*Review Article*

# Outcome Measures in Relapsing-Remitting Multiple Sclerosis: Capturing Disability and Disease Progression in Clinical Trials

**Amy M. Lavery,[1] Leonard H. Verhey,[2] and Amy T. Waldman[1,3]**

[1] *Division of Neurology, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA*

[2] *The Pediatric Demyelinating Disease Program, Program in Neuroscience & Mental Health, The Hospital for Sick Children, University of Toronto, Toronto, ON, Canada M5G 1X8*

[3] *Departments of Neurology and Pediatrics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA 19104, USA*

Correspondence should be addressed to Amy T. Waldman; waldman@email.chop.edu

Multiple sclerosis (MS) is a chronic inflammatory and neurodegenerative disease that manifests as acute relapses and progressive disability. As a primary endpoint for clinical trials in MS, disability is difficult to both characterize and measure. Furthermore, the recovery from relapses and the rate of disability vary considerably among patients. Given these challenges, investigators have developed and studied the performance of various outcome measures and surrogate endpoints in MS clinical trials. This review defines the outcome measures and surrogate endpoints used to date in MS clinical trials and presents challenges in the design of both adult and pediatric trials.

## 1. Introduction

Multiple sclerosis (MS) is a leading cause of morbidity and disability in young adults. There has been significant research in the development and use of outcome measures and surrogate endpoints for MS clinical trials. Such investigations have been necessary due to the evolution in clinical trial design in MS. Early trials sought to determine whether disease-modifying therapies could alter the number of relapses compared to placebo in patients already diagnosed with relapsing-remitting MS. Clinical trials later evolved to determine whether early treatment could delay the second attack in the relapsing portion of the disease compared to placebo. Current trials of new therapies, including oral medications, are mostly active-arm comparison trials, which require more sensitive metrics to determine efficacy. Alternatively, larger sample sizes may be required which may impact the feasibility of a study. Trials of symptomatic therapies, such as 4-aminopyridine for ambulation, and future trials studying agents for neuroprotection and neurorepair, will require different outcome measures to determine neurologic functional recovery compared to those currently being used in MS studies measuring clinical disease. In addition to their use in clinical trials, outcome measures and surrogate endpoints have been applied to the individual patient to evaluate disease progression and the need for a change in therapy.

MS is a heterogeneous disease with considerable variation in the clinical, radiographic, genetic, pathologic, and biologic features among patients. Such differences can present challenges in clinical trial design. It is not uncommon to have multiple primary and secondary endpoints reflecting the diversity. Furthermore, surrogate endpoints are often used to determine treatment efficacy. This review will define outcome measures and surrogate endpoints, discuss the metrics that have been used to date in MS trials, and present challenges in the design of adult and pediatric MS clinical trials.

## 2. Primary Outcome Measures

Table 1 lists the various adult clinical trials, along with their primary and secondary outcome measures, for relapsing-remitting MS. Selecting an appropriate outcome measure for clinical trials is important in determining whether

TABLE 1: Outcome measures selected for clinical trials in relapsing-remitting multiple sclerosis.

| Study name/group | Drug treatment | N | Publication year | Primary outcome | Secondary outcomes |
|---|---|---|---|---|---|
| Multiple sclerosis collaborative research group | Interferon beta-1a (avonex) | 301 | 1990 | Worsening in disability defined as deterioration by 1.0 point on EDSS | (i) Time to first relapse (ii) Number and volume of gadolinium enhancing lesions |
| Interferon beta-1b multiple sclerosis study group | Interferon beta-1b (Betaseron) | 372 | 1993 | (i) Annualized relapse rate (ii) Proportion of relapse-free patients | (i) Time to first relapse (ii) Relapse duration and severity by NRS score (iii) Change in EDSS and NRS scores from baseline (iv) Quantitative disease burden by MRI (% change in lesion area) (v) Disease activity as measured by MRI (annual rate of new or enlarging lesions) |
| Copolymer 1 multiple sclerosis study group | Copolymer 1 (glatiramer acetate, Copaxone) | 251 | 1995 | Average number of relapses at 2 years | (i) Proportion of relapse free patients (ii) Time to first relapse (iii) Proportion without sustained disease progression as measured by the EDSS at 3 months (iv) Average change in EDSS and ambulation index |
| PRISMS | Interferon beta 1a (Rebif) | 560 | 1998 | Number of relapses | (i) Times to first and second relapse (ii) Progression of disability as measured by sustained worsening of EDSS, ambulation index, or arm function index at 3 months (iii) Proportion of relapse-free patients (iv) MRI burden (proton density T2 MRI, Number of T2 active lesions) |
| CHAMPS | Interferon beta 1a (Avonex) | 383 | 2000 | Conversion to CDMS | (i) Volume of T2 lesions (ii) Number of new or enlarging T2 lesions (iii) Number of gadolinium enhancing T1 lesions |
| ETOMS | Interferon beta 1a (Rebif) | 309 | 2001 | Conversion to CDMS | (i) Change in the SNRS score (ii) Time to second relapse (iii) Number of new or enlarging T2 lesions (iv) Number of enhancing T1 lesions (v) Number of patients without MRI activity (no new or enlarging T2 lesions or enhancing lesions throughout the study) (vi) Yearly changes of hyperintense T2 lesion volume |
| EVIDENCE | Interferon beta-1a (Rebif) | 677 | 2005 | (i) Annualized relapse rate; (ii) Proportion remaining relapse free at 2 years | (i) Change in new, enlarged, or reappearing lesions (ii) Proportion of active scans per patient (new, enlarged, or reappearing lesions) |
| BENEFIT | Interferon beta-1b (Betaseron) | 487 | 2006 | (i) Time to conversion to CDMS; (ii) Time to McDonald defined MS | (i) Cumulative number of new or enhancing T2 lesions (ii) Change in T2 lesion volume (iii) Cumulative volume of gadolinium enhancing lesions (iv) Change in EDSS and MSFC scores |

Table 1: Continued.

| Study name/group | Drug treatment | N | Publication year | Primary outcome | Secondary outcomes |
| --- | --- | --- | --- | --- | --- |
| CHAMPIONS | Interferon beta 1a (Avonex) | 203 | 2006 | Conversion to CDMS | (i) Number of confirmed relapses<br>(ii) Disease course classification<br>(iii) Neurological disability as measured by EDSS at 5 years<br>(iv) Number of new or enlarging T2 lesions<br>(v) Changes in T2 volume<br>(vi) Percentage with gadolinium enhancing lesions |
| AFFIRM | Natalizumab (Tysabri) | 856 | 2006 | (i) Annualized relapse rate;<br>(ii) Progression of disability as measured by sustained worsening of EDSS at 12 weeks | (i) Number of new or enlarging T2 lesions (Year 1)<br>(ii) Number of gadolinium enhancing lesions (Year 1)<br>(iii) Proportion of relapse free patients (Year 1)<br>(iv) Annualized relapse rate (Year 2)<br>(v) Volume of lesions (Year 2)<br>(vi) Number of new hypointense lesions (Year 2)<br>(vii) MSFC as measure of disability (Year 2) |
| SENTINEL | Natalizumab and Interferon beta-1a (Tysabri and Avonex) | 1003 | 2006 | (i) Annualized relapse rate;<br>(ii) Progression of disability as measured by sustained worsening of EDSS at 12 weeks | (i) Number of new or enlarging T2 lesions (Year 1)<br>(ii) Number of gadolinium enhancing lesions (Year 1)<br>(iii) Proportion of relapse free patients (Year 1)<br>(iv) Annualized relapse rate (Year 2)<br>(v) Volume of lesions (Year 2)<br>(vi) Number of new hypointense T1 lesions (Year 2)<br>(vii) MSFC as measure of disability (Year 2) |
| REGARD | Interferon beta-1a and Glatiramer acetate (Rebif and Copaxone) | 764 | 2008 | Time to relapse at 96 weeks | (i) Number of new or enhancing lesions<br>(ii) Number of gadolinium enhancing lesions<br>(iii) Change in T2 lesion volume<br>(iv) Change in whole brain volume<br>(v) Other relapse outcomes<br>(vi) Progression of disability as measured by sustained worsening of EDSS at 6 months |
| BECOME | Interferon beta-1b and glatiramer acetate (Betaseron and Copaxone) | 75 | 2009 | Number of combined active lesions in the first year (total contrasting enhancing lesions plus new non-enhancing lesions that have appeared since most recent examination) | New lesions per subject in year 1 and 2 |
| PreCISe | Glatiramer acetate (Copaxone) | 619 | 2009 | Time to conversion to CDMS | (i) Number of new T2 lesions<br>(ii) Baseline-adjusted T2 lesions volume at last scan<br>(iii) Brain atrophy- percentage change from baseline<br>(iv) Proportion of patients converted to MS |
| TRANSFORMS | Fingolimod (Gilenya) | 1153 | 2010 | Annualized relapse rate | (i) Number of new or enlarged T2 hypointense lesions at 1 year<br>(ii) Time to confirmed disability progression as measured by EDSS at 3 months |

