorous methods of outcome assessments may play a role in decreasing the detection of spurious events (e.g. adjudication committees for relapses). Modern trials, on the other hand, tends to include populations from areas of the world not included in the past (mainly Eastern Europe and India). With this background, taking advantage of data from a pooled dataset of 3 randomized controlled trials (RCTs) in RRMS with the same inclusion criteria, we investigated whether there are differences in the proportion of patients with progression or improvement events (defined on the EDSS) and on other more easely quantifiable clinical endpoints, such as the Timed 25-Foot Walking (T25WT)[8] and the Multiple Sclerosis Functional Composite (MSFC)[9-10], related to the region of the world where the trial was conducted.

Methods

Patients

This is a post-hoc analysis of 3 RCTs, the ALLEGRO, BRAVO and CONCERTO studies (ClinicalTrials.gov identifiers: NCT00509145, NCT00605215 and NCT01707992, respectively). The study design and inclusion/exclusion criteria were the same for the 3 trials and have been described elsewhere.[4,11-12] . The ethics committees and institutional review boards of all participating centers approved the study protocols. All participants provided written informed consent[11-12]. Briefly, eligibility criteria included age 18–55 years, diagnosis of RRMS (revised McDonald criteria[13]) and EDSS scores of 0–5.5. Patients must have had at least one relapse in the previous 12 months, two relapses in the previous 24 months, or one relapse in the previous 12–24 months plus one gadolinium-enhancing lesion in the previous 12 months.

For the purpose of this study we analyzed only patients included in the placebo arms of the pooled studies. Patients were grouped according to the geographical region of enrollment, as reported in Table 1.

Outcomes

We analyzed disability progression/improvement, defined as a 1.0 point increase/decrease in EDSS score if baseline score was between 0 and 5.0, or a 0.5 point increase/decrease if baseline score was 5.5, sustained for 3 months;[3-4,11-12] and the T25FW progression, defined as the increase of at least 20% from baseline, sustained for 3 months.[14] We also analyzed the MSFC[9-10] change over 2 years in the subgroup of trials where this endpoint was collected (ALLEGRO and BRAVO).[11-12]

Statistical methods

Comparisons between patients were run by the non-parametric Kruskal Wallis test and the chi-square test. KM survival curves and Cox models adjusting for baseline covariates and study group were used to display and compare progression and improvement risk among groups. MSFC change over 2 years was compared among groups by an ANOVA model adjusted for baseline characteristics and study group.

Results

1746 RRMS patients treated with placebo were included in the analysis and their baseline characteristics, according to the 3 geographical regions, are reported in Table 1: 291 (17%) were enrolled in European centers, 145 (8%) in USA/Canada and 1310 (75%) were enrolled in East Europe.

Patients enrolled in the placebo arms in Eastern European countries were younger; with a more recent disease onset and shorter disease duration; with higher EDSS score; and higher MRI lesion volumes than placebo-treated patients from Europe or USA/Canada (Table 1).

At univariate analysis, the actuarial proportion of patients with 3-month confirmed EDSS progression was significantly heterogeneous ($p<0.0001$) among the 3 geographical regions: a higher 2-year cumulative probability of EDSS progression was observed for patients enrolled in USA/Canada (21.4%), followed by those enrolled in Europe (13.1%), while lowest risk was observed for those enrolled in East Europe (10.8%, Figure 1A). The same was true for the actuarial proportion of patients with 3-month confirmed EDSS improvement ($p=0.0002$): a higher 2-year cumulative probability of EDSS improvement was observed for patients enrolled in USA/Canada (13.4%), followed by those enrolled in Europe (12%), while the lowest proportion was observed for those enrolled in East Europe (6.8%, Figure 1B).

Additional details on heterogeneity across the countries within Eastern Europe are reported in the Supplementary material.

Since there were differences among the 3 regions in baseline characteristics (Table 1), we ran a multivariate model to check whether the heterogeneity in the actuarial proportion of patients with EDSS progression and improvement could be explained by differences of the baseline variables.

However, when adjusting for all the baseline variables, the EDSS progression risk differences among the geographical regions were maintained (p<0.0001): with the East Europe group as a reference, the hazard ratios (HRs) for Europe and USA/Canada patients, adjusted for the baseline characteristics, were 1.7 (95% confidence interval (CI): 1.2;2.5, p=0.006) and 2.3 (95%CI: 1.5;3.7, p=0.0003), respectively. The heterogeneity was maintained also for the proportion of patients with EDSS improvement (p=0.004): with the East Europe group as a reference, the hazard ratios (HRs) for Europe and USA/Canada patients, adjusted for the baseline characteristics, were 1.9 (95%CI: 1.2;3.1, p=0.01) and 1.6 (95%CI: 0.9;2.8, p=0.12), respectively.

