Journal of the Neurological Sciences 274 (2008) 76–79

Contents lists available at ScienceDirect

Journal of the Neurological Sciences

journal homepage: www.elsevier.com/locate/jns

Reconsidering clinical outcomes in Multiple Sclerosis: Relapses, impairment, disability and beyond

M. D'Souza, L. Kappos, A. Czaplinski *

Department of Neurology, University Basel, Switzerland

ARTICLE INFO

Article history:
Received 24 July 2008
Received in revised form 14 August 2008
Accepted 15 August 2008
Available online 24 September 2008

Keywords:
Multiple Sclerosis
Clinical trials
Outcome measures
Relapses
EDSS
MSFC

ABSTRACT

There is an increasing number of clinical trials testing new compounds which act at different stages of Multiple Sclerosis (MS). To prove their effectiveness several clinical outcome measures are used. The overall quality of clinical trials is increasing steadily due to the growing experience in this area, the increasing awareness of quality standards in the MS community and the more stringent requirements of regulatory authorities for approval of new treatments. Each successful clinical trial provided additional information that could be incorporated into the design of subsequent studies to improve their quality. However, the choice of appropriate outcome measures still presents major challenges. For an individual patient improvement or stability of their disability and to a lesser extent the relapse rate, are the main targets of treatment. As there is yet no scale or assessment, which objectively covers all major issues, it is recommended to use multiple instruments and endpoints as secondary outcome measures.

1. Introduction

There is an increasing number of clinical trials testing new compounds which act at different stages of Multiple Sclerosis (MS). To prove their effectiveness several clinical outcome measures are used. The overall quality of clinical trials is increasing steadily due to the growing experience in this area, the increasing awareness of quality standards in the MS community and the more stringent requirements of regulatory authorities for approval of new treatments [1]. Each successful clinical trial provided additional information that could be incorporated into the design of subsequent studies to improve their quality [2]. However, the choice of appropriate outcome measures still presents major challenges and controversies about availability of validated measures for unremitting disability [3], or the necessary duration of treatment exposure.

2. Relapses as outcome measures for clinical trials in MS

Relapse related outcome measures for clinical trials in MS are the annualized relapse rate, the number of relapse-free patients and the time to first relapse under treatment. Relapses are defined as new or worsening neurological symptoms with duration of more than 24 h, preceded by a minimum of 30 days of clinical stability or improvement, confirmed by objective findings on neurological examination. Symptoms should not be due to an alternative explanation [4]. In clinical trials the neurological abnormalities must usually be present for a minimum of 48 h and should result in an increase in the EDSS (Expanded Disability Status Scale) score by more than 0.5 points or in an increase of more than 1 point in one of the 7 Functional System Scores as compared to previous evaluation. Sometimes changes in the more difficult to assess vegetative and cerebral FS, are not counted. According to the McDonald criteria (2001, 2005) there is a 30-day limit from first relapse manifestation, after which newly appearing symptoms may be counted as a new relapse. However, this does not necessarily correspond with the actual biology of the disease, where often more than one areas of active inflammation in the CNS exist, each of which runs an independent time course [5]. Relapses are subject to reporting bias and cannot always easily be distinguished from "pseudo exacerbations" precipitated by heat exposure, infection, fever, fatigue and/or changes in mood.

In most studies relapse rates are related to age and time since onset, and therefore any drug which may be able to modify relapse rates has the greatest potential for a population-impact, in patients below 40 years [6] Studies lacking a randomization or an internal control group must consider that relapse rates decline at different rates over time according to the patient's onset age and that a relapse-quiescent period in MS is not uncommon [6].

Pronounced placebo effects on relapse rates have been observed in nearly all relapsing remitting MS controlled trials when comparing pre-study- with on study-exacerbation rates in the placebo group. Starting with 7% in a study with intravenous Immunoglobuline, which was not fully blinded [7,8] and ranging up to 56% in a study using oral MBP (myelin basic protein) which was probably indistinguishable from placebo [9]. Review of this graph (Fig. 1) strongly suggests a close dependence on efficiency of blinding. Other effects like regression to the mean or the impact of comprehensive care provided to trial participants must also be taken into consideration [10]. It has been

* Corresponding author. Tel.: +41 61 2654151.
E-mail address: CzaplinskiA@uhbs.ch (A. Czaplinski).

0022-510X/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jns.2008.08.023

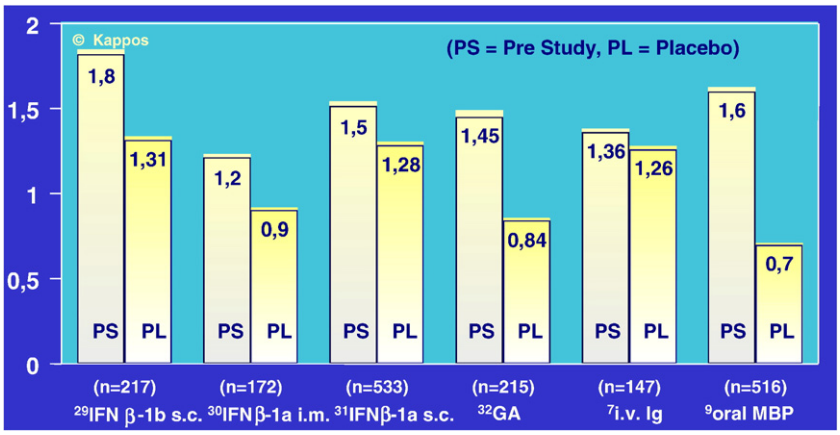

Fig. 1. Placebo-effect on the improvement of the relapse rates [29–32].

shown that higher frequency of clinical visits increases the probability to detect clinical relapses and may consequently result in higher relapse rates [11].