Table 1: Continued.

| Study name/group | Drug treatment | N | Publication year | Primary outcome | Secondary outcomes |
|---|---|---|---|---|---|
| FREEDOMS | Fingolimod (Gilenya) | 1272 | 2010 | Annualized relapse rate accompanied by change in EDSS | (i) Time to confirmed disability progression as measured by EDSS at 3 months<br>(ii) Time to first relapse and disability progression at 6 months<br>(iii) Change in EDSS and MSFC at 2 years<br>(iv) Number of patients with gadolinium enhancing lesions<br>(v) Number with new or enlarging T2 weighted lesions<br>(vi) Proportion of patients with gadolinium-enhancing lesions and T2-weighted lesions<br>(vii) T2 and T1 lesion volume<br>(viii) Change in brain volume at 2 years |
| TEMSO | Teriflunomide (Aubagio) | 1088 | 2011 | Annualized relapse rate | (i) Progression of disability as measured by EDSS at 12 weeks<br>(ii) Total lesion volume<br>(iii) Volume of hypointense T1 lesions<br>(iv) Number of gadolinium enhanced T1 lesions<br>(v) Number of active lesions (new gadolinium enhancing on T1 images or new or enlarged lesions on T2)<br>(vi) Brain parenchymal fraction |
| DEFINE | BG-12 (dimethyl fumarate, Tecfidera) | 1237 | 2012 | Proportion of patients who had a relapse at 2 years | (i) Number of gadolinium enhancing lesions<br>(ii) New or enlarged hyperintense T2 lesions<br>(iii) Annualized relapse rate<br>(iv) Time to progression of disability as measured by EDSS at 12 weeks |
| CONFIRM | BG-12 and glatiramer acetate (Tecfidera, Copaxone) | 308 | 2012 | Annualized relapse rate at 2 years | (i) New or enlarging hyperintense lesions on T2<br>(ii) Number of hypointense T1 lesions<br>(iii) Proportion of patients with relapse<br>(iv) Time to disability progression as measured by EDSS at 2 years |

PRISMS: Prevention of Relapses and Disability by Interferon beta-1a Subcutaneously in Multiple Sclerosis; CHAMPS: Controlled High Risk Avonex Multiple Sclerosis Study; ETOMS: Early Treatment of Multiple Sclerosis Study Group; EVIDENCE: Evidence of Interferon Dose-Response: European North American Comparative Efficacy; BENEFIT: Betaferon in Newly Emerging Multiple Sclerosis for Initial Treatment; CHAMPIONS: Controlled High Risk Avonex Multiple Sclerosis Prevention Study In Ongoing Neurological Surveillance; AFFIRM: Natalizumab Safety and Efficacy in Relapsing Remitting Multiple Sclerosis; SENTINEL: The Safety and Efficacy of Natalizumab in Combination with Interferon Beta-1a in Patients with Relapsing Remitting Multiple Sclerosis; REGARD: Rebif versus Glatiramer Acetate in Relapsing MS Disease; BECOME: Betaseron versus Copaxone in Multiple Sclerosis with Triple-Dose Gadolinium and 3-Tesla MRI Endpoints; PreCISe: Effect of glatiramer acetate on conversion to clinically definite multiple sclerosis in patients with clinically isolated syndrome; TRANSFORMS: Trial Assessing Injectable Interferon versus FTY720 Oral in Relapsing-Remitting Multiple Sclerosis; FREEDOMS: FTY720 Research Evaluating Effects of Daily Oral Therapy in Multiple Sclerosis; TEMSO: Teriflunomide Multiple Sclerosis Oral; DEFINE: Determination of the Efficacy and Safety of Oral Fumarate in Relapsing-Remitting Multiple Sclerosis; CONFIRM: Comparator and an Oral Fumarate in Relapsing-Remitting Multiple Sclerosis.
EDSS: Extended Disability Status Scale; MSFC: Multiple Sclerosis Functional Composite; CDMS: Clinically Definite Multiple Sclerosis; SNRS: Scripp's Neurological Rating Scale; NRS: Neurological Rating Scale.
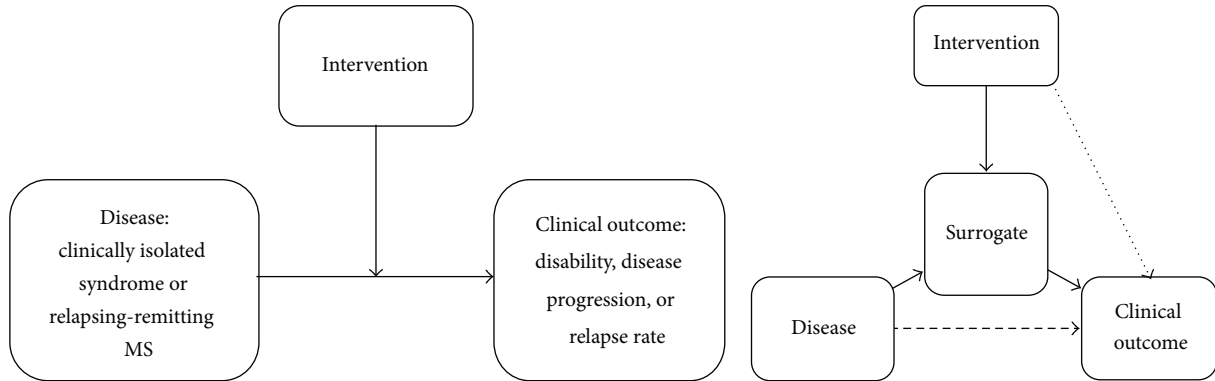
(a) Primary outcome measures for clinical trials are carefully selected. Clinically relevant outcomes that directly impact the patient are ideal for Phase 3 clinical trials. In MS trials, disability (as measured by the Expanded Disability Status Scale), disease progression (such as the conversion to clinically definite MS), and annualized relapse rate are commonly chosen as the primary endpoint

