

Section I

Chapter

1

Introduction

Aspects of multiple sclerosis that relate to experimental therapeutics

Richard A. Rudick and Jeffrey A. Cohen

The purpose of this chapter is to discuss key features of multiple sclerosis (MS) that relate to clinical trial design or treatment. The emphasis will be on aspects of the disease that create challenges for developing effective therapies, and for using them in practice. These include the subclinical nature of early-stage MS, phenotypic and disease heterogeneity, and complexities related to measuring disease severity. Many of these topics are covered in greater detail throughout the book. This chapter will end with a brief discussion about current controversies in the MS experimental therapeutics field.

Disease features relevant to clinical trials

MS pathology is largely subclinical in early MS

Relapsing remitting MS (RRMS) patients have periodic relapses occurring at variable rates, but generally less than one per year. Serial MRI demonstrates many more new lesions than clinical relapses, with most studies demonstrating a rate of new MRI lesions about 10–20-fold higher than clinical relapses.^{1,2} In some patients, MRI shows active disease for years with no clinical relapses, indicating that MS disease activity may be entirely subclinical in some patients during the early stage of the disease. New inflammatory lesions in white matter begin with gadolinium (Gd) enhancement, marking sites of inflammatory lesions (Chapter 9).^{3–5} Approximately 50% of untreated RRMS patients have Gd-enhancing lesions on a cranial MRI scan obtained when the disease is inactive clinically.^{6,7} Even the number of Gd-positive lesions drastically underestimates disease activity, however. First, gray matter pathology is present in MS patients, even early in the disease (Chapters 2, 11, and 13), and Gd-enhancement rarely occurs in gray matter lesions. Second, disease in normal appearing white matter is well recognized, and correlates with progressive atrophy. Therefore, Gd-enhancing lesions, themselves mostly asymptomatic, are just the “tip of the iceberg” with respect to MS pathology.

There are several implications of this for clinical trials. First, relapse counts are meaningful clinical outcomes in RRMS, but inherently insensitive, and it is difficult to measure clinical disability in RRMS patient groups, because RRMS patients don’t generally get disabled during the time-frame of a clinical trial. Second, insensitivity of clinical outcomes in RRMS drives up

sample sizes, which become prohibitive in active arm designs (Chapter 21). Third, the subclinical nature of disease activity in RRMS forms the basis for screening putative therapies using MRI outcomes, including MRI parameters as secondary outcomes, and potentially using MRI as a primary outcome measure in RRMS trials. In that regard, Sormani and colleagues conducted a pooled analysis of 23 clinical trials that included MRI lesion measurements, to test the effect of interventions on lesions and relapse rate.⁸ The effect of the intervention on MRI lesions was strongly correlated with the effect of the intervention on relapses, accounting for over 80% of the total variance. This indicates that formation of new lesions in RRMS is clinically relevant, and supports the argument that MRI disease activity could be used as a primary outcome for RRMS trials. While many have advocated for this, use of MRI as a primary outcome measure has not achieved regulatory agency acceptance (Chapter 18).

Progressive destructive pathology starts early in the disease

The rationale for early intervention in MS is the presence of widespread tissue damage at the earliest stages of the disease.^{9–11} Once RRMS is established, residual clinical disability is usually minimal or absent early in the disease, yet there is ongoing tissue damage, as evidenced by accumulation of T2-bright MRI lesions,¹ T1-hypointense lesions,¹² and brain atrophy (Chapter 11).^{13–16} Gray matter lesions are frequent in MS autopsy material.¹⁷ Although these lesions are not visualized with standard MRI methods, gray matter atrophy has been documented early in the disease.^{18,19} It is widely believed that the ongoing destructive pathology sets the stage for later conversion to secondary progressive MS (SPMS), in which disability accumulates. According to this model, tissue destruction proceeds without frank neurological disability progression until a threshold is surpassed. Beyond that stage, progressive neurological disability ensues. The threshold hypothesis was supported by studies correlating retinal nerve fiber layer (RNFL) thickness with visual acuity.²⁰ Visual acuity was maintained until RNFL thickness declined to about 75 microns, and then fell off rapidly with further loss of RNFL thickness. The implication of this for

Multiple Sclerosis Therapeutics, Fourth Edition, ed. Jeffrey A. Cohen and Richard A. Rudick. Published by Cambridge University Press.
© Cambridge University Press 2011.

Section I: Introduction

clinical trials is that disease-modifying drug therapy should be viewed as providing secondary neuroprotection, i.e. treatment prevents neurodegeneration by inhibiting inflammation, and thereby decreases the amount of brain injury and delays or prevents SPMS.

Disease manifestations are heterogeneous

Another factor complicating MS clinical trials is disease heterogeneity, which is one of the hallmarks of MS. Patients manifest varying patterns of clinical features, variable clinical course, and variable disease severity. This creates hurdles for clinical trials, as heterogeneity complicates outcomes assessment, and increases required sample sizes. The myriad clinical manifestations include neuropsychological impairment (itself multifaceted), visual loss, eye movement abnormalities, weakness, spasticity, incoordination, imbalance, sensory loss, paresthesias, gait impairment, bowel and bladder dysfunction, sexual dysfunction, fatigue, and paroxysmal phenomena. Individuals manifest these features in varying combinations, and the symptoms change over time. Even within multiply affected families, there is striking clinical heterogeneity between affected family members. Disease heterogeneity is poorly understood in MS, as genome-wide association studies have mostly focused on disease susceptibility genes, rather than disease modifying genes (Chapter 4).

Managing the wide variety of MS symptoms is crucially important for patient well-being, but is increasingly challenging with increasing complexity and emphasis on disease modifying drug therapy. Heterogeneity in clinical manifestations also presents significant challenges for the design of clinical trials. Subjects in separate trials and treatment arms within a given trial exhibit variable admixtures of clinical manifestations that are not necessarily evenly matched between study groups. Multidimensional clinical outcome measures are needed to capture the range of ways in which MS affects patients (Chapter 6). The traditional clinical outcome measure – Expanded Disability Status Scale (EDSS) – is heavily weighted to motor impairment, particularly gait dysfunction. Common symptoms such as cognitive dysfunction, sphincter disturbances, pain, and fatigue have significant effects on functional status and quality of life (QOL), but may not correlate well with measures of physical impairment and disability. This forms a strong rationale for including patient-reported outcomes in clinical trials as a measure of the impact of intervention on disease-related symptoms (Chapter 8). It is possible that therapies may have different effects on specific disease manifestations, i.e. benefit for some with no effect or deleterious effects on others. This is an under-explored area.

Evolution of the MS disease process – the “MS categories”

Because the clinical course of MS evolves over decades, there has been interest in subcategorizing MS into discrete groupings.