On the other hand, there was no detectable differences in T25FW progression among patients enrolled in the 3 geographical regions (East Europe=16.9%, Europe=13.4%, USA/Canada=13.9%): adjusting for the baseline characteristics and taking East Europe as reference, the HR for Europe was 0.9 (95%CI: 0.6;1.5, p=0.73) and it was 0.8 (95%CI: 0.4;1.5, p=0.51) for USA/Canada (Figure 2A). Also the MSFC change, assessed in 2 trials (n=826), was not significantly different among the geographical regions (Figure 2B): the mean MSFC change was +0.002 (95%CI= -0.031;0.034) in East Europe, it was -0.030 (95%CI= -0.082;0.023) in Europe, and it was +0.001 (95%CI= -0.063;0.065) in USA/Canada (p=0.66); the comparison was not affected by baseline adjustment. All the details of the regression models are reported in the Supplementary material.

Discussion

This study run on the placebo arms of 3 large RCTs revealed a high heterogeneity in the proportion of patients with both an EDSS progression and an EDSS improvement according to the geographic area of enrollment (Europe vs Eastern Europe vs USA/Canada); this heterogeneity cannot be explained by differences of baseline characteristics. Notably, the differences in EDSS progression/improvement among geographical regions are much higher than those usually observed due to treatment effects on this outcome. Moreover, disability progression defined on more "objective" and quantitative measures, such as the T25FW and the MSFC, did not differ across

regions. USA/Canada and Europe had a higher EDSS progression risk than East Europe and, at the same time, they have also a higher probability of EDSS improvement, suggesting a generalized higher proportion of EDSS change detection (in both directions) in such regions. This observation supports the hypothesis of a heterogeneity in the methods and criteria used for EDSS progression/improvement assessment, indicating a higher sensitivity in detecting changes in USA/Canada and Europe vs East Europe. It is difficult to speculate about the reasons for these differences. A higher sensitivity can be motivated by a higher confidence in EDSS assessment or in longer times dedicated to the clinical visits in Europe and USA/Canada.

This observation has implications for planning clinical trials and for the interpretation of results of observational studies. Many recent clinical trials, in fact, have been planned to recruit patients mostly in Eastern Europe. This choice was motivated by an expected higher recruitment rate in such regions due to the lower number of competitive trials and to limited access to highly expensive drugs in such geographical regions. This analysis reveals that the enrollment in Eastern Europe can in large part account for the decreased number of progression events detected in RRMS in many recent trials[4-6], that made them underpowered to detect any treatment effect on EDSS progression. The implications of such heterogeneities are even more relevant when evaluating the results of observational studies of comparative effectiveness of different drugs, especially when run on large multi-national registries. Heterogeneous EDSS assessment criteria in different geographical regions can bias the comparisons of drug effectiveness when the evaluated drugs are not evenly balanced across countries. As an example of such a situation, in the MSbase study comparing alemtuzumab to other drugs[15] the alemtuzumab cohort was entirely enrolled in the UK, while no other MSBase center enrolling patients in the comparators arms was from UK.

The results of this study, reflecting a large dataset of more than 1700 patients from 3 clinical trials, enrolled according to the same inclusion criteria and followed for 2 years under placebo, call for harmonization procedures for MS patients' clinical assessment. Large efforts have been devoted to defining composite outcomes to improve the clinical assessment of MS patients;[9,16] the MSFC,

incorporating 3 functional measures, reflecting cognition, ambulation and upper limb function, is an example of such an effort. However these attempts to create new measures were mainly focused on increasing the sensitivity of the outcome and its ability to better assess multiple disease aspects. Despite these efforts, EDSS progression remains the most commonly used outcome for disability assessment both in clinical trials and in clinical practice. This study shows that not only sensitivity of EDSS progression must be improved, but also, and perhaps more importantly, its objectivity in measurement.

At a time when disease progression and disability represent the targets for future interventions in MS, establishment of more quantitative and objective outcome measures remain a key priority for researchers in the field.

Acknowledgements

This study received no funds. TEVA allowed to re-analyze 3 clinical trial datasets.

Author Contributions

FB, AS, LC, and MPS contributed to the conception and design of the study; FB, IM, APT, JRS, TL and MPS contributed to the acquisition and analysis of data; FB, AS, LC, IM, and MPS contributed to drafting the text and preparing the figures.

Potential Conflicts of Interest

FB, AS, LC, IM, MPS have nothing to disclose.

TL and APT are TEVA employees.

JRS was a TEVA employer at the time of data analysis.