(b) Ideal relationship (solid lines) between an exposure and outcome, through a surrogate endpoint. The intervention does not affect the outcome without impacting the surrogate. A surrogate may fail if it has other effects on the outcome not mediated by the surrogate (dotted line) or the intervention only impacts the pathway involving the surrogate, and other causal mechanisms exist (dashed line)
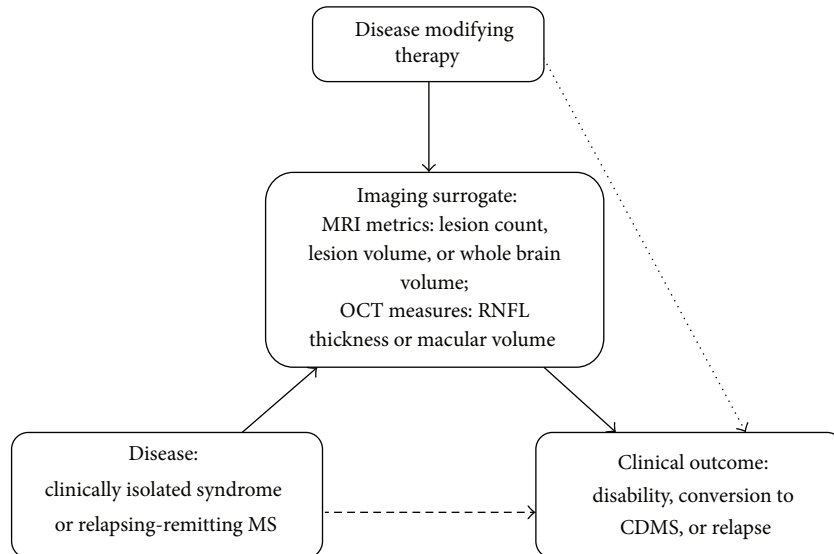
(c) While imaging has played a significant role in the diagnosis and prognosis for MS patients, MRI metrics have not been accepted as surrogates for Phase 3 clinical trials. For a surrogate to be valid, the effects of the intervention must be entirely mediated through the surrogate. Imaging surrogates are not sufficient to explain the complex pathophysiology resulting in disability, disease progression, and clinical relapses in MS

FIGURE 1: Clinical outcome measures and surrogate endpoints in Phase 3 clinical trials.

the intervention is actually modifying the disease course (Figure 1(a)) and evaluating the risk-to-benefit ratio [1]. A measure with poor reliability or interpretability may lead to inaccurate results and improper use of treatments. According to Fleming and DeMets, "For phase 3 trials, the primary end point should be a clinical event relevant to the patient, that is, the event of which the patient is aware and wants to avoid [2]." Disability is a patient-centered outcome that is appropriate for MS trials, akin to survival in cardiac disease or cancer trials.

## 2.1. Quantifying Disability

### 2.1.1. Expanded Disability Status Scale.

While disability may be considered an ideal primary endpoint based on the definition above, disability can be difficult to both define and measure. Patients with MS develop various degrees of visual, cognitive, physical, and psychological disability, and it is difficult, for example, to compare disability related to chronic fatigue affecting one patient to paralysis suffered by another patient. The Expanded Disability Status Scale (EDSS) was created to

address this issue and quantify a patient's disability based on scores of eight functional systems [3]. The scale ranges from 0 (normal neurological exam) to 5 (ambulatory without aid for 200 meters) to 10 as the most serious outcome, death due to MS. While this is often a measure used in clinical trials and regular clinical examinations, the scale is not very sensitive to change and lacks a clinically defined relevant change [4].

Clinical trials have defined disability progression as an increase in the EDSS scale of 0.5–1.0 point after 3 or 6 months. Ebers et al. challenge the use of the EDSS, concluding that such change over a short time frame is likely measurement error or random variation and not the result of sustained disability [5]. Consequently, the EDSS may be ineffective at predicting disease progression in RRMS patients after a short time period. The authors propose that longer trials (with a duration of at least 1 year) with greater changes in the EDSS scores (>1-2 points) may better capture patients with sustained disability. This and other studies have shown that the EDSS often has problems with reliability and validity. Interrater variation has been reported to be greater than a 1-point increase in the EDSS about 40% of the time [5, 6]. Despite the limitations of the EDSS, it is accepted as a "true clinical efficacy measure" and has been widely accepted as a primary outcome measure for Phase 3 clinical trials in MS [1].

*2.1.2. Multiple Sclerosis Functional Composite.* MS researchers recognized the limitations of the EDSS and sought to develop a more responsive clinical outcome measure. The National MS Society Task Force developed the Multiple Sclerosis Functional Composite (MSFC), which uses simple measures to help assess patient functionality [7, 8]. The MSFC consists of a 25-foot timed walk, a 9-hole peg board, and paced auditory serial addition task (PASAT), which reflect ambulation, hand function, and speed of processing. Tests can be administered by nonclinicians and the total time required to take the test is about 15 minutes [9]. A z-score is created based on the average value for each test, allowing for direct comparison of measures with different units [8, 9]. The z-score reports the standard deviation of the individual's performance compared to the mean of the reference population [8].

MSFC scores correlate with EDSS scores. In a study of 300 MS patients, Miller et al. found that the MSFC was highly correlated with the EDSS with a Spearman correlation coefficient of −0.80 [10]. The individual components of the MSFC (arm, leg, and cognitive) also correlate with composite EDSS scores (Spearman rank correlation coefficients −0.33, −0.52, and −0.23, resp.) [8]. Likewise, Rudick et al. determined that changes in MSFC scores (defined as a 15% or 20% change in a single component of the MSFC, sustained for 3 months) correlated with EDSS scores and relapse rates [11].

Furthermore, baseline MSFC scores in clinical trials and change in MSFC score during a 2-year period were shown to predict future disability. Decreased baseline and worsening MSFC scores early in a phase III clinical trial of interferon beta 1a predicted poor outcomes, including physical disability, progression to secondary progressive disease, brain atrophy, and decreased quality of life at a mean of 8.1 years [12].

Similar to the EDSS, however, researchers have indicated difficulty in quantifying a meaningful change in the MSFC components. The tests are also weighted equally in the composite score, which may not accurately capture the disease progression as a whole. Individual components of the test may change over time, and the composite score may not reflect these differences [13]. In a phase II trial of Rituximab as add-on therapy, Naismith et al. found that improvements in MSFC scores were largely attributed to the PASAT [14]. However, improvements in the PASAT may reflect practice effects rather than the true changes in cognitive dysfunction. Also, despite the efforts towards developing the MSFC, it has not been incorporated into many clinical trials as a primary endpoint. Although used in some of the trials listed in Table 1, the MSFC has generally been used as a secondary endpoint along with the EDSS or in further secondary analyses [11, 15].

*2.1.3. Low-Contrast Letter Acuity.* The MSFC has also been criticized for not including a visual measure; however, at the time of its development, a sensitive test to accurately capture vision loss in MS patients had not been identified. The commonly used Snellen (high-contrast) charts are known to have limited capacity to measure small changes in visual dysfunction [16]. Over the past 10 years, cross-sectional and longitudinal studies of adult MS patients and disease-free controls have demonstrated that low-contrast letter acuity (LCLA) is the most sensitive test for identifying visual dysfunction in patients with MS [4, 17–19] and have been proposed as the visual component of the MSFC [13]. As such, LCLA was performed in the AFFIRM and SENTINEL trials [15, 20, 21]. While high-contrast testing did not differ between study groups, visual function measured using LCLA worsened (defined as a 2-line or 10-letter worsening of visual acuity) over a two-year period in the placebo group. Moreover, Balcer et al. also demonstrated a decreased risk of sustained visual loss in the natalizumab group compared to placebo (a risk reduction of 47% for 2.5% LCLA, $P < 0.001$; and 35% for 1.25% LCLA, $P = 0.008$ for 1.25% LCLA in the AFFIRM trial). The authors concluded that LCLA has the ability to determine treatment effects and proposed that LCLA should be incorporated into future clinical trials [15].

*2.2. Relapse Rate and Disease Progression.* Sustained disability may not be detected early in the disease course for MS patients. Given the additional challenges in quantifying disability, annualized relapse rate (ARR), time to first relapse (TTFR), and conversion to clinically-definite MS (CDMS) have been the most common primary outcome measures used in clinical trials.

Annualized relapse rate is often included as an outcome measure for clinical trials because it is easy to quantify, and prevention of relapses benefits patients immediately. Relapses are generally defined as neurologic symptoms lasting more than 24 hours which occur at least 30 days after the onset of a preceding event [22], though definitions can vary slightly by study. Relapse rate early in the disease is thought to predict future disability [23, 24]. However, the probability of relapses is not a constant function over time. Patients are usually

enrolled in a trial at the time of MS diagnosis when the probability for relapses is high, and, as time progresses, this probability decreases due to the regression to the mean phenomenon. In addition, relapses may be separated by several years, which may be very time consuming and costly for a clinical trial. Sormani et al. determined that in order to have enough power to detect a significant reduction in relapses, a clinical trial needs to last at least 1 year, but this measure may also be less meaningful than looking at total number of relapses over a longer period of time [25]. They also suggest that due to low relapse rates recorded in recent trials, the sample size required for new studies may not be feasible [25, 26].