The current classification system was based on clinical phenomenology of the clinical disease course,²¹ not on the underlying biological mechanisms. According to this classification, MS begins with a clinically isolated syndrome (CIS), defined as an initial clinical episode with features typical for inflammatory demyelination (e.g. optic neuritis, partial transverse myelitis). With additional clinical episodes, CIS evolves to clinically definite RRMS. Even in the absence of a second relapse, a patient meets criteria for clinically definite MS when new MRI lesions are observed during follow-up.²² RRMS then evolves to SPMS in many but not all patients. About 15% of patients have primary progressive MS (PPMS), meaning that progressive disability ensues without prior relapses. In RRMS patients, periodic relapses occur at irregular and unpredictable intervals, averaging approximately one per year, but declining with disease duration. Episodic attacks of neurological dysfunction are followed by partial or complete recovery, separated by clinically stable intervals. Relapses become less conspicuous over the years, and over 60% of RRMS patients transition to SPMS. During this stage physical and cognitive disability gradually worsens, and disease worsening is refractory to known treatment.

RRMS and SPMS present different challenges in study design. In RRMS, relapses are infrequent, occur at irregular intervals, and pose significant measurement challenges, and disability progression tends to be minimal during the course of a clinical trial. There seems to be some “drift” in MS severity in the direction of more benign disease. This may be driven by increased awareness of MS and widespread use of MRI scanning for patients with non-specific symptoms such as fatigue, paresthesias, or headache.²³ The SPMS stage of the disease is also difficult to study, but for different reasons. Deterioration occurs slowly over the course of years, and there is significant within-patient and between-patient variability. Further, while trials tend to restrict patients by disease category, transition from RRMS to SPMS does not occur at a precise point in time. Clinical relapses become less distinct, recovery becomes less complete, and gradual worsening in the absence of relapses eventually becomes apparent. Transition to the SPMS stage, which commonly occurs during the fourth and fifth decade of life, can be estimated only in retrospect, once it is clear that the patient has gradually worsened in the absence of acute relapses. Because of the indistinct boundary between RR and SPMS, many patients could be entered into either a RRMS or a SPMS clinical trial, depending on how the clinician chooses to classify the individual patient.

A consensus has emerged that PPMS (Chapter 52) should be considered separately for clinical trials. This is based on the uncertainty about the etiological relationship between PPMS and SPMS. Prototypical PPMS patients have symptom onset at a later age, typically between ages 40 and 60, and the female preponderance seen with relapsing forms of MS is not evident. These patients commonly present with insidiously progressive spastic weakness, imbalance, and sphincter dysfunction; diffuse and less nodular T2-hyperintense lesions on cranial MRI; few if any Gd-enhancing lesions; and less indication of

inflammation in cerebrospinal fluid (CSF).²⁴ PPMS may be less dependent on inflammation, and neurodegenerative mechanisms may underlie the disease. Some PPMS patients have clinical, MRI, and CSF findings similar to SPMS. These patients may be similar to SPMS, but without clinically distinct relapses during the early disease stage. This is probably also true of another clinical category – progressive relapsing MS (PRMS) – in which there is gradual neurological progression from onset but with subsequent superimposed relapses. Thus, studies in PPMS are problematic for two reasons. These cases are relatively uncommon, and clinical trial groups contain admixtures of disease etiologies. It is unknown whether “SPMS-like PPMS” and “pure PPMS” patients have similar pathogenic mechanisms driving disease progression, or whether they would respond similarly to treatment intervention.

Common practice has been to select relatively homogeneous patient groups for inclusion in clinical trials by entering patients with a specified disease category, and creating disability limits based on the EDSS.²⁵ As a result of widespread acceptance of the disease categories, separate trials have been conducted for patients with CIS, RRMS, SPMS, and PPMS. This strategy aims to reduce between-patient variability and to increase the power to show therapeutic effects with a given sample size. However, there are some drawbacks. Narrow entry criteria impede recruitment; it may not be clear whether the results of a trial enrolling a highly selected cohort of patients can be extrapolated to other groups of MS patients; and the distinction between clinical disease categories is imprecise and based on clinical features that are disconnected from underlying disease mechanisms. Conversely, different clinical trials that nominally studied the same patient population almost certainly contain different mixes of patients. This point is well illustrated by the European and North American trials of interferon beta-1b (IFNβ-1b) in SPMS. These two trials used similar entry criteria, but enrolled distinct patient populations that yielded different results with the same therapeutic agent.²⁶ The problem of classifying patients is most problematic at the interface between RRMS and SPMS, as discussed above. Biological markers for the different MS categories would be valuable, but are not currently available.

Disease severity can not be accurately predicted in individuals or groups

Because of the highly variable future course for newly diagnosed MS patients, there is a compelling need for prognostic markers for treatment decision-making at the individual patient level. Prognostic markers would not only serve the need for better clinical decision-making, but also would help with informative enrollment into clinical trials. Data from the pre-therapeutic era suggested that 50% of MS patients were unable to carry out household and employment responsibilities 10 years after disease onset, 50% required an assistive device to walk after 15–20 years, and 50% were unable to walk at all after

25 years.²⁷ About 10% of patients have an unusually severe disease course, deteriorating to severe disability in only a few years, while 10%–20% exhibit mild disease with minimal disability decades after symptom onset. Distinguishing these severe and mild cases early after symptom onset has proved difficult.

Selective enrollment has been attempted in clinical trials. The approach has been to enroll patients at risk for disease activity, excluding patients not likely to change during the trial. In groups of patients, milder disease has been associated with sensory symptoms or optic neuritis at onset, good recovery from relapses and infrequent relapses early in the disease course.^{13–15} Conversely, symptom onset at an older age, progressive disease from onset, and poor relapse recovery mark a relatively worse prognosis. Clinical features have not been useful for informative enrollment, however. The presence of multiple white matter lesions at the time of first MS symptom has proven very useful, as it is associated with much higher risk of disease activity in the next 5 years.²⁸ Also, the amount of T2 lesion accrual during the initial 5 years after onset is a modest predictor of EDSS 20 years later.²⁹ Despite this, T2 lesion load has not been used for informative enrollment strategies in clinical trials.

Most trials employ relapses or progression during a specified time period prior to the trial, or Gd-enhancing lesions on screening MRIs to identify patients with increased likelihood of disease activity during the trial. This is supported by a study showing that relapse rate prior to the trial and disease duration were the best predictors of on-study relapse rate.³⁰ In that study, disease course and Gd-enhancement status did not provide additional information. That study used a pooled data set from natural history studies and the placebo groups of randomized clinical trials, with a substantially larger sample size compared with previous analyses. A second study examined factors that predicted on-study Gd-enhancement, a common efficacy end-point in Phase 2 studies.³¹ A combination of younger age at onset, shorter disease duration, recent relapses, and T2 lesion volume predicted Gd-enhancement. In other studies, the presence of Gd-enhancing lesions at baseline predicted frequency of clinical relapses, as well as increased T2 lesion volume and brain atrophy progression over the subsequent two years.^{13,32} However, all of the identified predictors, alone or in combination, are only modestly predictive of disease activity during a trial. The advantages of informative enrollment need to be balanced against the difficulty of finding eligible patients, and the problem of generalizing results when the entry criteria are restrictive.