References

1. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology. 1983; 33(11):1444-52.
2. Haber A, LaRocca NG. eds. Minimal Record of Disability for multiple sclerosis. New York: National Multiple Sclerosis Society; 1985.
3. Hauser SL, Bar-Or A, Comi G et al. Ocrelizumab versus Interferon Beta-1a in Relapsing Multiple Sclerosis. N Engl J Med. 2017 Jan 19;376(3):221-234.
4. ECTRIMS Online Library. Comi G. Oct 27, 2017; 202596
5. Cohen JA, Arnold DL, Comi G et al. Safety and efficacy of the selective sphingosine 1-phosphate receptor modulator ozanimod in relapsing multiple sclerosis (RADIANCE): a randomised, placebo-controlled, phase 2 trial. Lancet Neurol 2016;15(4):373-381
6. ECTRIMS Online Library. Comi G. Oct 27, 2017; 202595
7. Röver C, Nicholas R, Straube S, Friede T. Changing EDSS Progression in Placebo Cohorts in Relapsing MS: A Systematic Review and Meta-Regression. PLoS One 2015; DOI: 10.1371/journal.pone.0137052
8. Kiesseier B, Pozzilli C. Assessing walking disability in multiple sclerosis. Mult Scler 2012;18:914–924.
9. Cutter GR, Baier ML, Rudick RA et al. Development of a multiple sclerosis functional composite as a clinical trial outcome measure. Brain. 1999 May;122 (Pt 5):871-82.
10. Fisher JS, Rudick RA, Cutter GR et al. The multiple sclerosis functional composite (MSFC): an integrated approach to MS clinical outcome assessment. Multiple Sclerosis, 5 (1999), pp. 244-250
11. Vollmer TL, Sorensen PS, Selmaj K, et al. A randomized placebo-controlled phase III trial of oral laquinimod for multiple sclerosis. J Neurol. 2014; 261(4):773-83.
12. Comi G, Jeffery D, Kappos L, et al. Placebo-Controlled Trial of Oral Laquinimod for Multiple Sclerosis. N Engl J Med. 2012; 366(11):1000-9.

13. Polman CH, Reingold SC, Edan G, et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria." Ann Neurol 2005;58:840-6.

14. 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.

15. Kalincik T, Brown JWL, Robertson N et al. Treatment effectiveness of alemtuzumab compared with natalizumab, fingolimod, and interferon beta in relapsing-remitting multiple sclerosis: a cohort study. Lancet Neurol. 2017 Apr;16(4):271-281.

16. La Rocca NG, Hudson LD, Rudick R et al. The MSOAC approach to developing performance outcomes to measure and monitor multiple sclerosis disability. Mult Scler. 2017 Aug 1:1352458517723718. doi: 10.1177/1352458517723718. [Epub ahead of print]

Figure 1. PANEL A) Kaplan-Meier plot for the risk of 3-month confirmed EDSS progression according to geographical region; PANEL B) Kaplan-Meier plot for the probability of 3 month-confirmed EDSS improvement according to geographical region

Figure 2. PANEL A) Kaplan-Meier plot for the risk of 3-month-confirmed Time 25 Foot Walk progression according to geographical region; PANEL B) Multiple Sclerosis Functional Composite change over 2 year according to geographical region

Table 1. Demographical and clinical characteristics of 1746 RRMS placebo patients enrolled in the ALLEGRO, BRAVO and CONCERTO trials, by geographical region

Baseline variables[a]	**EAST EUROPE* N=1310**	**EUROPE* N=291**	**USA/CANADA* N=145**	**p-value***
Gender (Male)	449 (34.3%)	88 (30.2%)	32 (22.1%)	**0.008**
Age (Years)	36.5 ± 9.2	38.4 ± 8.8	40.5 ± 9.2	**<.0001**
Time to Diagnosis (Years)	3.6 ± 4.1	4.8 ± 5.5	5.2 ± 5.1	**0.0015**
Time to first symptoms (Years)	6.6 ± 5.4	7.8 ± 6.5	8.5 ± 7.1	**0.0051**
EDSS	2.8 ± 1.2	2.1 ± 1.2	2.5 ± 1.0	**<.0001**
2-years prior relapses	1.9 ± 0.9	1.8 ± 1.0	1.8 ± 0.9	**0.0018**
1-years prior relapses	1.3 ± 0.6	1.3 ± 0.7	1.2 ± 0.6	0.3658
MSFC Timed 25 Foot Walk	6.9 ± 4.5	5.4 ± 2.4	7.1 ± 14.8	**<.0001**
Brain Volume, cm^3	1509.4 ± 117.5	1553.1 ± 115.6	1582.3 ± 101	**<.0001**
T2 lesion volume, cm^3 (log-transformed)	1.7 ± 1.2	1.3 ± 1.2	1.0 ± 1.4	**<.0001**
T1 lesion volume, cm^3 (log-transformed)	0.9 ± 1.4	0.1 ± 1.5	-0.4 ± 1.5	**<.0001**
GD+ lesion presence	540 (41.2%)	104 (35.9%)	47 (32.4%)	**0.04**

EDSS: Expanded Disability Status Scale; MSFC: Multiple Sclerosis Functional Composite; Gd+: gadolinium-enhancing;

[a] Data reported as mean (SD) and percentage for continuous and categorical variables, respectively

*East Europe: Belarus, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Estonia, Georgia, Hungary, Latvia, Lithuania, Macedonia, Montenegro, Poland, Republic of Moldova, Republic of Serbia, Romania, Russian Federation, Slovakia, Turkey, Ukraine.
Europe: Austria, France, Germany, Greece, Israel, Italy, Netherlands, Spain, Sweden, United Kingdom.
USA/Canada: Canada, Republic of Korea, South Africa, USA.
A small number of patients (n=17) was enrolled outside Europe, USA or Canada, and they were grouped in the USA/Canada group. The analysis results did not change excluding them from the sample. Israel (n=27) was included in the Europe group, and its exclusion did not affect the results.