Time to first relapse has been recently proposed as a primary endpoint for clinical trials. Time to first relapse is an appealing alternative to annualized relapse rate because it generally requires less study time to reach the trial endpoint, and the first relapse is typically more accurately documented than the subsequent relapses included in the ARR calculation. Additionally, Sormani et al. determined that TTFR may require a lower number of subjects than trials based on the annualized relapse rate [25]. The authors also note that for clinical trials where it may not be ethical to use a placebo for extended periods of time, using TTFR as the primary outcome measure allows patients enrolled in the placebo arm to switch to the active treatment immediately after a relapse occurs (i.e., the study endpoint).

Conversion to CDMS as a trial endpoint is related to TTFR; for patients enrolled into a trial at the time of an incident demyelinating event (i.e., clinically isolated syndrome [CIS]), the first relapse corresponds to the second clinical attack, or confirmation of CDMS. Early trials of disease-modifying therapies used progression to CDMS as a primary outcome measure. The measure was generally defined as the time to a second clinical attack following the first demyelinating event. While the conversion to CDMS is an ideal outcome measure for many clinical trials, it is often costly and time consuming. Furthermore, with varying definitions of a second attack between studies, even slight differences can affect the reproducibility of treatment efficacy. For example, some studies defined a second attack as neurological symptoms that have persisted for a minimum of 24 hours in the absence of fever with objective findings whereas other studies have required neurological symptoms to persist for at least 48 hours and subjective events were permitted [27–29].

## 3. Surrogate Endpoints

The previously-mentioned outcome measures represent patient-centered outcomes that have been used in Phase 3 clinical trials to determine treatment efficacy. Other secondary measures, often surrogate endpoints, have also been included in the same trials. As defined by Temple: "A surrogate endpoint of a clinical trial is a laboratory measurement or a physical sign used as a substitute for a clinically meaningful endpoint that measures directly how a patient feels, functions, or survives [30]." Various MRI parameters have been proposed as surrogates in MS trials.

*3.1. Magnetic Resonance Imaging.* Magnetic resonance imaging has been incorporated into the diagnostic criteria for MS [31] and serves as a routine paraclinical tool to follow disease progression. Consideration of MRI findings also allows the diagnosis of MS to be made earlier than if relying solely on clinical relapses. The next several paragraphs will address the rationale for including MRI metrics in MS clinical trials followed by a general discussion of the challenges when using MRI measures as surrogates for disease activity.

*3.1.1. MRI Lesion Counts.* Changes in MRI-visible brain lesions reflect changes in the underlying disease pathology and therefore provide theoretical rationale for using MRI lesions as measures of disease activity. MRI lesions are quantified as the number of T1-weighted gadolinium-enhancing lesions, new T2 lesions, or active (i.e., new or enlarging) T2 lesions. The treatment effect on gadolinium-enhancing lesions is highly associated with the treatment effect on active T2 lesions ($R^2 = 0.93$) [32], suggesting that either contrast-enhancing or new T2 lesion endpoints are suitable for monitoring MRI activity in MS clinical trials. It should be noted, however, that using gadolinium-enhancing lesions as an outcome requires monthly MRI scans which in turn increases the cost of the trial. All immunomodulatory agents decrease the number of gadolinium-enhancing and new T2 lesions; the degree and rate of their effect are variable and dependent on the drug's mechanism of action (i.e., at the blood-brain barrier for beta-interferons and more centrally-mediated action for glatiramer acetate). Decreased MRI activity represents the earliest treatment effects in clinical trials and therefore has served as an attractive endpoint in clinical trials.

Lesion load measured early in the disease course is associated with future relapses [23, 33], disability accumulation [34, 35], and cognitive deficits [36]. Based on a meta-analysis of 23 clinical trials in relapsing-remitting MS, the association between the treatment effect on relapse rate is strongly correlated ($R^2 = 0.81$) with the treatment effect on MRI lesions (i.e., new or active T2 lesions or gadolinium enhancing lesions if monthly scans were acquired) [37]. Other studies have shown that the effect of interferon-beta on MRI lesions mediates 60% of the effect on relapse rate [38] and 57% of the effect on disability progression [39]. Taken together, these findings suggest the potential role of MRI lesions as a surrogate for disability progression and relapse rate in clinical trials, and also that MRI lesion and relapse activity can serve as early indicators of treatment response in regards to disability progression.

*3.1.2. "Black Hole" Formation.* "Black holes" are nonenhancing hypointense lesions on T1-weighted imaging that are correlated with areas of focal chronic axonal damage and loss on histopathology [40]. Therefore, the evolution of active lesions into T1-hypointense lesions represents irreversible tissue damage, and their accumulation is associated with disability accrual ($r_s = 0.46$) [41–43]. In a recent multivariable analysis, worsening of EDSS score over 10 years in 58 patients with RRMS was associated with a combination of baseline

T1-hypointense lesion count and increasing T1 lesion volume ($r = 0.61$, $P < 0.001$) [44]. Taken together, these findings support the role for a decline in "black hole" formation as a potential marker for neuroprotective effects. In this context, it is noteworthy that a placebo-controlled trial of glatiramer acetate with monthly MRI monitoring demonstrated a significant reduction in "black hole" lesion formation [45].

### 3.1.3. T2 Lesion Volume.

*3.1.3. T2 Lesion Volume.* Measuring changes in total T2 lesion volume is another method of evaluating MRI activity. Robust (semi-) automated volumetric analyses rely on the acquisition of MRI scans according to a standardized and carefully quality-monitored MRI protocol. T2 lesion volume has been shown to be reduced in patients with MS receiving teriflunimide compared to placebo [46]. T2 lesion volume is positively correlated with disability measured by EDSS at 2 [47] and 10 years of follow-up [48] and number of relapses after 2 years of follow up [34, 47]. A recent study evaluated MRI correlates of disability in a cohort of 159 patients with relapsing-remitting MS (median EDSS = 4) followed for mean of 26 years from first attack and found that T2 lesion volume was associated with long-term disability and independent of cervical spinal cord atrophy and grey matter atrophy [49]. Similarly, a study of 107 MS patients followed for mean of 20 years from first attack showed that T2 lesion volume correlated with 20-year EDSS ($r_s = 0.48$–$0.67$) and MSFC z-score ($r_s = -0.5$–$0.61$).

*3.1.4. Brain Volume.* Brain volume and atrophy measurements correlate with measures of disability [48, 50, 51] and cognitive function [52, 53]. Considering the previously discussed limitations to using EDSS as a primary outcome measure, the validity of brain atrophy as a surrogate marker for disability progression is of interest in studies on clinical trial design for relapsing-remitting MS. A recent meta-analysis of thirteen randomized clinical trials in relapsing-remitting MS showed that the treatment effect on disability progression (i.e., a 3- or 6-month sustained 1-point increase in EDSS) correlated with the treatment effects on brain atrophy ($R^2 = 0.48$) and active T2 lesions. ($R^2 = 0.61$) [54]. In fact, the authors showed in a multivariable model that 75% of the variance in treatment effect on disability was explained by the combined effect on active T2 lesions and brain atrophy. These findings, if validated in individual patient-based analyses of clinical trial data, support the use of brain atrophy alone or in combination with active T2 lesions as surrogate markers of disability progression in relapsing-remitting MS.

An important practical implication is that trials powered on the outcome of a 50% reduction in MRI lesions [55] and atrophy [56] require ten times fewer subjects than that required for a trial based on a disability endpoint [54, 57]. Recent data has suggested that placebo-controlled trials evaluating brain atrophy (using the SIENA method) would require 32 subjects per arm (80% power) to detect a 50% treatment effect over 2 years [58].