Heterogeneity in pathological mechanisms

Studies of a large number of biopsy and autopsy specimens suggested that the mechanisms leading to tissue damage differ from patient to patient.^{33–36} Four distinct patterns of pathology were proposed. Analogous to experimental autoimmune encephalomyelitis, in patterns I and II the myelin sheath appears to be the target of the destructive process, mediated by macrophages in pattern I and antibody and complement

Section I: Introduction

deposition in pattern II. Pattern III is characterized by an ill-defined lesion border with early loss of adaxonal myelin-associated glycoprotein. This pattern is similar to that seen in some viral encephalitides and in cerebral ischemia. In pattern IV, there is a sparse inflammatory reaction, with prominent non-apoptotic degeneration of oligodendrocytes in the periplaque white matter. At present, pathologically distinct MS subgroups cannot be defined on the basis of biomarkers, or functional assays. However, most now recognize neuromyelitis optica (NMO) as a distinct disorder (Chapter 53). It has long been known that NMO differs from typical MS clinically, by imaging features, pathology,³⁷ and response to MS disease-modifying drugs. But the watershed event was the observation that NMO is associated with antibodies to aquaporin-4, an astrocyte water channel.^{38,39} Presumably, better understanding of MS pathological heterogeneity will lead to more rationale approaches to personalized use of disease-modifying drugs.

Complexities related to measurement tools that impact clinical trials

Clinical measures: relapses, physical function, neuropsychological performance (Chapters 6–8)

The annualized relapse rate or the number of relapses are the most common primary outcome measure for RRMS clinical trials. Relapse frequency was the primary outcome measure in pivotal trials of two of the three IFN β products,^{40,41} the glatiramer acetate trial,⁴² and the natalizumab trials.^{43,44} These studies led to world-wide approval by regulatory agencies, and marketing of the products. Relapses are considered clinically relevant by regulatory agencies, because they are defined by new neurological symptoms and signs and are therefore assumed to have clinical impact. The relationship between relapse number and future disability is weak, however.⁴⁵ Relapses may be subjective, and influenced by bias, over- or under-reporting, and treatment unmasking. There are no accepted methods to quantify relapse severity or recovery from relapse. Lastly, the amount of relapse rate reduction considered “clinically important” has never been defined. The rate of relapse in MS clinical trial populations has fallen over time. This indicates that more recent trials have enrolled patients with less active disease, lowering the power of recent trials to show treatment arm differences, and making comparison across trials completely impractical.

The EDSS is an ordinal scale ranging from 0 to 10 that classifies disability severity according to 19 steps.²⁵ A score of 0 means a normal neurological examination; a score of 3.5 is computed when there is moderate disability in more than one functional system (e.g. visual, motor, cerebellar, sensory, bowel, bladder, etc.), but the patient is able to walk an unlimited distance without assistance. A score between 4.0 and 6.0 indicates limited distance walking. Level 6.0 indicates the need for unilateral assistance to walk, 6.5 bilateral assistance, and ≥ 7.0 measures severity in non-ambulatory patients. There is debate whether the EDSS measures disability accurately at the low end, because it

has been very difficult to standardize the scoring for the functional system scales and small changes within the functional systems have unclear clinical relevance; and the middle and high ranges are insensitive to change, and so lower the power of clinical trials. Despite criticism, the EDSS has been the standard measure of neurologic disability in nearly all MS clinical trials.

Since the mid-1990s, the EDSS has been used to determine “disability progression,” by identifying patients with confirmed worsening from the baseline score. The proportion of patients in different treatment arms are compared directly, or using survival curves. The most common definition of “disability progression” in RRMS trials is worsening from baseline by at least 1.0 EDSS point, confirmed at the next three-month study visit. A minority of trials have required six-month confirmation. The EDSS may revert to baseline more commonly if the three-month definition is used,⁴⁶ probably because of residual effects of relapses still present at three-months. The relevance of confirmed EDSS worsening in the early stages of MS remains controversial. One study showed a strong observed correlation between six-month confirmed EDSS worsening and clinical outcome eight years later.⁴⁷ There are no similar studies using three-month confirmation. A pooled analysis of multiple clinical trials demonstrated a strong association between treatment effect on relapse rate, and treatment effect on confirmed EDSS worsening, suggesting these two measures are inter-related in RRMS patients.⁴⁸ Despite continued criticism of EDSS as a clinical outcome measure, confirmed EDSS worsening has been accepted by regulatory agencies as a primary “disability progression” end-point for RRMS trials.

Because of perceived limitations of the EDSS, a National MS Society task force recommended the MS Functional Composite (MSFC), a three-part composite consisting of timed measures of ambulation, upper extremity function, and cognition.⁴⁹ The MSFC has been extensively tested and validated but has yet to achieve its intended purpose – to replace the EDSS as a primary clinical measure of MS-related disability. A substantial part of the problem lies in difficulty interpreting the clinical relevance of the results. As originally recommended, the three MSFC measures are transformed to a single Z score, defined as the average of the Z scores from the ambulation, upper extremity, and cognitive tests. Not only is the clinically relevant Z-score difference not defined, but also the choice of reference population influences the weighting of the different components within the MSFC,⁵⁰ so the optimal population used to normalize clinical trial test scores is debatable. Recently, a group analyzed MSFC data collected during the natalizumab placebo-controlled trial, and proposed using the MSFC to identify a disability progression event, analogous to how the EDSS is used.⁵¹ Disability progression defined using MSFC scores correlated with traditional measures of disease activity and progression, and demonstrated treatment effects similar to EDSS. It is hoped that the addition of a sensitive visual function measure (e.g. low contrast visual acuity), and perhaps substituting a cognitive measure with less learning effect than PASAT, will improve

MSFC performance characteristics and may promote use of the MSFC as a more sensitive primary outcome measure.