Annualized relapse rates may not be suitable for trials with expected high drop out rates since patients who had 1 or 2 relapses and discontinued the study may bias the results by increasing the annualized relapse rate due to extrapolating.

Time to first relapse under treatment is a robust parameter, even in trials with high drop out rate. However, it does not make use of the second and next relapses in the course of a trial. Therefore, it may favor drugs with an immediate effect, as compared to those with a more delayed but perhaps in the long run better sustained effect.

3. EDSS, Expanded Disability Status Scale

For an individual suffering from Multiple Sclerosis, the main target of treatment is a reduction of impairment and disability or at least a slowing of disability progression. A scale used to measure impairment and disability in MS should be multidimensional, applicable across the range of disease severity, reliable and easy to use. Moreover it should have a predictive value for sustained changes and its individual components should be sensitive to change over time. The EDSS [12] is the most frequently used scale for rating disability in MS and up to now the only one accepted by drug agencies like FDA and EMEA. It was first developed in 1955 as the Disability Status Scale by John Kurtzke, a 10-step overall disability rating scale, devised to evaluate Isoniazid as a possible treatment, and initially validated on 500 hospitalized second world war veterans suffering from MS. In the following years some definitions were modified and since 1983, half steps were introduced to create the EDSS for a total of 20 steps, ranging from 0 to 10, separated by half-points. It summarizes findings in 7 main neurological systems, so called Functional Systems (FS): visual, brainstem, pyramidal, cerebellar, sensory, bowel/bladder and cerebral. In addition evaluation of walking distance and required assistance contributes as an independent factor to the EDSS score. To point out its advantages one should note that the EDSS is the longest established and most widely used rating scale (for >50 years), it is familiar to most MS specialists, it allows for a simple comparison on a scale from 0 to 10 and most importantly, it is easy to use. The scale has well known disadvantages. It shows a poor inter- and intra-rater reliability, especially in its lower part [13], it is an ordinal and not metric scale with bimodal distribution of scores, both crossectionally and if time on each step is considered. It relies heavily upon ambulation with a poor assessment of upper limb function in higher EDSS-steps and is insensitive to cognitive decline.

Time to change by one step in the EDSS in a "survival analysis" is used in the majority of studies as an outcome measure. Confirmation after 3 or 6 months is required in an attempt to reduce the impact of direct relapse related ephemere variability and to depict sustained or even unremitting progression. We have to consider that "survival analysis" has been developed for irreversible events. Concerning EDSS changes in MS, even if confirmed after 3 or 6 months this is not the case [3]. Although terms like "*sustained* progression" suggest irreversibility of progression, a high percentage of patients even of those with secondary progressive disease, will return to a lower EDSS score in their further disease course [3]. Irrespective of these caveats EDSS remains a gold standard for clinical assessments in therapeutic trials. To increase reliability for clinical trials standardized definitions and training tools for scoring FS and the EDSS have been developed [14]. It is recommended that the same physician evaluates individual patients throughout a trial based on a standardized comprehensive neurological examination and the walking distance should be measured under standardized conditions and not just estimated.

4. Multiple Sclerosis Functional Composite (MSFC)

The National Multiple Sclerosis Society Clinical Assessment Task Force (NMSS Task Force) presented 1997 the Multiple Sclerosis Functional Composite (MSFC). In response to the weaknesses of the EDSS, the MSFC was developed as a metric scale to increase mainly the reliability and the sensitivity to change over time when used as an outcome measurement in clinical trials.

The MSFC includes quantitative tests of ambulation (Timed 25-Foot Walk), arm function (9-Hole Peg Test) and cognitive function (Paced Auditory Serial Addition Test). It indicates how many standard deviations above or below the mean of a standard population the patient scores in each of these tests. The component raw scores are converted to *z*-scores based on the mean and standard deviation of a reference population. As a metric scale it may be more informative than an ordinal scale (e.g. EDSS score) and facilitates statistical analysis of longitudinal data. It relies on well standardized procedures. In contrast to the EDSS, it captures information on cognitive function. MSFC changes have been found to reflect the severity of the disease, as reported by patients through quality-of-life questionnaires (QoL-54 [1]). In one trial (IMPACT study), in which it was defined as primary outcome, it showed a higher sensitivity to change than the EDSS [15]. In several other completed and ongoing trials, MSFC was a secondary endpoint. Interestingly, the assumption that it has a higher sensitivity to change has not generally been confirmed [16,17]. The main disadvantages are the influence of visual deficit on the 9-Hole Peg Test, and of speech problems and learning effects on its cognitive component, the Paced Auditory Serial Attention Test (PASAT). Results depend on the reference population that is chosen for the calculation of *z*-scores (original sample from which the scale was derived or