Given that axonal degeneration and loss are now understood as major contributors to disability in MS [59], considerable attention has been given to defining neuroprotective therapeutic strategies that will slow or prevent disability progression. Changes in brain volume in patients with MS reflect the neurodegenerative biology of the disease, and therefore, MRI markers of neuronal damage may represent potential surrogates to neuroprotective therapy response [60]. Moreover, current brain atrophy measurement techniques are suited to multicentered trials, further supporting the potential role of brain volume surrogate markers in phase II trials of neuroprotective agents [21]. Trials of oral fingolimod [61], oral laquinimod [62], and natalizumab [20] reported a favorable effect of treatment on brain volume loss in relapsing-remitting MS patients, compared to placebo. However, some challenges, such as the effects of brain edema, pseudoatrophy from corticosteroids, and ongoing neurodegeneration from injury prior to trial enrollment that challenge establishment of a stable baseline need to be addressed before brain atrophy is considered as a surrogate marker of treatment effect in clinical trials.

*3.1.5. MRI Metrics as Surrogate Outcome Measures.* While MRI variables (e.g., lesion count, lesion volume, and brain volume) are commonly used in clinical trials to assess treatment efficacy and disease progression, they are not generally recognized as validated surrogate outcome measures. For a surrogate endpoint to be valid for a Phase 3 trial, it must exist within the causal pathway in the absence of other mechanisms of action between the exposure and the outcome, and an intervention must exert its effects on the clinical outcome through the surrogate (Figure 1(b)). If the intervention has other mechanisms of action, the surrogate may fail [2].

As reflected in Table 1, clinical trials in relapsing-remitting MS incorporate multiple MRI outcome measures as surrogates of disease activity. However, there is a poor correlation between MRI activity and relapses as the appearance of new MRI lesions often outnumber clinical relapses. This "clinico-radiological paradox" in MS became apparent when MRI was first used in MS and attempts to correlate T2 lesions (a nonspecific marker of focal brain injury) with EDSS revealed a dissociation between the two [63]. Advances in our understanding of normal-appearing brain tissue damage, clinically-silent spinal cord damage and atrophy, grey matter and retinal nerve fiber layer involvement, cognitive impairment, and cortical adaptation in MS have helped explain some of the confounders to the clinical-radiological association. MRI lesion activity measures have not been accepted as validated surrogates because such measures lack the pathologic specificity for the processes that contribute to disability in MS. As illustrated in Figure 1(c), the relationship between a first attack or relapsing disease and the clinical outcomes in MS is complex. MRI measures alone have failed as surrogate markers of disability since other mechanisms for tissue injury exist beyond the inflammatory lesions that are visualized on MRI scans.

*3.2. Optical Coherence Tomography.* Optical coherence tomography (OCT) is an imaging modality which uses near infrared light to measure thickness and volumes of structures in the eye. Of particular interest is the retinal nerve fiber

layer (RNFL), which consists of nonmyelinated axons, and macular volume, which comprises the axons and ganglion cell bodies. By quantifying the RNFL layer thickness and macular volume, OCT noninvasively captures anterior visual pathway axonal loss [64–66].

RNFL thickness has been proposed as a structural biomarker or surrogate measure for regional (optic nerve) and global (whole brain) axonal loss in MS [67]. In adults with MS, RNFL thickness is decreased compared to healthy controls by 5 um to 40 um on average, with greater thinning observed in the eyes of patients who have had optic neuritis [65, 67–69]. Longitudinal studies have also shown RNFL thinning in MS patients over time in the absence of clinical optic neuritis [70, 71]. RNFL thinning in adults with MS who have not had a clinical history or radiographic evidence of optic neuritis suggests that OCT also captures global brain atrophy in adults rather than local effects from optic nerve damage. The ability of OCT measures to inform upon global axonal injury in adult MS is further supported by the relationship between RNFL thickness and macular volumes with brain parenchymal fraction (BPF). In a study of 44 patients with a clinically isolated syndrome or multiple sclerosis, Young et al. found that RNFL thickness and total macular volume (TMV) as measured by OCT were both significantly associated with BPF as measured by MRI ($P = 0.005$ and $P = 0.034$ for RNFL thickness and TMV, resp.) [72]. Zimmerman et al. also found RNFL and TMV to be associated with white matter volume and normalized brain volume ($P < 0.002$ for both); however, contrary to Young et al, they also found RNFL thickness and TMV to be associated with grey matter volume ($P \leq 0.001$ for both). Furthermore, Zimmerman et al. found these associations to be significant regardless of ON history [73].

Despite these data and further studies showing correlations between OCT measures, functional vision scores, quality of life scales, and the EDSS, RNFL thickness has not been validated as a surrogate for disability or disease progression in Phase 3 trials given the complex pathobiology in MS [15]. Pathologic changes in axons occur early in the disease and are prominent during the acute and progressive stages of MS [74]. Similar to the complexities using MRI metrics as surrogate outcome measures, OCT does not capture demyelination, and other pathways exist between the disease (either clinically isolated syndrome or relapsing-remitting MS) and the outcomes (disability, conversion to CDMS, or relapses). Nevertheless, OCT is still a promising tool for future trials, especially studies evaluating the effects of neuroprotective agents [50]. During an international meeting of selected experts ("Imaging outcomes for protection and repair in multiple sclerosis," Amsterdam, The Netherlands, August, 2008), Barkhof et al. concluded that OCT fulfilled the five criteria for outcome measures in such trials: (1) pathologic specificity; (2) reproducibility; (3) sensitivity to change; (4) clinical relevance; and (5) response to treatment [50].

## 4. Pediatric MS Clinical Trials

Currently, disease-modifying therapies have not been approved for pediatric MS as clinical trials have not

been performed in children less than 18 years. However, interferons, glatiramer acetate, natalizumab, and other therapies are prescribed off-label in the pediatric population based on adult trials and pediatric observational studies [75, 76]. While these therapies have been monitored for adverse effects and tolerability in children, additional data are needed, especially with respect to the risk of progressive multifocal leukoencephalopathy and exposure to JC virus. With respect to the newer agents approved for MS, such as fingolimod, teriflunomide, and dimethyl fumarate, long-term effects on neurological development and the maturation of a child's immune, endocrine, and reproductive systems, and neurological development are unknown [36]. Therefore, clinical trials in pediatric MS are needed and are currently underway given the recent EMA and FDA mandate that all newly approved drugs must include a pediatric investigation plan.

The conduct of clinical trials in pediatric MS will be dependent on the use of sensitive clinical outcome measures that capture the most relevant dimensions of function and disability. While the clinical symptoms of MS are similar in children and adults, there are notable differences in the relapse rate, disease burden on MRI, and disability based on the age at diagnosis. Some of these differences include the following.

  (i) Children have a higher number of relapses in the first 2–5 years of disease compared to adults [77, 78].

 (ii) Children generally present with a second attack within 12 months of the first attack, although younger children may have a longer interval [79]. Adults generally take about 2–2.5 years to convert to clinically-definite MS [80, 81].

(iii) Children have an increased T2 lesion burden in the brain at disease onset on MRI scans and more gadolinium enhancing foci [16, 78].

(iv) Children, especially prepubertal, are more likely to initially have large ill-defined T2 lesions that resolve on follow-up scans and then develop focal ovoid lesions typical of adult-onset MS. In contrast, adults have focal ovoid T2 lesions that do not typically resolve [16, 78].

The EDSS may not be as sensitive in pediatric subjects compared to adults with MS due to a lesser degree of physical disability in children with MS. However, cognitive deficits have been recognized in pediatric MS subjects. Cognitive impairment occurs in approximately 30% of children with MS [53, 82]. Patient IQ is also significantly reduced in pediatric-onset MS, and lower IQ scores are associated with younger age at onset of disease [82]. Till et al. found decreased cognitive function to be associated with reduced thalamic and global brain volumes. Since T1 and T2 lesional volumes were not strong predictors of cognitive impairment, the authors concluded that neurodegenerative processes, rather than inflammation and relapses, are responsible for brain atrophy early in pediatric MS [53].