Neuropsychological impairment, particularly changes in processing speed, complex attention, and verbal learning, has been noted in approximately 50% of MS cases in population-based studies, and has been associated with significant vocational and social disability.^{52,53} The effects of treatment on neuropsychological test performance have been reported, although the popularity of neuropsychological testing in MS clinical trials has declined because of research subject burden, and cost considerations. Only six published randomized MS clinical trials included measures of neuropsychological outcome.⁵⁴ Results were mixed. Neuropsychological testing is obviously critical for studies specifically targeting neurocognitive deficits, but these tests have not achieved widespread use in clinical trials. Efforts are under way to develop and validate more brief neuropsychological test batteries that might be more practical for MS trials.⁵⁵

Patient-reported quality of life measures (Chapter 8)⁵⁶

Generic health related (HR)-QOL measures include the Symptom Impact Profile and the Medical Outcomes Study 36-Item Short-Form Survey (SF-36). Hybrid measures are the MS Quality of Life Index and MSQOL-54; MS-specific instruments include the Functional Assessment of MS and MS Impact Scale-29. No consensus exists concerning the optimal patient self-report HR-QOL instrument for MS clinical trials. At least eight clinical trials have reported the effects of IFN β or glatiramer acetate on HR-QOL in MS. The AFFIRM study of natalizumab found a strong association between the physical component score of the SF-36 and the EDSS score, relapse rate, and treatment effects.⁵⁷ Patient-reported HR-QOL measures are appealing in that they capture the overall burden of MS, but they are insensitive because clinical changes may occur while HR-QOL remains the same. In addition, many HR-QOL measures are non-specific, affected by non-disease factors, and are therefore considered most appropriate as secondary outcome measures.

Conventional MRI measures (Chapters 9,11)

All contemporary MS trials include lesion and brain atrophy measures. Gd-enhancing lesions, T2-bright lesions, and lesions that appear dark on T1-weighted scans – the so-called “black holes” – comprise the standard lesion assessment, although image acquisition parameters and the lesion analysis methods have not been standardized. Consequently, lesion “numbers” can not be compared directly across studies. Lesions that enhance after intravenous Gd infusion indicate blood–brain barrier disruption and inflammatory activity. Enhancement lasts 1–4 weeks, so frequent MRI scans are required to capture all Gd-enhancing lesions. During and following enhancement, lesions appear bright on T2-weighted scans, and once formed, T2 hyperintense lesions persist indefinitely. The typical clinical trial includes counts of both Gd-enhancing

lesions and new or newly enlarging T2 lesions. Both are considered measures of new inflammatory activity. All currently approved MS disease-modifying therapies have been shown to reduce enhancing lesions.

The volume of T2 lesions is an estimate of overall MS disease burden. Reductions in the accumulation of T2 lesion volume have been reported in active treatment arms compared to placebo for most MS trials. The significance of this is not clear because the rate of accumulation of T2 lesions correlates weakly with disability progression over the short term.⁵⁸ This may be due to non-specificity of T2 lesions – only about half of all T2 bright lesions are associated with demyelination demonstrated pathologically.⁵⁹ However, T2 lesion volume correlates modestly with future brain atrophy,⁶⁰ and accumulation of T2 volume in the five years following MS onset is predictive of the clinical status 20 years later.⁶¹ Also, the T2 lesion volume is one of the best measures for confirming that clinical trial treatment arms are well matched at baseline.

Lesions appearing dark on T1-weighted scans not related to current or recent Gd-enhancement (“black holes”) correspond to regions with axonal loss.⁶² Black hole volume correlates strongly with T2 volume, however, and has not been very useful as a clinical trial outcome measure. A newer use of black hole data is to determine the proportion of Gd-enhancing lesions that develop into chronic black holes. It has been proposed that this metric may be a useful indicator of neuroprotection.⁶³

Brain atrophy is used as a marker of severe tissue destruction. The techniques vary for quantifying brain atrophy, but generally break into two categories. One is to measure normalized brain volume at two points in time, and subtract the two measures;¹⁴ the other is a more direct measure of brain volume change, in which the two MRI studies are co-registered, and the software measures the changes in brain edges from the MRI pairs.⁶⁴ Both of these methods have been applied in MS clinical trials to estimate whole brain atrophy. Atrophy measures have some advantages over lesion measurements. Most significantly, brain atrophy reflects the net effect of the CNS pathology in MS. Furthermore, brain atrophy correlates more strongly with disability than do lesion measures, and also predicts subsequent disability.⁶⁰ However, interpretation of brain atrophy results is complicated. First, in the initial period after starting anti-inflammatory therapies, loss of brain volume accelerates, presumably because of inflammation and associated edema resolves. This has been termed pseudoatrophy.⁶⁵ Pseudoatrophy has been observed in the initial year with nearly all drugs that strongly inhibit new lesion formation; an interesting exception to this was reported for fingolimod,^{66,67} which significantly reduced new lesion formation, but which slowed brain volume loss in the first year compared with placebo. Generally, however, treatment effects on brain atrophy of anti-inflammatory drugs are observed in the second year of treatment. An additional problem with brain atrophy measures is the extremely low rate of change – brain volumes decline about 0.2% per year in healthy controls, and about 0.5% to 1% per year in

Section I: Introduction

MS patients. Because the changes are very small, highly reproducible methods and studies of adequate duration are required.

Current controversies in MS trials

Relapse rate

Relapses are subjective. Because symptoms fluctuate, and are influenced by many factors – fever, high ambient temperature, anxiety, intercurrent illness, and sleep deprivation, among others – it is often not clear whether an individual MS patient has experienced a relapse or not. Also, the required duration beyond which symptoms must persist has not been standardized. Some studies use 24 hours, others 48 hours. The precise methodology to evaluate possible relapses is not standardized either. Some studies use an examining neurologist who is blinded to the treatment arm. In some cases, the examining neurologist is instructed to “not talk to the patient,” but then how does the neurologist determine some of the functional system scores, e.g. bladder / bowel? All definitions require an objective change on neurological examination, but practice is variable in allowing examining neurologists access to prior neurological examination data vs. conducting an unbiased new examination without reference to prior scores. Also, recovery is rarely quantified, and no studies require a neurological exam to establish a new baseline after an initial relapse. This makes the finding of “new neurological signs” very difficult for relapses that occur after an initial on-study relapse. For all these reasons and others, relapses remain quite subjective, and differences in relapse ascertainment almost certainly exist between examiners, sites, and studies. A recent practice has been to use an “objective” adjudication committee to review case report forms and confirm relapses. While this approach helps standardize relapse scoring, it does not make relapse assessment more consistent, still allowing potential ascertainment biases. The impact of adjudication committees to quantify relapses in MS trials has not been studied.

Another major issue with relapse as an outcome measure is its uncertain relationship to long-term clinical outcome. Natural history studies have shown only a weak relationship between relapse frequency and subsequent disability, or conversion to SPMS. While IFN β and glatiramer acetate were approved based on an approximate one-third reduction in relapse rate, no studies have documented whether a one-third reduction in relapse rate correlates with a significant benefit in later disability progression.

EDSS

The EDSS has been criticized vociferously, and for decades. Some of the more significant concerns relate to the ordinal nature of the scale, which essentially means that simple statistical analysis of EDSS change can not be done. The significance of a patient worsening from a 1.0 to a 2.0 is vastly different than for a patient worsening from a 6.0 to a 7.0, yet the difference is 1.0 EDSS units. Therefore, magnitude of EDSS change should

be avoided as a clinical trial metric. Another very significant problem is that MS patients remain at particular levels of the EDSS for variable amounts of time. Therefore, the proportion of patients at each EDSS level in the treatment arms of a clinical trial becomes relevant when a metric based on frequency of EDSS worsening is the primary outcome. This detail is not always reported. The EDSS score itself is difficult to derive, particularly at the lower end of the scale. Standardized training is important, and has been implemented in many, but not all MS trials. However, the impact of EDSS training is not clear.

Confirmed EDSS worsening is now a standard metric, but the details are still somewhat variable. Also, most studies require a ≥ 1.0 point change from baseline at EDSS levels below 5.5, and a ≥ 0.5 point change at EDSS 5.5 and above. Most studies require the EDSS change to persist at least three months. What happens when the three-month sustained change reverts back below the threshold at a subsequent visit, or at the final visit? This has been documented to occur in clinical trials,⁴⁶ presumably because of noise in the measurement tool, and because of the residual effects of relapses. The latter problem may be reduced by requiring EDSS worsening to persist for at least six months.

Another common criticism of the EDSS is that it fails to capture information about major dimensions of MS. In particular, the EDSS is relatively insensitive to visual impairment, and even more insensitive to neuropsychological impairment. Most studies indicate that EDSS correlates, even in RRMS patients, with walking ability.

As with relapse reduction, the clinical significance of confirmed EDSS worsening vis-à-vis long-term clinically meaningful disability is not clear. An analysis of disability progression in the Phase 3 IM IFN β -1a study⁴⁷ demonstrated strong correlation between disability progression in the clinical trial (defined with six-month confirmation of EDSS worsening) and clinical status eight years later. This is reassuring, and suggests that confirmed EDSS worsening in RRMS is meaningfully related to disability progression, at least as used in that particular trial.

Relevance of MRI lesions

At the time MRI was developed and applied to the study of MS patients, there was great hope that MRI visible lesions represented a window into disease pathology, and that MRI visible lesions would suffice as a sensitive, specific, and predictive imaging marker useful for patient care and clinical trials. Complexities and uncertainties soon became apparent, and continue to this day. The main problem has been the limited predictive value of MRI lesions early in the disease, and the low-modest correlations between lesions and MS-related disability. There are many potential reasons for the so-called “MRI-clinical paradox.” First, MRI-detectable lesions represent only a small portion of MS pathology. Standard MRI does not detect lesions in the gray matter, which constitutes about 65% of brain parenchyma. MRI visible lesions are restricted to white matter, and hence demonstrate pathology in about

35% of the target end-organ. Gray matter atrophy has been shown to more strongly correlate with disability than white matter atrophy.^{18,61,68} Finally, even within white matter, there are many reports of diffuse abnormalities in the tissue distant from lesions. Lesion size fluctuates over time, which introduces sampling noise when MRI scans are done infrequently, as in most clinical trials. Finally, there is almost certainly temporal dissociation between development of lesions and clinical events such as relapses or progression. This degrades the magnitude of cross-sectional correlations.

Despite these limitations, MRI lesions are included as primary outcome measures in Phase 2 proof-of-concept trials, and as important secondary outcome measures in registration trials. The effect of subcutaneous IFN β -1b on T2 bright lesions was an important consideration in the approval of subcutaneous IFN β -1b in 1993 by the USA FDA, a watershed event in the field of MS therapeutics. The recent pooled analysis showing a strong relationship between therapeutic effects on lesions and relapses strongly supports the measurement of MRI lesions in MS clinical trials.

The role of brain atrophy measures in MS trials

A National error MS Society-supported workshop on MRI measures for neuroprotection concluded that brain atrophy measures are currently a logical metric for studies of possible neuroprotective interventions.⁶⁹ However, many caveats were discussed at the meeting, and some were listed in the publication. First, brain atrophy measures are non-specific – loss of myelin and axons, gliosis, edema, and state of hydration all affect brain volumes. Second, there are various techniques used to measure brain atrophy, and results may not necessarily correlate strongly. There are very few studies in which multiple techniques were compared using the same patient sample and imaging set. Third, there are many published reports documenting pseudoatrophy – accelerated brain volume loss in the 4–6 months after initiating anti-inflammatory therapy, presumably due to resolution of edema. The kinetics of Wallerian degeneration within the central nervous system also complicate the use of brain atrophy measures in clinical trials. Assuming a treatment is instantaneously effective in stopping neuronal injury, the impact of tissue injury occurring prior to treatment will play out over an undefined length of time. Use of brain atrophy methods in clinical trials probably requires establishing a stable baseline months after starting intervention because of pseudoatrophy and the kinetics of Wallerian degeneration. The optimal trial design using brain atrophy has not been defined at present. Regional measures of brain atrophy, e.g. gray matter atrophy, cortical atrophy, or specific structures such as the thalamus, may be more useful, but development of techniques, including their validation and comparison between techniques is at an early stage. Application to multicenter trials is only beginning. Despite the caveats, brain atrophy measures are standard secondary outcomes in MS trials, and will likely play an increasingly important role.

Gray matter pathology

It is now clear that gray matter pathology is prominent in MS, and more closely relates to disability than white matter pathology, but measuring gray matter pathology using imaging techniques within clinical trials is at an early stage. There are no specific, sensitive measures of gray matter or cortical lesions, although work in this area is rapidly accelerating and there are promising methods.^{70,71} Because sensitive imaging methods are lacking, fundamental questions about gray matter pathology remain, e.g. is gray matter affected before, simultaneously, or after white matter injury? What is the mechanistic relationship between gray matter and white matter pathology?

Contemporary issues in MS trials

Optimal study designs for primary neuroprotection

There are no demonstrably effective therapies for primary neuroprotection, though it appears possible to slow the neurodegenerative process in early-stage MS with immunomodulatory or immunosuppressive drugs. Presumably, this form of neuroprotection is secondary to the anti-inflammatory effect. How can we best measure primary neuroprotection? What is the optimal patient population? What is the optimal study design? What are the optimal outcome measures? A recent trial of lamotrigine used multiple measures of CNS atrophy, and multiple clinical measures to test the hypothesis that sodium channel blockade in progressive MS would be neuroprotective.⁷¹ The results were largely negative for this pioneering study.