The differences between pediatric and adults MS and selection of outcome measures were considered by an expert panel of physicians, pharmaceutical representatives, and regulatory agencies while discussing the future of pediatric trials. The Steering Committee of the International Pediatric MS Study Group [36] agreed that annual relapse rate or time to next relapse should be used as the primary outcome measure for clinical trials but acknowledges that this may not always be feasible due to time constraints and cost. The Steering Committee of the International Pediatric MS Study Group also recommended several measures of cognitive function, which examine attention, functioning, verbal learning and memory, language, and general intelligence [36]. Sample size calculations for clinical trials of pediatric-onset MS have recently been proposed and may aid in planning phase II and III trials powered on clinical and MRI endpoints [83].

The comparison group for pediatric MS trials has also been debated. The inclusion of a placebo comparison group may be considered unethical since the disease-modifying therapies are generally considered safe and well-tolerated in children. An alternative approach is to use an interferon or glatiramer actetate for the control group when comparing a newer agent or novel treatment [84]. The latter would require a larger sample size, and the conduct of pediatric MS clinical trials has been further challenged by the low prevalence of disease in children.

## 5. Conclusion

Multiple sclerosis is a disabling disease and current treatment is aimed at slowing disease progression. Several outcome measures have been used in recent relapsing-remitting MS clinical trials to quantify clinical disease activity (relapses), MRI disease burden (lesion counts, lesion and brain volumes), and neurodegeneration (OCT). The selection of the outcome measures and the results of these trials should be interpreted with an understanding of the following. (1) MS is a heterogeneous disease and inclusion criteria should be carefully reviewed when results are applied to individual patients. (2) MS evolves over time from a predominantly inflammatory disease to a secondary progressive or neurogenerative process. As such, a clinical outcome measure may have different efficacy in the early stages of the disease compared to the advanced stages. (3) A surrogate may seem to indicate a positive correlation with the intervention, but the relationship could be attributed to other mechanisms. If the measure is not something that is in the causal pathway for MS, it may mislead researchers about the efficacy of the treatment or progression of the disease [1]. (4) The conduct of pediatric MS trials presents additional challenges that are being carefully reviewed and considered in the design of future studies. The combination of multiple endpoints in clinical trials has led to the successful approval of MS therapies for relapsing-remitting MS resulting in decreased disease burden, morbidity, and disability. Future trials, both adult and pediatric, will likely continue to use a combination of outcome measures to determine the efficacy of investigational therapies in slowing disease progression and modifying disability.

## Conflict of Interests

Dr. Waldman has received an honorarium in the past from Teva Pharmaceutical Industries for a speaking engagement. Ms. Lavery and Dr. Verhey declare that there is no conflict of interests regarding the publication of this paper.

## References

[1] T. R. Fleming and J. H. Powers, "Biomarkers and surrogate endpoints in clinical trials," *Statistics in Medicine*, vol. 31, pp. 2973–2984, 2012.

[2] T. R. Fleming and D. L. DeMets, "Surrogate end points in clinical trials: are we being misled?" *Annals of Internal Medicine*, vol. 125, no. 7, pp. 605–613, 1996.

[3] J. F. Kurtzke, "Historical and clinical perspectives of the expanded disability status scale," *Neuroepidemiology*, vol. 31, no. 1, pp. 1–9, 2008.

[4] L. J. Balcer, "Clinical outcome measures for research in multiple sclerosis," *Journal of Neuro-Ophthalmology*, vol. 21, no. 4, pp. 296–301, 2001.

[5] G. C. Ebers, L. Heigenhauser, M. Daumer, C. Lederer, and J. H. Noseworthy, "Disability as an outcome in MS clinical trials," *Neurology*, vol. 71, no. 9, pp. 624–631, 2008.

[6] D. E. Goodkin, D. Cookfair, K. Wende et al., "Inter- and intrarater scoring agreement using grades 1.0 to 3.5 of the Kurtzke Expanded Disability Status Scale (EDSS)," *Neurology*, vol. 42, no. 4, pp. 859–863, 1992.

[7] R. Rudick, J. Antel, C. Confavreux et al., "Recommendations from the national multiple sclerosis society clinical outcomes assessment task force," *Annals of Neurology*, vol. 42, no. 3, pp. 379–382, 1997.

[8] G. R. Cutter, M. L. Baier, R. A. Rudick et al., "Development of a multiple sclerosis functional composite as a clinical trial outcome measure," *Brain*, vol. 122, no. 5, pp. 871–882, 1999.

[9] J. S. Fischer, A. J. Jak, J. E. Knicker, and R. Rudick, "Multiple Sclerosis Functional Composite: Administration and Scoring Manual," 2001.

[10] D. M. Miller, R. A. Rudick, G. Cutter, M. Baier, and J. S. Fischer, "Clinical significance of the multiple sclerosis functional composite: relationship to patient-reported quality of life," *Archives of Neurology*, vol. 57, no. 9, pp. 1319–1324, 2000.

[11] R. A. Rudick, C. H. Polman, J. A. Cohen et al., "Assessing disability progression with the multiple sclerosis functional composite," *Multiple Sclerosis*, vol. 15, no. 8, pp. 984–997, 2009.

[12] R. A. Rudick, G. Cutter, M. Baier et al., "Use of the multiple sclerosis functional composite to predict disability in relapsing MS," *Neurology*, vol. 56, no. 10, pp. 1324–1330, 2001.

[13] D. Ontaneda, N. LaRocca, T. Coetzee, R. Rudick, and N. M. T. Force, "Revisiting the multiple sclerosis functional composite: proceedings from the National Multiple Sclerosis Society (NMSS) Task Force on Clinical Disability Measures," *Multiple Sclerosis*, vol. 18, pp. 1074–1080, 2012.

[14] R. T. Naismith, L. Piccio, J. A. Lyons et al., "Rituximab add-on therapy for breakthrough relapsing multiple sclerosis: a 52-week phase II trial," *Neurology*, vol. 74, no. 23, pp. 1860–1867, 2010.

[15] L. J. Balcer, S. L. Galetta, P. A. Calabresi et al., "Natalizumab reduces visual loss in patients with relapsing multiple sclerosis," *Neurology*, vol. 68, no. 16, pp. 1299–1304, 2007.

[16] E. A. Yeh, T. Chitnis, L. Krupp et al., "Pediatric multiple sclerosis," *Nature Reviews Neurology*, vol. 5, no. 11, pp. 621–631, 2009.

[17] L. J. Balcer, M. L. Baier, V. S. Pelak et al., "New low-contrast vision charts: reliability and test characteristics in patients with multiple sclerosis," *Multiple Sclerosis*, vol. 6, no. 3, pp. 163–171, 2000.

[18] L. J. Balcer, M. L. Baier, J. A. Cohen et al., "Contrast letter acuity as a visual component for the multiple sclerosis functional composite," *Neurology*, vol. 61, no. 10, pp. 1367–1373, 2003.

[19] M. L. Baier, G. R. Cutter, R. A. Rudick et al., "Low-contrast letter acuity testing captures visual dysfunction in patients with multiple sclerosis," *Neurology*, vol. 64, no. 6, pp. 992–995, 2005.

[20] D. H. Miller, D. Soon, K. T. Fernando et al., "MRI outcomes in a placebo-controlled trial of natalizumab in relapsing MS," *Neurology*, vol. 68, no. 17, pp. 1390–1401, 2007.

[21] R. A. Rudick, A. Pace, M. R. S. Rani et al., "Effect of statins on clinical and molecular responses to intramuscular interferon beta-1a," *Neurology*, vol. 72, no. 23, pp. 1989–1993, 2009.

[22] L. Kappos, C. H. Polman, M. S. Freedman et al., "Treatment with interferon beta-1b delays conversion to clinically definite and McDonald MS in patients with clinically isolated syndromes," *Neurology*, vol. 67, no. 7, pp. 1242–1249, 2006.

[23] F. Lublin, "Multiple sclerosis trial designs for the 21st century: building on recent lessons," *Journal of Neurology*, vol. 252, no. 5, pp. v46–v53, 2005.