Declining disease severity in contemporary MS trials

As discussed in Chapter 21, there is an urgent need for more sensitive and predictive outcome measures. This need is driven by the widespread availability of partially effective therapies, which has two consequences. First, placebo-controlled studies have become controversial, supplanted in many cases by active arm comparison studies. Active arm comparison studies require many more patients. The more important impact of available disease-modifying therapy is the selection of more benign patients for clinical trials of unproven agents. Neurologists are naturally reluctant to enroll highly active patients in clinical trials where one or more arms entails an unproven therapy, opting to treat highly active patients with established treatment. This results in a selection bias in the direction of enrolling less severe patients for current clinical trials. This was thought to explain the low event rate in a recent study comparing s.c. IFN β -1a with glatiramer acetate.⁷² In addition, the move toward early diagnosis and treatment, combined with widespread use of MRI has led to the diagnosis of MS in a large number of patients who in a prior era would not have gotten a diagnosis. Some of these patients have mild MS, and some may not have MS at all.

Section I: Introduction

Personalized medicine – from trials to patient care

Results from clinical trials rarely provide insights into individual treatment response. But in an era of multiple disease-modifying drugs, it becomes desirable to select the proper drug for the proper patient at the proper time. Therefore, methods for rational selection of disease modifying drugs, and techniques to rationally monitoring treatment effectiveness are badly needed (Chapter 23). Presently, there are no validated biological markers that predict individual responsiveness to available drugs, though efforts are under way to correlate genotype, gene expression, proteomics, and pathways with effects of disease modifying drugs. There is emerging literature on the use of MRI to monitor patients treated with IFN β , for the purposes of predicting long-term benefits,^{73–75} but no current MRI markers that predict treatment response at the time therapy is initiated.

Observational and follow-up studies

Clinical trial durations of 2–3 years are feasible, notwithstanding the increasing difficulty of maintaining a placebo arm, but the impact of MS evolves over the course of a decade or longer, and rare serious adverse effects of immunomodulatory or immunosuppressive drugs may emerge only years after use of the treatment, as with natalizumab.⁷⁶ Consequently, long-term follow-up studies for efficacy and toxicity are much needed. But long-term follow-up studies cannot definitively determine causality, i.e. between intervention and outcome, because there is no concurrent comparison group.

This has lead to newer approaches, such as propensity matching to allow comparison of more similar groups.⁷⁷ There have been published long-term follow-up studies for all the IFN β products,^{78–80} and for glatiramer acetate.⁸¹ These studies have suggested a beneficial effect of disease-modifying drug therapy in delaying the onset of SPMS, or in lowering progression to EDSS milestones. As noted,⁸² causal inferences about the therapeutic effectiveness of disease-modifying drugs in long-term follow-up studies are difficult due to design constraints. Long-term follow-up studies are very useful for determining long-term tolerability and emergence of rare adverse effects.

Post-approval monitoring

As discussed in Chapter 18, there has been an expanded emphasis on adverse event monitoring, and risk minimization procedures for approved MS drugs. The natalizumab progressive multifocal leukoencephalopathy experience, and occurrence of leukemia and cardiotoxicity with mitoxantrone have ushered in a new era in MS therapeutics, with more effective but more toxic therapies. This has occurred at a time of heightened awareness of risk with marketed pharmaceuticals (e.g. rofecoxib;⁸³ rosiglitazone.⁸⁴) Clearly, efficacy needs to be balanced with risk, and risk needs to be defined over the long duration of MS. Availability of more effective, riskier drugs further drives the need for accurate prognostic markers for disease severity, and rational methods to personalize the use of disease-modifying drugs.

References

1. Paty DW, Li DK, UBC MS/MRI Study Group, The IFNB Multiple Sclerosis Study Group. Interferon beta-1b is effective in relapsing–remitting multiple sclerosis. II. MRI analysis results of a multicenter, randomized, double-blind, placebo-controlled trial. *Neurology* 1993;43:662–7.

2. McFarland HF, Frank JA, Albert PS, *et al.* Using gadolinium-enhanced magnetic resonance imaging lesions to monitor disease activity in multiple sclerosis. *Ann Neurol* 1992;32(6): 758–66.

3. Miller DH, Rudge P, Johnson G, *et al.* Serial gadolinium enhanced magnetic resonance imaging in multiple sclerosis. *Brain* 1988;111(4):927–39.

4. Katz D, Taubenberger JK, Cannella B, McFarlin DE, Raine CS, McFarland HF. Correlation between magnetic resonance imaging findings and lesion development in chronic, active multiple sclerosis. *Ann Neurol* 1993;34(5): 661–9.

5. Bruck W, Bitsch A, Kolenda H, Bruck Y, Stiefel M, Lassmann H. Inflammatory central nervous system demyelination: correlation of magnetic resonance imaging findings with lesion pathology. *Ann Neurol* 1997;42(5): 783–93.

6. Harris JO, Frank JA, Patronas N, McFarlin DE, McFarland HF. Serial gadolinium-enhanced magnetic resonance imaging scans in patients with early, relapsing–remitting multiple sclerosis: implications for clinical trials and natural history. *Ann Neurol* 1991;29(5):548–55.

7. Thompson AJ, Miller D, Youl B, *et al.* Serial gadolinium-enhanced MRI in relapsing/remitting multiple sclerosis of varying disease duration. *Neurology* 1992;42(1):60–3.

8. Sormani MP, Bonzano L, Roccatagliata L, Cutter GR, Mancardi GL, Bruzzi P. Magnetic resonance imaging as a potential surrogate for relapses in multiple sclerosis: a meta-analytic approach. *Ann Neurol* 2009;65(3): 268–75.

9. Ge Y, Gonen O, Inglese M, Babb JS, Markowitz CE, Grossman RI. Neuronal cell injury precedes brain atrophy in multiple sclerosis. *Neurology* 2004;62(4):624–7.

10. Gallo A, Rovaris M, Riva R, *et al.* Diffusion-tensor magnetic resonance imaging detects normal-appearing white matter damage unrelated to short-term disease activity in patients at the earliest clinical stage of multiple sclerosis. *Arch Neurol* 2005;62(5): 803–8.

11. Rovaris M, Gambini A, Gallo A, *et al.* Axonal injury in early multiple sclerosis is irreversible and independent of the short-term disease evolution. *Neurology* 2005;65(10):1626–30.

12. Van Den Elskamp IJ, Lembecke J, Dattola V, *et al.* Persistent T1 hypointensity as an MRI marker for treatment efficacy in multiple sclerosis. *Mult Scler* 2008;14(6):764–9.

Chapter 1: Challenges in experimental therapies

13. Simon JH, Jacobs LD, Campion MK, *et al.* A longitudinal study of brain atrophy in relapsing multiple sclerosis. The Multiple Sclerosis Collaborative Research Group (MSCRG). *Neurology* 1999;53(1):139–48.
14. Rudick RA, Fisher E, Lee JC, Simon J, Jacobs L. Use of the brain parenchymal fraction to measure whole brain atrophy in relapsing–remitting MS. Multiple Sclerosis Collaborative Research Group. *Neurology* 1999;53(8):1698–704.
15. Brex PA, Jenkins R, Fox NC, *et al.* Detection of ventricular enlargement in patients at the earliest clinical stage of MS. *Neurology* 2000;54(8):1689–91.
16. Chard DT, Griffin CM, Parker GJ, Kapoor R, Thompson AJ, Miller DH. Brain atrophy in clinically early relapsing–remitting multiple sclerosis. *Brain* 2002;125(2):327–37.
17. Geurts JJ, Bo L, Pouwels PJ, Castelijns JA, Polman CH, Barkhof F. Cortical lesions in multiple sclerosis: combined postmortem MR imaging and histopathology. *Am J Neuroradiol* 2005;26(3):572–7.
18. Fisher E, Lee JC, Nakamura K, Rudick RA. Gray matter atrophy in multiple sclerosis: a longitudinal study. *Ann Neurol* 2008;64:255–265.
19. Chard DT, Griffin CM, Rashid W, *et al.* Progressive grey matter atrophy in clinically early relapsing–remitting multiple sclerosis. *Mult Scler* 2004;10(4):387–91.
20. Costello F, Coupland S, Hodge W, *et al.* Quantifying axonal loss after optic neuritis with optical coherence tomography. *Ann Neurol* 2006;59(6):963–9.
21. Lublin FD, Reingold SC. Defining the clinical course of multiple sclerosis: results of an international survey. *Neurology* 1996;46, 907–11.
22. Polman CH, Reingold SC, Edan G, *et al.* Diagnostic criteria for multiple sclerosis: 2005 revisions to the “McDonald Criteria.” *Ann Neurol* 2005;58(6):840–6.
23. Marrie RA, Cutter G, Tyry T, Hadjimichael O, Campagnolo D, Vollmer T. Changes in the ascertainment of multiple sclerosis. *Neurology* 2005;65(7):1066–70.
24. Thompson AJ, Kermode AG, Wicks D, *et al.* Major differences in the dynamics of primary and secondary progressive multiple sclerosis. *Ann Neurol* 1991;29(1):53–62.
25. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 1983;33(11):1444–52.
26. Bagnato F, Butman JA, Gupta S, *et al.* In vivo detection of cortical plaques by MR imaging in patients with multiple sclerosis. *Am J Neuroradiol* 2006;27(10):2161–7.
27. Weinshenker B. The natural history of multiple sclerosis. *Neurol Clin* 1995;13:119–46.
28. Filippi M, Horsfield MA, Morrissey SP, *et al.* Quantitative brain MRI lesion load predicts the course of clinically isolated syndromes suggestive of multiple sclerosis. *Neurology* 1994;44: 635–641.
29. Fisniku LK, Chard DT, Jackson JS, *et al.* Gray matter atrophy is related to long-term disability in multiple sclerosis. *Ann Neurol* 2008;64(3):247–54.
30. Held U, Heigenhauser L, Shang C, Kappos L, Polman C. Predictors of relapse rate in MS clinical trials. *Neurology* 2005;65(11):1769–73.
31. Barkhof F, Held U, Simon JH, *et al.* Predicting gadolinium enhancement status in MS patients eligible for randomized clinical trials. *Neurology* 2005;65(9):1447–54.
32. Simon JH, Jacobs LD, Campion M, *et al.* Magnetic resonance studies of intramuscular interferon beta-1a for relapsing multiple sclerosis. The Multiple Sclerosis Collaborative Research Group. *Ann Neurol* 1998;43(1):79–87.
33. Lucchinetti C, Bruck W, Rodriguez M, Lassmann H. Distinct pattern of multiple sclerosis pathology indicates heterogeneity in pathogenesis. *Brain Pathol* 1996;6(3):259.
34. Lucchinetti C, Bruck W, Parisi J, Scheithauer B, Rodriguez M, Lassmann H. Heterogeneity of multiple sclerosis lesions: implications for the pathogenesis of demyelination. *Ann Neurol* 2000;47(6):707–17.
35. Lassmann H, Bruck W, Lucchinetti CF. The immunopathology of multiple sclerosis: an overview. *Brain Pathol* 2007;17(2):210–18.
36. Lucchinetti CF, Bruck W, Lassmann H. Evidence for pathogenic heterogeneity in multiple sclerosis. *Ann Neurol* 2004;56(2):308.
37. Lucchinetti CF, Mandler RN, McGavern D, *et al.* A role for humoral mechanisms in the pathogenesis of Devic’s neuromyelitis optica. *Brain* 2002;125(7):1450–61.
38. Weinshenker BG, Wingerchuk DM, Pittock SJ, Lucchinetti CF, Lennon VA. NMO-IgG: a specific biomarker for neuromyelitis optica. *Dis Markers* 2006;22(4):197–206.
39. Lennon VA, Wingerchuk DM, Kryzer TJ, *et al.* A serum autoantibody marker of neuromyelitis optica: distinction from multiple sclerosis. *Lancet* 2004;364(9451):2106–12.
40. PRISMS Study Group. Randomised double-blind placebo-controlled study of interferon beta-1a in relapsing/ remitting multiple sclerosis. PRISMS (Prevention of Relapses and Disability by Interferon beta-1a Subcutaneously in Multiple Sclerosis) Study Group [see comments]. *Lancet* 1998;352(9139):1498–504.
41. The IFN- β Multiple Sclerosis Study Group. Interferon beta-1b is effective in relapsing–remitting multiple sclerosis. I. Clinical results of a multicenter, randomized, double-blind, placebo-controlled trial. *Neurology* 1993;43:656–61.
42. Johnson KP, Brooks BR, Cohen JA, *et al.* Copolymer 1 reduces relapse rate and improves disability in relapsing–remitting multiple sclerosis: results of a phase III multicenter, double-blind placebo-controlled trial. *Neurology* 1995;45(7):1268–76.
43. Rudick RA, Stuart WH, Calabresi PA, *et al.* Natalizumab plus interferon beta-1a for relapsing multiple sclerosis. *N Engl J Med* 2006;354(9):911–23.
44. Polman CH, O’Connor PW, Havrdova E, *et al.* A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. *N Engl J Med* 2006;354(9):899–910.
45. Kremenetzky M, Cottrell D, Rice G, *et al.* The natural history of multiple sclerosis: a geographically based study. 7. Progressive-relapsing and relapsing-progressive multiple sclerosis: a re-evaluation. *Brain* 1999;122 (10):1941–50.
46. Ebers GC, Heigenhauser L, Daumer M, Lederer C, Noseworthy JH. Disability as an outcome in MS clinical trials. *Neurology* 2008;71(9):624–31.