[24] B. G. Weinshenker, B. Bass, G. P. A. Rice et al., "The natural history of multiple sclerosis: a geographically based study. 2. Predictive value of the early clinical course," *Brain*, vol. 112, no. 6, pp. 1419–1428, 1989.

[25] M. P. Sormani, A. Signori, P. Siri, and N. DeStefano, "Time to first relapse as an endpoint in multiple sclerosis clinical trials," *Multiple Sclerosis*, vol. 19, pp. 466–474, 2013.

[26] S. Inusah, M. P. Sormani, S. S. Cofield et al., "Assessing changes in relapse rates in multiple sclerosis," *Multiple Sclerosis*, vol. 16, no. 12, pp. 1414–1421, 2010.

[27] K. P. Johnson, B. R. Brooks, J. A. Cohen 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," *Neurology*, vol. 57, no. 12, pp. S16–S24, 2001.

[28] G. Comi, M. Filippi, F. Barkhof et al., "Effect of early interferon treatment on conversion to definite multiple sclerosis: a randomised study," *The Lancet*, vol. 357, no. 9268, pp. 1576–1582, 2001.

[29] D. Cadavid, L. J. Wolansky, J. Skurnick et al., "Efficacy of treatment of MS with IFN$\beta$2-1b or glatiramer acetate by monthly brain MRI in the BECOME study," *Neurology*, vol. 72, no. 23, pp. 1976–1983, 2009.

[30] R. J. Temple, *A Regulatory Authority's Opinion About Surrogate Endpoints*, John Wiley & Sons, New York, NY, USA, 1995.

[31] Y. Sadaka, L. H. Verhey, M. M. Shroff et al., "2010 McDonald criteria for diagnosing pediatric multiple sclerosis," *Annals of Neurology*, vol. 72, pp. 211–223, 2012.

[32] L. Bonzano, L. Roccatagliata, G. L. Mancardi, and M. P. Sormani, "Gadolinium-enhancing or active T2 magnetic resonance imaging lesions in multiple sclerosis clinical trials?" *Multiple Sclerosis*, vol. 15, no. 9, pp. 1043–1047, 2009.

[33] L. Kappos, D. Moeri, E. W. Radue et al., "Predictive value of gadolinium-enhanced magnetic resonance imaging for relapse rate and changes in disability or impairment in multiple sclerosis: a meta-analysis," *The Lancet*, vol. 353, no. 9157, pp. 964–969, 1999.

[34] P. A. Brex, O. Ciccarelli, J. I. O'Riordan, M. Sailer, A. J. Thompson, and D. H. Miller, "A longitudinal study of abnormalities on MRI and disability from multiple sclerosis," *The New England Journal of Medicine*, vol. 346, no. 3, pp. 158–164, 2002.

[35] J. I. O'Riordan, M. G. Cain, A. Coles et al., "T1 hypointense lesion load in secondary progressive multiple sclerosis: a comparison of pre versus post contrast loads and of manual versus semi automated threshold techniques for lesion segmentation," *Multiple Sclerosis*, vol. 4, no. 5, pp. 408–412, 1998.

[36] T. Chitnis, M. Tardieu, M. P. Amato et al., "International pediatric Ms study group clinical trials summit: meeting report," *Neurology*, vol. 80, pp. 1161–1168, 2013.

[37] M. P. Sormani, L. Bonzano, L. Roccatagliata, G. R. Cutter, G. L. Mancardi, and P. Bruzzi, "Magnetic resonance imaging as a potential surrogate for relapses in multiple sclerosis: a meta-analytic approach," *Annals of Neurology*, vol. 65, no. 3, pp. 268–275, 2009.

[38] M. P. Sormani, B. Stubinski, P. Cornelisse, S. Rocak, D. Li, and N. De Stefano, "Magnetic resonance active lesions as individual-level surrogate for relapses in multiple sclerosis," *Multiple Sclerosis*, vol. 17, no. 5, pp. 541–549, 2011.

[39] M. P. Sormani, P. Bruzzi, K. Beckmann et al., "MRI metrics as surrogate endpoints for EDSS progression in SPMS patients treated with IFN $\beta$-1b," *Neurology*, vol. 60, no. 9, pp. 1462–1466, 2003.

[40] M. A. A. Van Walderveen, L. Truyen, B. W. Van Oosten et al., "Development of hypointense lesions on T1-weighted spin-echo magnetic resonance images in multiple sclerosis: relation to inflammatory activity," *Archives of Neurology*, vol. 56, no. 3, pp. 345–351, 1999.

[41] J. van Waesberghe, W. Kamphorst, C. J. De Groot, M. A. van Walderveen, J. A. Castelijns et al., "Axonal loss in multiple sclerosis lesions: magnetic resonance imaging insights into substrates of disability," *Annals of Neurology*, vol. 46, pp. 747–754, 1999.

[42] L. Truyen, J. H. T. M. Van Waesberghe, M. A. A. Van Walderveen et al., "Accumulation of hypointense lesions ("black holes") on T1 spin-echo MRI correlates with disease progression in multiple sclerosis," *Neurology*, vol. 47, no. 6, pp. 1469–1476, 1996.

[43] M. A. A. Van Walderveen, F. Barkhof, O. R. Hommes et al., "Correlating MRI and clinical disease activity in multiple sclerosis: relevance of hypointense lesions on short-TR/short-TE(T1-weighted) spin- echo images," *Neurology*, vol. 45, no. 9, pp. 1684–1690, 1995.

[44] A. Giorgio, M. L. Bartolozzi, F. Rossi, M. Battaglini et al., "Relevance of hypointense brain MRI lesions for long-term worsening of clinical disability in relapsing multiple sclerosis," *Multiple Sclerosis*, vol. 20, no. 2, pp. 214–219, 2014.

[45] M. Filippi, M. Rovaris, M. A. Rocca, M. P. Sormani, J. S. Wolinsky, and G. Comi, "Glatiramer acetate reduces the proportion of new MS lesions evolving into "black holes"," *Neurology*, vol. 57, no. 4, pp. 731–733, 2001.

[46] P. O'Connor, J. S. Wolinsky, C. Confavreux et al., "Randomized trial of oral teriflunomide for relapsing multiple sclerosis," *The New England Journal of Medicine*, vol. 365, no. 14, pp. 1293–1303, 2011.

[47] A. M. Pascual, M. C. Martínez-Bisbal, I. Boscá et al., "Axonal loss is progressive and partly dissociated from lesion load in early multiple sclerosis," *Neurology*, vol. 69, no. 1, pp. 63–67, 2007.

[48] V. Popescu, F. Agosta, H. E. Hulst, I. C. Sluimer, D. L. Knol, and al. e, "Brain atrophy and lesion load predict long term disability in multiple sclerosis," *Journal of Neurology, Neurosurgery, and Psychiatry*, vol. 84, pp. 1082–1091, 2013.

[49] H. Kearney, M. Rocca, P. Valsasina et al., "Magnetic resonance imaging correlates of physical disability in relapse onset multiple sclerosis of long disease duration," *Multiple Sclerosis*, vol. 20, no. 1, pp. 72–80, 2014.

[50] F. Barkhof, P. A. Calabresi, D. H. Miller, and S. C. Reingold, "Imaging outcomes for neuroprotection and repair in multiple sclerosis trials," *Nature Reviews Neurology*, vol. 5, no. 5, pp. 256–266, 2009.

[51] R. A. Rudick, J.-C. Lee, K. Nakamura, and E. Fisher, "Gray matter atrophy correlates with MS disability progression measured with MSFC but not EDSS," *Journal of the Neurological Sciences*, vol. 282, no. 1-2, pp. 106–111, 2009.

[52] O. Yaldizli, I. K. Penner, K. Frontzek, Y. Naegelin et al., "The relationship between total and regional corpus callosum atrophy, cognitive impairment and fatigue in multiple sclerosis patients," *Multiple Sclerosis*, vol. 20, no. 3, pp. 356–364, 2014.