Section I: Introduction

47. Rudick RA, Lee JC, Cutter GR, *et al.* Disability progression in a clinical trial of relapsing–remitting multiple sclerosis: eight-year follow-up. *Arch Neurol* 2010;87:1329–35.

48. Sormani MP, Bonzano L, Roccatagliata L, Mancardi GL, Uccelli A, Bruzzi P. Surrogate endpoints for EDSS worsening in multiple sclerosis. A meta-analytic approach. *Neurology* 2010;75(4):302–9.

49. Rudick R, Antel J, Confavreux C, *et al.* Recommendations from the National Multiple Sclerosis Society Clinical Outcomes Assessment Task Force. *Ann Neurol* 1997;42(3):379–82.

50. Fox RJ, Lee JC, Rudick RA. Optimal reference population for the multiple sclerosis functional composite. *Mult Scler* 2007;13(7):909–14.

51. Rudick RA, Polman CH, Cohen JA, *et al.* Assessing disability progression with the multiple sclerosis functional composite. *Mult Scler* 2009;15(8):984–97.

52. Rao SM, Leo GJ, Bernardin L, Unverzagt F. Cognitive dysfunction in multiple sclerosis. I. Frequency, patterns, and prediction. *Neurology* 1991;41(5):685–91.

53. Rao SM, Leo GJ, Ellington L, Nauertz T, Bernardin L, Unverzagt F. Cognitive dysfunction in multiple sclerosis. II. Impact on employment and social functioning. *Neurology* 1991;41(5):692–6.

54. Cohen JA, Rudick RA (eds.) *Multiple Sclerosis Therapeutics*. 3rd ed. United Kingdom: Informa Healthcare, 2007.

55. Benedict RH, Cookfair D, Gavett R, *et al.* Validity of the minimal assessment of cognitive function in multiple sclerosis (MACFIMS). *J Int Neuropsychol Soc* 2006;12(4):549–58.

56. Rudick RA, Miller DM. Health-related quality of life in multiple sclerosis: current evidence, measurement and effects of disease severity and treatment. *CNS Drugs* 2008;22(10):827–39.

57. Rudick RA, Miller D, Hass S, *et al.* Health-related quality of life in multiple sclerosis: effects of natalizumab. *Ann Neurol* 2007;62(4):335–46.

58. Barkhof F. MRI in multiple sclerosis: correlation with expanded disability status scale (EDSS). *Mult Scler* 1999;5(4):283–6.

59. Fisher E, Chang A, Fox RJ, *et al.* Imaging correlates of axonal swelling in chronic multiple sclerosis brains. *Ann Neurol* 2007;62(3):219–28.

60. Fisher E, Rudick RA, Cutter G, *et al.* Relationship between brain atrophy and disability: an 8-year follow-up study of multiple sclerosis patients. *Mult Scler* 2000;6(6):373–7.

61. Fisniku LK, Chard DT, Jackson JS, *et al.* Gray matter atrophy is related to long-term disability in multiple sclerosis. *Ann Neurol* 2008;64(3):247–54.

62. van Walderveen MA, Kamphorst W, Scheltens P, *et al.* Histopathologic correlate of hypointense lesions on T1-weighted spin-echo MRI in multiple sclerosis. *Neurology* 1998;50(5):1282–8.

63. van de I, Lembecke J, Dattola V, *et al.* Persistent T1 hypointensity as an MRI marker for treatment efficacy in multiple sclerosis. *Mult Scler* 2008;14(6):764–9.

64. Smith SM, Zhang Y, Jenkinson M, *et al.* Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *Neuroimage* 2002;17(1):479–89.

65. Zivadinov R, Reder AT, Filippi M, *et al.* Mechanisms of action of disease-modifying agents and brain volume changes in multiple sclerosis. *Neurology* 2008;71(2):136–44.

66. Cohen JA, Barkhof F, Comi G, *et al.* Oral fingolimod or intramuscular interferon for relapsing multiple sclerosis. *N Engl J Med* 2010;362(5):402–15.

67. Kappos L, Radue EW, O’Connor P, *et al.* A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis. *N Engl J Med* 2010;362(5):387–401.

68. Rudick RA, Nakamura K, Lee J-C, Fisher E. Gray matter atrophy correlates with MS disability progression measured with MSFC but not EDSS. *J Neurol Sci* 2009;282:106–11.

69. Barkhof F, Calabresi PA, Miller DH, Reingold SC. Imaging outcomes for neuroprotection and repair in multiple sclerosis trials. *Nat Rev Neurol* 2009;5(5):256–66.

70. Calabrese M, Filippi M, Gallo P. Cortical lesions in multiple sclerosis. *Nat Rev Neurol* 2010;6(8):438–44.

71. Chard DT, Miller DH. What you see depends on how you look: Gray matter lesions in multiple sclerosis. *Neurology* 2009;73(12):918–19.

72. Mikol DD, Barkhof F, Chang P, *et al.* Comparison of subcutaneous interferon beta-1a with glatiramer acetate in patients with relapsing multiple sclerosis (the REBif vs Glatiramer Acetate in Relapsing MS Disease [REGARD] study): a multicentre, randomised, parallel, open-label trial. *Lancet Neurol* 2008;7(10):903–14.

73. Rudick RA, Lee JC, Simon J, Ransohoff RM, Fisher E. Defining interferon beta response status in multiple sclerosis patients. *Ann Neurol* 2004;56(4):548–55.

74. Durelli L, Barbero P, Bergui M, *et al.* MRI activity and neutralising antibody as predictors of response to interferon beta treatment in multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2008;79(6):646–51.

75. Prosperini L, Gallo V, Petsas N, Borriello G, Pozzilli C. One-year MRI scan predicts clinical response to interferon beta in multiple sclerosis. *Eur J Neurol* 2009;16(11):1202–9.

76. Yousry TA, Major EO, Ryschkewitsch C, *et al.* Evaluation of patients treated with natalizumab for progressive multifocal leukoencephalopathy. *N Engl J Med* 2006;354(9):924–33.

77. Trojano M, Pellegrini F, Fuiani A, *et al.* New natural history of interferon-beta-treated relapsing multiple sclerosis. *Ann Neurol* 2007;61(4):300–6.

78. Rudick RA, Cutter GR, Baier M, *et al.* Estimating long-term effects of disease-modifying drug therapy in multiple sclerosis patients. *Mult Scler* 2005;11(6):626–34.

79. Ebers GC, Traboulsee A, Li D, *et al.* Analysis of clinical outcomes according to original treatment groups 16 years after the pivotal IFNB-1b trial. *J Neurol Neurosurg Psychiatry* 2010;81(8):907–12.

80. Kappos L, Traboulsee A, Constantinescu C, *et al.* Long-term subcutaneous interferon beta-1a therapy in patients with relapsing–remitting MS. *Neurology* 2006;67(6):944–53.