[53] C. Till, R. Ghassemi, B. Aubert-Broche et al., "MRI correlates of cognitive impairment in childhood-onset multiple sclerosis," *Neuropsychology*, vol. 25, no. 3, pp. 319–332, 2011.

[54] M. Sormani, D. L. Arnold, and N. De Stefano, "Treatment effect on brain atrophy correlates with treatment effect on disability in multiple sclerosis," *Annals of Neurology*, vol. 75, no. 1, pp. 43–49, 2014.

[55] M. P. Sormani, D. H. Miller, G. Comi et al., "Clinical trials of multiple sclerosis monitored with enhanced MRI: new sample size calculations based on large data sets," *Journal of Neurology Neurosurgery and Psychiatry*, vol. 70, no. 4, pp. 494–499, 2001.

[56] N. De Stefano, A. Giorgio, M. Battaglini et al., "Assessing brain atrophy rates in a large population of untreated multiple sclerosis subtypes," *Neurology*, vol. 74, no. 23, pp. 1868–1876, 2010.

[57] C. H. Polman, P. W. O'Connor, E. Havrdova et al., "A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis," *The New England Journal of Medicine*, vol. 354, no. 9, pp. 899–910, 2006.

[58] D. R. Altmann, B. Jasperse, F. Barkhof et al., "Sample sizes for brain atrophy outcomes in trials for secondary progressive multiple sclerosis," *Neurology*, vol. 72, no. 7, pp. 595–601, 2009.

[59] C. Bjartmar, G. Kidd, S. Mork, R. Rudick, and B. D. Trapp, "Neurological disability correlates with spinal cord axonal loss and reduced N-acetyl aspartate in chronic multiple sclerosis patients," *Annals of Neurology*, vol. 48, pp. 893–901, 2000.

[60] R. A. Bermel and R. Bakshi, "The measurement and clinical relevance of brain atrophy in multiple sclerosis," *The Lancet Neurology*, vol. 5, no. 2, pp. 158–170, 2006.

[61] L. Kappos, E.-W. Radue, P. O'Connor et al., "A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis," *The New England Journal of Medicine*, vol. 362, no. 5, pp. 387–401, 2010.

[62] M. Filippi, M. A. Rocca, E. Pagani et al., "Placebo-controlled trial of oral laquinimod in multiple sclerosis: MRI evidence of an effect on brain tissue damage," *Journal of Neurology, Neurosurgery & Psychiatry*, 2013.

[63] F. Barkhof, "The clinico-radiological paradox in multiple sclerosis revisited," *Current Opinion in Neurology*, vol. 15, no. 3, pp. 239–245, 2002.

[64] V. Parisi, G. Manni, M. Spadaro et al., "Correlation between morphological and functional retinal impairment in multiple sclerosis patients," *Investigative Ophthalmology and Visual Science*, vol. 40, no. 11, pp. 2520–2527, 1999.

[65] S. A. Trip, P. G. Schlottmann, S. J. Jones et al., "Retinal nerve fiber layer axonal loss and visual dysfunction in optic neuritis," *Annals of Neurology*, vol. 58, no. 3, pp. 383–391, 2005.

[66] C. V. Warner, S. B. Syc, A. M. Stankiewicz et al., "The impact of utilizing different optical coherence tomography devices for clinical purposes and in multiple sclerosis trials," *PLoS ONE*, vol. 6, no. 8, Article ID e22947, 2011.

[67] J. B. Fisher, D. A. Jacobs, C. E. Markowitz et al., "Relation of visual function to retinal nerve fiber layer thickness in multiple sclerosis," *Ophthalmology*, vol. 113, no. 2, pp. 324–332, 2006.

[68] B. M. Burkholder, B. Osborne, M. J. Loguidice et al., "Macular volume determined by optical coherence tomography as a measure of neuronal loss in multiple sclerosis," *Archives of Neurology*, vol. 66, no. 11, pp. 1366–1372, 2009.

[69] D. Fernandes, A. Raza, R. Nogueira, D. Wand et al., "Evaluation of inner retinal layers in patients with multiple sclerosis or neuromyelitis optica using optical coherence tomography," *Ophthalmology*, vol. 120, no. 2, pp. 387–394, 2013.

[70] R. Herrero, E. Garcia-Martin, C. Almarcegui et al., "Progressive degeneration of the retinal nerve fiber layer in patients with multiple sclerosis," *Investigative Ophthalmology & Visual Science*, vol. 53, pp. 8344–8349, 2012.

[71] L. S. Talman, E. R. Bisker, D. J. Sackel et al., "Longitudinal study of vision and retinal nerve fiber layer thickness in multiple sclerosis," *Annals of Neurology*, vol. 67, no. 6, pp. 749–760, 2010.

[72] K. L. Young, A. U. Brandt, A. Petzold et al., "Loss of retinal nerve fibre layer axons indicates white but not grey matter damage in early multiple sclerosis," *European Journal of Neurology*, vol. 20, pp. 803–811, 2013.

[73] H. Zimmerman, A. Freing, F. Kaufhold, G. Gaded et al., "Optic neuritis interferes with optical coherence tomography and magnetic resonance imaging correlations," *Multiple Sclerosis*, vol. 19, no. 4, pp. 443–450, 2013.

[74] E. M. Frohman, J. G. Fujimoto, T. C. Frohman, P. A. Calabresi, G. Cutter, and L. J. Balcer, "Optical coherence tomography: a window into the mechanisms of multiple sclerosis," *Nature Clinical Practice Neurology*, vol. 4, no. 12, pp. 664–675, 2008.

[75] T. Chitnis, S. Tenembaum, B. Banwell et al., "Consensus statement: evaluation of new and existing therapeutics for pediatric multiple sclerosis," *Multiple Sclerosis*, vol. 18, no. 1, pp. 116–127, 2012.

[76] A. Ghezzi, "Therapeutic strategies in childhood multiple sclerosis," *Therapeutic Advances in Neurological Disorders*, vol. 3, no. 4, pp. 217–228, 2010.

[77] M. P. Gorman, B. C. Healy, M. Polgar-Turcsanyi, and T. Chitnis, "Increased relapse rate in pediatric-onset compared with adult-onset multiple sclerosis," *Archives of Neurology*, vol. 66, no. 1, pp. 54–59, 2009.

[78] E. Waubant, D. Chabas, D. T. Okuda et al., "Difference in disease burden and activity in pediatric patients on brain magnetic resonance imaging at time of multiple sclerosis onset vs adults," *Archives of Neurology*, vol. 66, no. 8, pp. 967–971, 2009.

[79] R. C. Dale, F. Brilot, and B. Banwell, "Pediatric central nervous system inflammatory demyelination: acute disseminated encephalomyelitis, clinically isolated syndromes, neuromyelitis optica, and multiple sclerosis," *Current Opinion in Neurology*, vol. 22, no. 3, pp. 233–240, 2009.

[80] M. Gómez-Moreno, M. Díaz-Sánchez, and A. Ramos-González, "Application of the 2010 McDonald criteria for the diagnosis of multiple sclerosis in a Spanish cohort of patients with clinically isolated syndromes," *Multiple Sclerosis*, vol. 18, no. 1, pp. 39–44, 2012.

[81] R. P. Kinkel, "IM interferon $\beta$-1a delays definite multiple sclerosis 5 years after a first demyelinating event," *Neurology*, vol. 66, no. 5, pp. 678–684, 2006.

[82] M. P. Amato, B. Goretti, A. Ghezzi et al., "Cognitive and psychosocial features of childhood and juvenile MS," *Neurology*, vol. 70, no. 20, pp. 1891–1897, 2008.

[83] L. H. Verhey, A. Signori, D. L. Arnold et al., "Clinical and MRI activity as determinants of sample size for pediatric multiple sclerosis trials," *Neurology*, vol. 81, pp. 1215–1221, 2013.

[84] S. Tenembaum, "Ethical challenges in paediatric clinical trials in multiple sclerosis," *Therapeutic Advances in Neurological Disorders*, vol. 5, pp. 139–146, 2012.